Language, literacy and technology: embodied peer-interaction and collaborative writing in an ESOL classroom

Stephen Woulds
MSc MA BA PGCE(FE)

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of Leeds
School of Education
September 2015

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Thank you

To my supervisors Dr James Simpson and Dr Aisha Walker.

To my wife and children: Carol, Stephanie, Jake, Emily and William.

To Teresa and Reza for their translations.

To my ESOL colleagues at Leeds City College who continue to brave the coal-face of teaching in turbulent times.

To my parents William and Elaine, and sisters Angela and Melanie.

Thank you dad for the years of wisdom and support.

Abstract

This study focuses on digital literacies, and real-time multimodal design, within the context of migrant adult learners in the UK. It seeks to understand the frameworks of peer-interaction when second-language learners are paired at a computer and how they negotiate second-language writing. In this research, pairs of students sharing the same language were tasked with an environmental project which included the digital design of an image, designing a four-page booklet using Publisher, a website and to produce all of these using English as a second language. The process was videoed across a three-hour classroom session with four pairs of learners: Kurdish, Polish, French and Arabic. New literacies, embodied peer-interaction and second-language writing are the primary fields informing this research.

The outcomes of the research are: (1) a methodology is developed for the collection and analysis of multimodal data when learners collaborate at a computer; (2) the field of new literacies is extended through an analysis of the *design-process*, as opposed to *product-analysis*; (3) a peer-interaction framework is presented which broadens our understanding of classroom interaction, including linguistic, paralinguistic and mediating resources when learners share technology; (4) the field of second-language writing is extended through an analysis of peer-writing with technology.

The research concludes with a peer-interaction framework comprised of learner alignment and misalignment across language, literacy and technology. Spoken and written language goes through an iterative cycle of transformation. The central finding from the research is the naming and defining of *transmodal talk* within a peer-interaction framework. The sequentiality of this process has common features across all the pairs of learners. *Transmodal talk* is presented to identify the fluid process of transposing off-screen dialogue to on-screen text. They both shape and mediate each other through temporal mapping and polyvocality.

Table of Contents

List of tables

List of figures

List of transcripts

List of multimodal extracts

List of abbreviations

ACMC	asynchronous computer-mediated communication
CEFR	Common European Framework of Reference for Languages
CMC	computer-mediated communication
CW	collaborative writing
DfEE	Department for Education and Employment
DfES	Department for Education and Skills
ESOL	English for Speakers of Other Languages
FELTAG	Further Education Learning Technology Action Group
F2FCCW	face-to-face collaborative computer writing
HLA	higher-level actions
IELTS	International English Language Testing System
L1	first language
L2	second language
LLA	lower-level actions
LLT	Language, Literacy and Technology
LRE	language-related episodes
LSA	learning support assistant – this is someone employed by the organisation to support learners on a one-to-one basis as required, often circulating the classroom and verbally checking on learners
MDA	multimodal discourse analysis
ROLSI	Research on Language and Social Interaction
SCMC	synchronous computer-mediated communication
SFA	Skills Funding Agency
SLA	second-language acquisition
ZPD	zone of proximal development

1 Introduction

1.1 explaining the purpose

This study explores embodied peer-interaction, and collaborative writing, within the context of migrant adult learners in the UK. It seeks to understand how second-language learners collaborate when paired at a computer and negotiate the production of multimodal texts. This is important to understand because 'what constitutes English is not to be found in language alone, but exists in many modes, and in many tasks' (Kress et al, 2005: 2). In this research, same-language learners were tasked with an environmental project which included the digital design of an image, designing a four-page booklet using Publisher, a website and to produce all of these collaboratively using English as a second language. The process was videoed across a three hour classroom session with four pairs of learners: Kurdish, Polish, French and Arabic. The research works towards an understanding of embodied peer-interaction within the context of face-to-face collaborative computer writing (F2FCCW).

Technology has enabled new pedagogies with a convergence of modes and a profusion of hybrid texts. It is widely acknowledged that there has been a paradigm shift from page to screen, from pen to keyboard, from words to visuals, from consumer to producer and more (Jewitt, 2006). One outcome of this paradigm shift in digitised, textual communication is the repositioning of literacy:

> Conceptions of literacy as a singular canonical English that exclusively concerns linguistics or alphabetic print are no longer sufficient in an increasingly multimodal and digitally-mediated world of textual design. In particular, cultural differences and a proliferation of communication media provide impetus for a pedagogy of multiliteracies (Mills, 2006: 13).

Many theorists and educators have responded to this multimodal, digital landscape, including Warschauer (1999); Cope & Kalantzis (2000); Hawisher & Selfe (2000); Unsworth (2001); Snyder (2002); Martin (2006); Jewitt (2006); Kist (2005); Burn & Durran (2007); Baynham & Prinsloo (2009); Kress (2010); Goodfellow & Lea (2013). The tension between new literacy studies and learning technologies is considered by Gourlay, Hamilton & Lea (2013) who discuss the implications and future directions for researchers working in this field. This research seeks to understand how migrant learners of English collaborate at a computer to produce multimodal texts in real-time and what resources and modes they utilise to help them in the process. In this sense the

research belongs to what has been termed '*third generation* empirical work which is pushing the boundaries of literacy research in a number of key directions: from the local to the translocal, from print based literacies to electronic and multimedia literacies and from the verbal to the multimodal' (Baynham & Prinsloo, 2009: 1). A research focus on second-language writing as an outcome of embodied peer-interaction also belongs to a body of work which Nevile (2015) calls the 'embodied turn' in multimodal research; discussed in detail in section 2.7.

1.2 researcher context

This research is situated in the Further Education (FE) sector and belongs to what could be termed practitioner-researcher, a branch of research in which the educator investigates their own teaching practice. Ecclesfield observes the paucity of FE based research compared to Higher Education (HE) and notes that 'it seems essential to encourage and promote practitioner research both into the sector to support exploration and to change and find new ways of helping practitioners to turn their experiences into learning for themselves, their learners and their sector colleagues' (2013: 10). Procter-Legg similarly identifies the *insider* benefits and 'implicit knowledge that those working in FE can bring to research' (2013: 11). At the same time, the FE sector seeks to professionalise the use of technology for teaching and learning. FELTAG (Further Education Learning Technology Action Group) for example makes a number of institutional and training recommendations to improve standards in the use of learning technologies. The FE workforce is at the core of this, calling for teachers' 'continuing professional development so that their ability to understand and optimise the use of learning technology can be enhanced and refreshed regularly' (2013: 4). Grounded research in one's own practice can generate *emic* ways of seeing and thinking about teaching and learning, including the use of digital technologies. Section 3.2 provides more contextual detail on the locality and demographics of this study as well as the ethnomethodological approach to practitioner research.

I have taught language and literacy in the FE sector, using technology, for twenty years. I have therefore experiential knowledge of how literacy has proliferated through digital media in the last two decades. I completed an MA in English language and later an MSc in educational technology. I was at the time teaching English for Speakers of Other Languages (ESOL) when the ratio of students to computers was rarely 1:1. Through these post-graduate qualifications I applied the research in these courses to my own

teaching practice. As researcher-practitioner I began to realise there were significant gaps in my understanding of how learners were collaborating using language and technology. I could also see my colleagues pairing learners at a computer. But in the real-time synchronicity of classroom learning it is impossible for a tutor to sit and observe the totality of learner-collaboration; even more so when there can be several pairings all happening at the same time.

As teachers we might use peer-collaboration (with or without a computer) in the best interests of the learners. It is a pedagogy assumed to be beneficial. Bruner (1996) discussed *scaffolding* to describe the process of an individual supporting another individual to achieve a learning goal; this could be the tutor but could equally be another student. Vygotsky talked in similar terms about the *zone of proximal development* to describe 'the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers' (Vygotsky, 1978: 86). I began to ask myself a range of questions: is peer-interaction as pedagogy always beneficial? How does scaffolding actually work in second-language contexts, if at all? What is an action in interaction? Are there inherited ways of thinking? What theories of learning are we taking for granted?

A PhD seemed ideal to formalise a research path to begin to answer some of these questions in relation to my own teaching practice, and beyond that, to other practitioners using the same pedagogy. This research journey was triggered because of two problems I was seeing in my own teaching: (1.) literacy was not a singular entity as I had been led to understand but multiple, digital and multimodal (2.) the participation frameworks for learner collaboration were far more complex than I realised because language was only one mode in a sophisticated repertoire of multimodal communication. Shortly after I began the PhD I became an ESOL Curriculum Manager at Leeds City College and then Head of Department. With 1200 learners and twenty-eight teachers I began to see on a daily basis the pedagogy of second-language collaboration and sometimes the use of computer-sharing in that pedagogy. Research into my own teaching practice could also have relevance to my peers.

1.3 learner context

ESOL is a term used to define a diverse collection of learners from different backgrounds and with different languages; typically, the learner is living in the country

of the dominant language and is in the minority. Based on the findings of the ESOL Effective Practice Project: 'two thirds of ESOL students are women, half are under thirty, about one in seven cannot read or write in L1 while one in eight have a tertiary level education. And the vast majority are not currently working. Many migrants to English dominant countries do not already have competence in English when they arrive' (Simpson, 2007: 2). The ESOL learners in the present research enrol on a thirty-five week Further Education (FE) college course, working towards ESOL and IT qualifications. This type of multicultural and multilingual classroom is increasing in the FE sector and mirrors a global pattern of migration and education: 'globalisation and patterns of mass forced and voluntary migration have resulted … in large numbers of migrants coming to the UK' (Simpson, 2007: 1). A fuller demographic and biographic discussion is provided at section 3.2.

The classroom in this study is an instance of literacy practice which is populated by multicultural, multilingual learners. Languages, literacies and technologies fuse at a local level within this classroom. That localised fusion is reflective of a more global pattern. Similar instance of new literacies are commonplace in the US, Canada and Australia, where dominant English competes with vernacular and digital literacies in traditional and virtual spaces for learning; Kist (2005); Mills (2006). A recent study is Bhatt's (2014) PhD thesis. Also based in the FE sector and with ESOL learners in a digital context, Bhatt argues: 'By understanding learner practices it is possible to better understand digital innovations in education, the extent to which learners embrace or avoid imposed technologies, and how such practices re-shape assignments as evolving pedagogic forms' (2014: 4). This approach towards learner use of technology begins to answer recent concerns in the FE sector, that 'research and conversations consistently referred to the under-exploitation of learners' skills, devices and technical knowledge when it came to the use of learning technology' (FELTAG, 2013: 5).

A classroom such as this in the UK is a place where teachers and students struggle to 'survive and thrive in a world that increasingly puts emphasis on one's ability to work in multiple forms of representation across a wide variety of spaces, all text-driven in one way or another' (Kist, 2005:13). The classroom can put at the disposal of students a range of meaning-making resources, all text-mediated and often digitally-mediated. Multiple modes for communicating in and outside the classroom include traditional literacy repertoires (note-taking, essay writing, reading and producing different genres)

as well as newer types of multimodal literacies mediated by technology: web design, images and leaflet design. Such text-mediated resources (many of which include images) have been collectively named *multiliteracies*, discussed in chapter 3 (Cope & Kalantzis, 2000). It is the latter which is the focus of this research, how ESOL learners engage with the creation of digital literacies in the classroom.

1.4 research rationale

Central to this research is how to define, capture and analyse the sequential actions which collectively and accumulatively enable (and evidence) learning in real-time when using language, literacy and technology. By *define* I mean delineate interaction to arrive at a fuller understanding of what an action is, and how it is mediated by, different timescales of interaction, communicative modes and resources available to learners. This calls for a reassessment of classroom peer-interaction as part of the *third generation* approach to empirical research in literacy studies (Baynham & Prinsloo, 2009). By *capture* I mean how we might gather the multimodal video data we need to observe actions and interactions in real-time; including synchronous multiple streams of audio and video for analytical purposes. By *analysis* I mean how we might analyse and represent embodied peer-interaction, transposing observed phenomena into multimodal transcripts. Ultimately, the research seeks to explicate the participation frameworks of learning, language and interaction when learners collaborate in real-time using technology. Capturing and analysing these 'real-time' interactions to evidence learning is not without significant problems; discussed in chapters 3 and 4. The specific research questions are: (1.) *How are off-screen talk and on-screen text coordinated? (2.) How are peer-interaction frameworks for learning structured?* Chapter 2 will evidence the framing of these questions as contributions to the fields of multimodal interactional analysis and collaborative second-language writing.

1.5 structure of the thesis

Chapter 1 provides a rationale for the PhD and explains that the research grew out of classroom practice, driven by a desire to understand the real-time multimodal collaboration of second-language learners as they design digital texts. It was discussed that the current research will contribute to the capturing, looking at and thinking about such interaction.

In chapter 2 the relevant literature is reviewed. It begins with a consideration of how others have defined 'literacy' and finds competing views and ideologies. Theorists might favour one model, such as *critical literacy* (Freire 1986) or *situated literacy* (Street 1998) but these do not sit easily with the government funded *functional literacy* (DfEE 2001) and the similar focus on mechanical skills acquisition of *autonomous literacy (*Olson 1994). The chapter then moves on to consider more recent challenges to literacy from a technological perspective. Multimodality is problematised to clarify the focus in this research as one of multimodal interactional analysis rather that multimodal text analysis. Consideration is then given to what is meant by multimodal. The field of collaborative second-language writing is discussed.

Chapter 3 begins with the research methodology and identifies the research participants and the educational context. The research approach is identified as an *ethnomethodological case study*. A qualitative orientation is explained with video observation and interviewing as the primary methods of data-collection. The methods used to trial data collection in the piloting phase are considered and the epistemological and ontological changes that were an outcome of this experience. Audio-visual tools and methods were developed to collect the data from one classroom session and a rationalisation is provided to explain why the research needed extensively rich and detailed data. The videography (Knoblauch, 2011) of how to capture and manage video data is considered.

Chapter 4 describes the micro-analytical approach used to analyse the video data of the interaction, applied later in chapters 5 and 6. An ethnomethodological justification is given to the observation of videoed interaction; i.e. a grounded approach to understand the patterns and organisation of peer-interaction. A methodological framework for analysing the learners' interaction (Norris, 2004) is discussed which involves considering scales of time and analysable units (Lemke, 2009). Consideration is given to what is meant by an action, where it starts, ends and bleeds into other actions. Analysing the moment-to-moment interaction of learners to understand the sequentiality of collaboration is one reason why the totality of the lesson was recorded and transcribed. The structuring of lower and higher-level actions in Norris' methodological framework (2004) is used to enable this. Attention is then given to the preparation of multimodal data for transcription and analysis followed by a discussion on which focal events, and which participants, were chosen for detailed analysis.

Chapter 5 is the first analysis chapter. A ten-minute interaction of two female Polish learners is chosen as they write two sentences in Microsoft Publisher. Using detailed, second-by-second microanalysis, the research identifies the real-time modal coordination of screen-based, collaborative writing as the learners work together. Their interaction is positively structured with significant findings on the features of their multimodal collaboration and how off-screen dialogue is transposed to on-screen text.

Chapter 6 is the second analysis chapter. A similar ten-minute interaction of two male Kurdish learners is chosen as they write three sentences in Microsoft Publisher. The same micro-analytical approach is used, combining multimodal and conversation analysis. Their interaction has a contested structure with features both similar and different to the Polish learners at chapter 5. The focal points chosen in both chapters provide a unique insight into the real-time design process of on-screen writing where language, literacy and technology come together. Using the similar focal event in both analysis chapters offers interesting contrast between the Kurdish learners and the Polish learners.

Chapter 7 summarises the findings from the previous two analysis chapters with comparisons and contrasts made between the two pairs of learners. The term *transmodal talk* is identified to explain the coordination of off-screen talk and on-screen text. There is a sequential pattern in how learners transpose talk to text. Spoken language between learners, as a drafting process prior to writing, undergoes significant mutation across phoneme, lexeme and lexical chunking as learners attempt to map speech to on-screen graphemes. The sequencing of that process is structured as cognitive orientation, off-screen drafting, on-screen writing, off-screen noticing, on-screen correcting; discussed further in section 8-2.

Chapter 8 is the conclusion chapter which offers a final summary of the key findings. The two research questions are answered. Section 8.2 collates the findings on transmodal talk and provides a diagram and terminology to answer the first research question on the coordination of talk and text. The identified features and terminology provided will support other educators and researchers wishing to explore face-to-face computer collaborative writing F2FCCW in similar contexts. 8.3 collates the findings on the peer-interaction framework (PiF) and provides a PiF diagram and terminology to

answer the second research question on how peer-interaction frameworks are structured. An inclusionary and exclusionary framework is identified and terminology provided for other educators and researchers wishing to explore PiF in similar contexts. The limitations of the study are considered, its implications for education and research, contributions and further areas for research.

1.6 key terms

Table 1-1 lists and defines the key terms and ideas used in the thesis. This defining does simplify and hence is minimalist, however, problematic and additional terms are more fully explained as they appear in the thesis.

terms	definitions	source
actions	a much theorised term but defined here in an educational context as behaviour which has purpose, intentionality; there is agency and cognitive engagement with goal-orientated outcomes realised through physical means	Norris (2004)
distributed cognition	the extent to which knowledge is shared through interaction with people and semiotic resources	Atkinson (2010)
embodied cognition	the extent to which knowledge is structured by physical interaction with the world	Atkinson (2010)
embodied	sensory engagement in the immediate environment through physical interactions with objects and people	Goodwin (2000a)
higher-level actions	actions with an identifiable beginning and end towards achieving goal-orientated outcomes	Norris (2004)
lower-level actions	chains of smaller actions which are fluidly performed to achieve the higher-level actions; utterances, gestures, etc	Norris (2004)
materiality	the physical dimensions of modes with different affordances; e.g. print has enduring materiality compared to speech	Björkvall & Karlsson (2011)
multimodal literacies	emphasis on how text interplays with the visual and other modes in print and digital media	Lankshear & Knobel (2006)
multiliteracies	classroom practice advocating a pedagogy of Learning by Design including digital literacies and the multimodal	Cope et al (2000)
mode	a loose concept denoting a grouping of signs, such as the visual mode, linguistic mode, etc	Kress (2010)
modal configuration	also termed multimodal ensemble; the relationship of several modes as they interplay in interaction	Norris (2004)
multimodal interactional analysis	identifying how different modes are structured in social interaction between individuals	Norris (2004)

multimodal text analysis	identifying how different modes are structured in digital and printed media	Kress & van Leeuwen (2001)
modal affordance	the potential benefits of a mode in representation and communication	Oliver (2005)
modal constraint	the limitations of a mode in representation and communication	Kress (2010)
new literacies	emphasis on literacy in digital media	Kist (2005)
scales of time	the temporal duration between an action's beginning and ending	Lemke (2009)
semiotic resources	social, material and cultural resources to make meaning, including language, visuals, gesture, etc	Kress (2010)
sociomaterial	the interconnection of people and artefacts in networks of activity	Fenwick et al (2011)
transmodal	moving the same meaning from one mode to another; e.g. words as images and vice versa	Newfield (2013)

Table 1-1 (key terms and definitions)

2 Literature review

2.1 introduction

This chapter will contextualise the research by considering literature across a number of related fields: new literacies, multimodality, second-language writing and peer-interaction. The discussion will move from the broad to the narrow, considering first the wider literature on literacy and technology, before narrowing the focus to the classroom research of this study: which is peer-interaction and writing as an embodied, cognitive activity within the context of second-language learners in a computer sharing setting. A research concern is with learner-related contexts and where to draw the boundaries of interest and relevance. A context-based framework for research seeks to place learners and learning at the centre of the focus because 'it is impossible to understand how people work or learn without taking into account the people and artefacts that are part of the completion of their work or learning' (Luckin, 2010: 3). A similar sociomaterial approach is taken here.

As practitioner-based research, the current study grew out of watching my own learners and my curiosity with the complexity of how they were collaborating at a computer to design multimodal texts. Understanding this requires close observation of 'people and artefacts' to identify how they are completing their learning. Therefore, the learners and their negotiation/coordination of learning and the role of the environment/artefacts are considered to be most central. In this research there is a resistance to shoehorning the students into an established field of discipline which might be a poor fit; for example, Human Computer Interaction (HCI); Computer Mediated Communication (CMC); Mediated Discourse Analysis (MDA); Activity Theory (AT). These are all feasible disciplines of study for this research and yet each of those might be considered a top-down approach which is not wholly relevant. Instead, the approach is to look at the learners and what is most relevant to understand how they are learning, and in a broadly ethnomethodological approach to identify social order, relate the literature where appropriate so that theory is grounded in the data. To this end, the literature review was significantly rewritten after a preliminary analysis of the data because it was then when the most relevant contexts became apparent.

Multiple contexts, however, run the risk of being too interdisciplinary and consequentially being 'too thin' when reviewing the literature in those fields, and in

addition, using established ideas and terms inappropriately. Luckin explains her similar dilemma at the start of her research endeavour. Each discipline 'works within its own frame of reference with its associated and differing language, philosophy, concepts and methodologies. I also acknowledge that I run the risk of using these tools inappropriately as I blend them together in order to try and understand more …' (2010: 4). She considers it a risk worth taking.

Returning to the idea of context and relevance, Luckin references Cole (1996) and suggests two concepts for identifying context: (1.) that which surrounds and (2.) that which weaves together. 'Because what we call mind works through artifacts, it cannot be unconditionally bounded by the head or even the body, but must be seen as distributed in the artifacts which *are woven together* and which weave together individual human actions' (Cole, 1996: 136; quoted in Luckin, 2010: 10). Thinking about two students sharing a computer and using English as a second language, there is a physical context (the classroom comprised of layout, computer screen, mouse, keyboard) and cognitive contexts (languages, literacies, technologies) which could be weaved together to make better sense of the learners' embodied actions with each other and their environment. Another way of looking at this is to think of off-screen modes and on-screen modes and the coordinating actions of the learners between the two. A common approach is to use concentric circles to visually demonstrate contexts of relevance with the most important at the epicentre and radiating circles moving outwards in decreasing order of relevance. This can suggest separation. Here, overlapping circles are used to identify the contexts of relevance as they relate to the learners and the broader fields of literature.

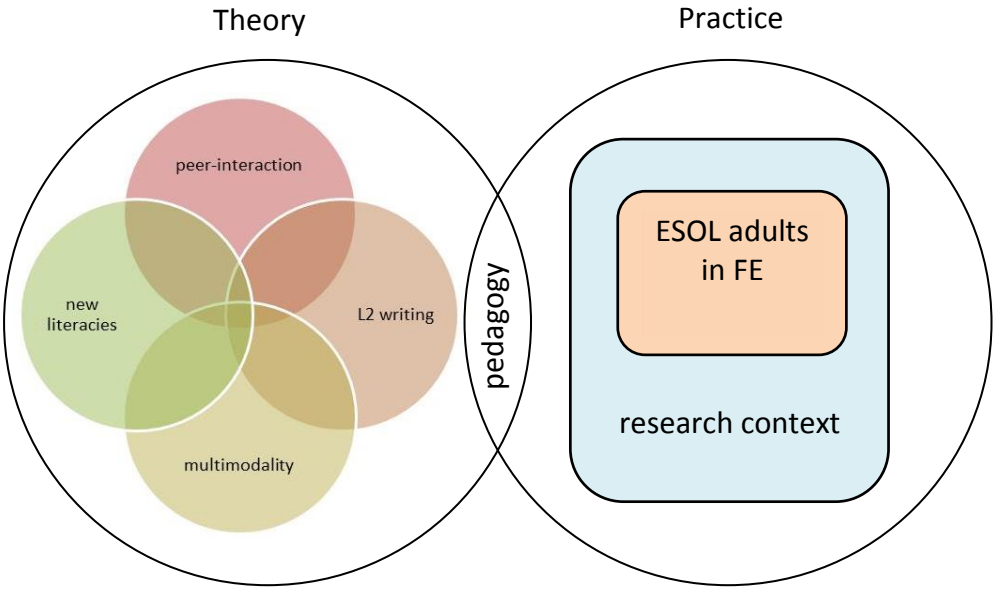


Figure 2-1 (research foci)

It is the intention of this research to work towards a theory of embodied peer-interaction and for those findings to inform the research context. The fields of theory informing this literature review and the research include: new literacies, multimodality, L2 writing and peer-interaction. The following sections will *weave* together the most relevant aspects in these fields as they relate contextually to the learners and their learning. This will require succinct discussion across the fields to prioritise what is most appropriate to the learners in their learning contexts. Like Luckin, I consider it a risk worth taking if the outcome is a more holistic understanding of the learners and their learning.

Section 2.2 offers a theoretical framework for the study, from the perspective of the students' textual habitats, and positions literacy within the research as plural, multimodal and digital. 2.3 narrows the focus to locate the research from a technology perspective as one of collaboration *at* a computer, rather than *through* a computer; the latter being customary in the field of Computer-Mediated Communication (CMC). 2.4 problematises what is meant by collaboration, as a broad term used loosely in many learning contexts, and defines peer-interaction as most relevant to the present study. Participation frameworks are considered as an overarching structure within which peer-interaction unfolds. Section 2.5 considers peer-interaction within second-language collaborative writing. Reviewing the literature reveals that this research can contribute towards the field of collaborative writing for lower level language learners, and especially when considering face-to-face computer collaborative writing (F2FCCW). 2.6 considers the fields of multimodality and embodied language to pinpoint the relevance to the current research. Multimodal *interactional* analysis is identified, compared to multimodal *text* analysis. Section 2.7 concludes the chapter and offers two research questions as an outcome of the literature review.

2.2 the multimodal classroom

In this section the discussion of literacy takes into consideration the increasing interest in the plurality of literacies and the multimodal dimension in new technological modes. Giving agency to the learners is a priority and one way of doing this is to position them as active designers of meaning. The discussion will contextualise the learners in this study from a digital literacy perspective as multimodal designers and consider what that

might look like and how, as a researcher, one might go about identifying the process of designing.

In their 'textual habitat' learners are surrounded by multimodal designs (Unsworth, 2001: 7). In their wider communicative landscape there is a plethora of literacies: websites, magazines, posters, emails, etc. This positions them as consumers of literacies. In the classroom they are similarly consumers of given designs, both tutor produced and existing *realia*. However, when learners are engaged in digital literacy events in class, producing their own designs, web-sites, writing, this is an opportunity to increase their textual habitat. Of interest is how they are problem-solving, negotiating, collaborating and writing, particularly in a second-language. Learners are 'inheritors of patterns' but they are also 'active designers of meaning' (Cope & Kalantzis, 2000: 7). To look for learner imitation and reproduction of existing designs is to simplify the literacy event. *Transformed practice* can evidence learning as can *transduction* of meaning across modes; in other words, where we find 'much of what we regard as "creativity" happens' (Kress, 2003: 36). An analysis of learner output following teaching activities is a feature of much multimodal analysis. It is a product-based approach to multimodality. Whilst this study is interested in the completed visual designs and writing of the learners it will not go into a detailed multimodal analysis of the completed products. How learners negotiate the design of new literacies, and transform practice in a process-based approach to multimodal texts, is of more interest than the completed texts themselves. There is a plethora of terms as researchers have tried to name literacy and technology:

- electronic literacies (Warschauer, 1999)
- global literacies (Hawisher & Selfe, 2000)
- multiliteracies (Cope et al, 2000); (Unsworth, 2001)
- multimodal literacy (Kress, 2001)
- silicon literacies (Snyder, 2002)
- new literacies (Kist, 2005)
- digital literacies (Martin, 2006)
- media literacy (Burn & Durran, 2007)

Each author addresses distinct aspects of technological literacy and yet each overlap in many respects. In particular, each draws varying levels of correlation between digital literacy and social literacy, that is, digital literacy as a socioculturally mediated practice no different to print literacy as defined by Street (1995) and Barton et al (2000) and

Papen (2005). The *plasticity* of the term literacy in technological contexts is taken up by Gourlay, Hamilton & Lea who ask if digital literacies is 'ambiguous and infinitely elastic' (2013: 7). If the pedagogical goals of classrooms are to enable a degree of mastery over numerous meaning-making resources, but the second-language of English is dominant across the multiple modes of communication, then a question might be asked about the 'elasticity' of literacy and technology in multimodal design collaboration. The classroom is a site where languages, literacies and technologies (LLT) converge. What participation frameworks are being constructed and contested and mediated by LLT? At a very dynamic level, in the day-to-day business of teaching and learning, there is language-switching, mode-switching, tool-switching and unexpected connections with numerous knowledge sources. For instance, learners design a website in the mode of paper and through first-language dialogue, utilising second-language translation products, but the completed design is transformed into a digital website with writing in the second-language of English and uploaded for public consumption on the internet. Here is evidence of *transduction* where modes work together to transform each other and transmute shared ideas into digitalised literacy artefacts. In the process of textual design, there may be scaffolding between learners to generate support and output in language, literacy and technology. For some, perhaps the 'new' requirement to become multiliterate is just too demanding, too great an ontological shift for those with a mind-set schooled in traditional literacy practices (Lankshear & Knobel, 2006: 29-62). This research seeks in particular to understand the sociomaterial process of interaction when individuals are paired according to their first-language and work collaboratively using a computer to design digital literacy products using English as a second language.

Pockets of languages often develop in ESOL classrooms where learners informally support each other. Sometimes a 'community interpreter' is found, such as an Arabic speaker with stronger digital or stronger literacy skills. Individuals thereby help each other, dependent and independent of the tutor, by switching languages from L2 English to a shared L1 and back again to support digital literacy processes such as using software or writing in a second language. This interpreter strategy, or use of a community scribe, is customary in multilingual settings. Baynham and others 'point to the use of mediators of literacy as a typical strategy for the accomplishment of literacy tasks' (Baynham, 1993: 296). In the instance of digital literacy practice described so far, the classroom is a place where the learners have reduced linguistic proficiency to

understand the dominant language of spoken English, are learning to use technology to express themselves, but have limited literacy skills to write in English and read instructions. In addition, the tutor as 'master' is circumnavigated by the different symbol uses of his students when they collaborate in their first language. In the pilot study, when asked about switching languages to L1 to support learning, there was an approximate 50% split across the respondents:

AM: It should be better but the problem is the people here in the class speak a different type of Arabic because we all have different backgrounds and accent. So Arabic is also a problem.

BC: Yes this might help.

JC: I don't think this is going to help you. If you are here in England then you need to speak English. You have to work hard and use English. If you want to use Spanish go to Spain.

MT: No.

NM: Yes, definitely. Like my friend when she need help she always call me.

RN: Yes, it might help but I'm not interested in speaking Kurdish in the classroom. I need to speak English. I can already speak my own language.

These few responses show the attitudinal beliefs of the learners. In the classroom described so far, the learning environment is perceived as one with many potential semiotic resources and where many of the learners have limited linguistic skills to work with the dominant spoken language and limited literacy skills to work with writing. Bilingualism is a strategy for dealing with these at the intersection of language, literacy and technology and merits further investigation within this research.

An important consideration is the *multimodal* which integrates two or more semiotic resources and multiple modes. Multimodality is identified by Cope & Kalantzis (2000) in their *multiliteracies* framework. Kress asks 'if the meaning of a message is realised, 'spread across', several modes, we need to know on what basis this spreading happens, what principles are at work' (Kress, 2003: 35). However, a significant number of literacy models and multimodal models view individuals as consumers rather than producers. The term *design* has been suggested to emphasis the agency of individuals as producers of multimodal texts. The term *design* is intended to suggest that learners are 'inheritors of patterns and conventions of meaning while at the same time active designers of meaning' (Cope & Kalantzis, 2000: 7). These two differences ('inheritors'

and ‘designers’) might also be polarised as product and process, and consumers and producers.

A trend in the literature is to position learners as consumers of products in an inheritance model of literacy. This research is interested in learners as producers of meaning in a designer model; this entails looking at the process of their visual and textual production. Pairing learners is commonplace in many language-learning contexts; a computer may configure in that process. The participation frameworks which learners establish when collaborating and designing are constituted on an ongoing flux of semiotic resources where off-screen activity and meaning is spread across action, language and the environment and the meaning-making is transposed across real-time into an on-screen product. Of interest is how those frameworks are structured, including both verbal and nonverbal modes. This study identifies the ESOL classroom as a multimodal learning environment where language, literacy and technology meet. Adult learners are dynamic meaning-makers who inherit knowledge from inside the classroom and outside its physical walls, as well as from each other and from the tutor, in a variety of languages and through traditional and technological modes. Increasingly, literacy studies focus on the sociomaterial aspects of digital practice. Gourlay, Hamilton & Lea discuss *literacies* alongside technological research interests to consider if ‘the concept of ‘literacies’ has lost its ethnographic and disruptive edge through its complex re-emergence and co-option in the messy term ‘digital literacies’’ (2013: 8). Considering classroom digital literacy as one which interconnects the multiple temporal and spatial practices of *online* and *offline* resources, Bhatt questions the *where* and *when* of a literacy event. The classroom is just one practice in a wider network of physical and digital practices; what Law calls *material semiotics* (2009) and Johri calls *sociomaterial bricolage* (2011). The potentially rich and multiple strands which weave together the tapestry of social phenomena, bridging the inside of the classroom with the outside of the learners’ lives, are not a central focus of this research because the network of associations under investigation is relatively small. The intention of this research is to understand how language interconnects with other modes in the immediate environment of the classroom rather than the learner’s complex interconnection with other practices outside the classroom. This requires detailed microanalysis of actions in the vein of Norris (2004) and Goodwin (2000, 2007a).

Of primary interest is how interaction is framed, actions sequenced and talk unfolded when learners collaborate on a single computer. 'If language is no longer the only or even the central semiotic mode, then theories of language can at best offer explanations for one part of the communication landscape' (Kress, 2000b: 153). The increasing forms of multimodal, digital literacies have led to a rethinking and re-naming of literacy processes, events and repertoires. As Kress states, language is central in this intersection of meaning-making, but it is not the only mode. There is always 'choice' and a coexistence of modes. The second-language classroom of today is a crucible of diverse languages mixed with traditional literacy practice(s) and emerging technologies. Plurality and polyvocality are central. In such instances it has become difficult to discuss literacy in the singular. There are a range of semiotic resources being used in classrooms. Individuals may have varying proficiency in, and reliance on, these resources. In pedagogical terms, to be multiliterate, one should ideally be able to consume and produce meanings across a range of semiotic resources which include linguistic, visual, spatial, etc: 'It is important to raise students' awareness of the variety of semiotic systems employed in texts and help them recognise that texts may be paper, electronic, and live' (Anstèy & Bull, 2006: 27). Here there are three text types 'and a range of semiotic systems, with which and through which, meanings are communicated in a classroom in a second language. There is then a crucial intersection where three text types merge. Figure 2.2 illustrates this intersection of modes as three overlapping triangles.

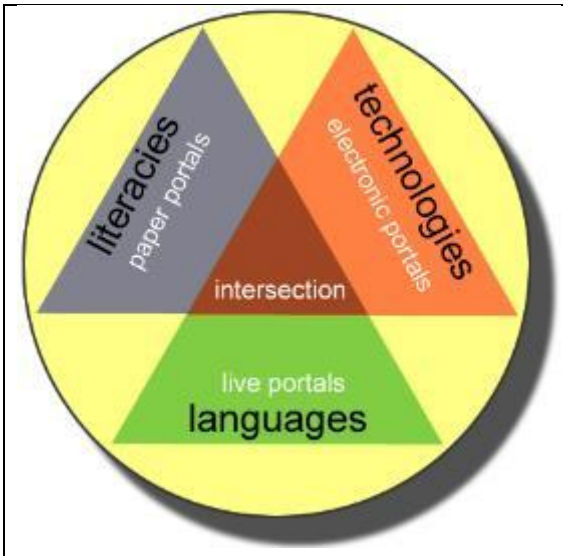


Figure 2-2 (multimodal intersection)


Figure 2-3 (Polish learners)

By ‘live’ the authors mean communication which is synchronous from person to person, typically spoken. By ‘portals’ they mean the different asynchronous media (traditional and modern) and individuals (peers, tutors, etc) which people may access. *Portals* are doorways to communication: ‘a portal is anything that gives access to the content and to ways of interacting with that content’ (Gee, 2004: 81). Of central concern to this research is what happens at this intersection when learners collaborate to design digital texts and write in a second language. At Figure 2.3, two adult Polish learners negotiate a screen-based form of literacy using technology, language(s) and the paper they are holding. They draw on each other’s existing knowledge, language(s), computer software, an electronic translator, the internet as well as the paper instructions provided by the tutor. Collaboratively, they ‘talk’ their way through *multimodal designs* (creating visuals and writing) on the computer, switching between English and Polish and paper and computer. For educational purposes, it is important to understand what learning constraints and what affordances are generated by this intersection. How is the multimodal design *realised* on-screen through off-screen negotiation and learner-collaboration? In such contexts we know very little about the ‘interactive organization of participation frameworks, including how they are structured and contested in the midst of moment-to-moment interaction’ (Goodwin, 2007a: 53). There is a need to understand the participation framework of technology and peer-collaboration at Figure 2-2 where digital designs are structured and contested in real-time. The sociomaterial configuration of space, modes and tools within the classroom are considered.

2.3 narrowing the focus

There are many language learning studies which justify technology as a tool for second-language acquisition (SLA). Sauro for example gives an overview of the last two decades to synthesise research findings on ‘the role of synchronous computer-mediated communication (SCMC) for second-language acquisition’ (2011, 369). More recently, but talking generally on the justification for using technology, Crawford-Thomas and Bloxham argue the case for learners, ‘If you’re not bringing digital into your teaching then it’s going to be really hard to meet their needs’ (2015). It has become common knowledge that technology is ‘good’ and the research on SLA mediated by technology is certainly rich (Chapelle, 2004; Sauro, 2011). This research will not consider traditional researched areas of collaborative writing with technology such as wikis, email exchanges, online forums, which generally fall under the field of asynchronous computer-mediated communication (ACMC). Nor will it consider SCMC such as text

chat, chat rooms, face-time videos, etc. Whilst this study may fall under the rubric of SCMC because the learners are synchronously working together using a computer, there is more interest in what happens off-screen as learners collaborate to produce a joint piece of writing which is written on-screen. It could be argued that this research is still positioned within CMC theory because 'CMC can take place with the person seated at the computer next to you or with someone on the other side of world' (Philp et al, 2013: 141). The contexts within which CMC occur are diverse so I want to briefly unpack this to narrow the scope and clarify the focus.

Prepositions of place can be helpful. Much of the literature on learner collaboration with ACMC and SCMC could be identified as collaboration *through* computers. The screen (be it computer, tablet, phone, TV) is a portal to wider communities through which communication and collaboration can take place. It is also possible to talk about learner collaboration *with* computers. Here the focus is on asymmetrical configurations through tutoring (or instructional) software where the programme contains the knowledge and the learner is positioned didactically in a top-down system where the computer 'initiates, the student replies, the computer evaluates, the computer initiates again, and so on' (Levin et al, 1990: 210). Computer-as-tutor is still prevalent in primary schools, for example, in literacy and numeracy programmes. Drill and practice *with* computers is not relevant here. Crook (1996) also discusses collaboration *in relation to* computers (chapter 5) and *at* computers (chapter 7).

One piece of research was found which is very comparable to the current research. Gardner & Levy (2010) analysed the collaborative interaction of two learners *at* a computer as they designed a webpage on recycling and the environment. The following quote is lengthy but is important to understand where there is understanding and where there are opportunities for enriching our understanding in collaborative computing research:

> Collaborative computing is a significant and pervasive social phenomenon in education and the workplace, from early childhood to adulthood. Whether the communication is face-to-face or computer-mediated, it involves human-to-human collaboration and interaction. This interaction may take the form of synchronous face-to-face talk, as two individuals work together at the computer, or asynchronous forms of collaboration, for example in the joint construction of a document using the Track Changes option in a word processing application. A considerable body of work has emerged that focuses upon the ways in which participants remote from each other use specialised software to complete tasks: this is the kind of research that is reported under the rubric of computer

> supported collaborative work. In contrast, much less consideration has been given to the impact of the IT tools – computer, keyboard, mouse, screen and software – on the participants' manner of working and the sequences of talk that unfold at the computer. What seems to have been overlooked in the transition from pen and paper to the computer over the past three decades is our understanding of some of the key differences in the working environment that the computer imposes (Gardner & Levy, 2010: 1).

Bhatt & de Roock (2014) also use screen recording software to explore digital literacy practice, but without the collaborative element which is the focus here. In education it is common practice to pair learners around a computer to complete a shared task; Mercer et al (2004) for example discuss quantitative and qualitative approaches to collecting data from such pedagogy; though the focus is on children in school settings. Collaboration *through* computers is well documented, as is collaboration *with* software, but not so well documented is how learners collaborate *at* the computer. In L2 adult contexts the motivation for pairing learners is typically to 'get' learners to use English as a second language; both using English to talk together and to use English for reading and writing. Learner collaboration for the purpose of talk, be it dyadic or group, is a primary method used in teaching practice but is also backed by research as beneficial to language development; (Swain & Lapkin, 2002; Swain et al, 2009; Swain, 2010). Shehadeh for example reviewed the literature and summarised how it 'has been argued that students' collaborative dialogues mediate the construction of linguistic knowledge and that this process of joint accomplishment of a task contributes to L2 learning' (2011: 2). Storch however draws attention to the fact that although 'pair and group work are commonly used in language classrooms, very few studies have investigated the nature of such collaboration when students produce a jointly written text' (2005: 153). Although there is now more research in the field since Storch wrote that a decade ago, the recent literature reveals uncertainty in our understanding of how students produce a joint text in real-time (Rouhshad, Wigglesworth, Storch, 2015). Understanding adult learners' 'manners of working' with technology and the 'sequences of talk that unfold' (Gardner & Levy, 2010: 1) are two overlooked considerations when pairing learners at a computer. The continuing sections will discuss these issues in more detail.

2.4 defining collaboration

This section will unpack what is meant by collaboration in educational, second-language contexts. It will then consider collaboration in second-language writing. *Scaffolding* as a metaphor for knowledge and skills sharing, as a building process, was

first used by Wood, Bruner & Ross (1976) to explain how teachers helped children solve a block construction problem. The term is etymologically rooted in the idea that a more capable individual helps a less capable individual learner; typically teacher and pupil. This needs teasing out.

Essentially, in educational terms, collaboration makes reference to knowledge which is shared but we need to foreground that knowledge-sharing exists in a material world, not just inside people's heads. Atkinson discusses shared knowledge acquisition to challenge cognitivist approaches: 'For cognitivists, language learning is invisible … Yet if cognition occurs not just *in* but *between* people, and between people and their sociocognitive environments, then it is also in the world. People learn, from a sociocognitive approach, by participating in extended cognition' (2010: 618). Co-cognition is at the core of what is meant by peer-interaction in an educational context; between two or more learners and less so the teacher. Although the tutor may have a significant role in establishing the grouping of learners and monitoring them (O'Donnell, 2006) he or she does not typically figure in collaboration research in education.

In this vein of sociocognitivism, common practice in the collaboration literature is to start with Vygotsky (1978) and the *zone of proximal development* (ZPD). Summarising the Vygotskian approach in relation to their research on the collaborative writing of college students, Ajmi & Ali contend that the 'cognitive and linguistic development of children appears through social interaction as they obtain scaffolding from the older members of society' and yet, the learners in their study are adults of symmetrical ability (2014: 2). A social constructivist approach to learning is important but noticeably the ZPD is an asymmetrical configuration and scaffolding is the process through which symmetry of knowledge is sought. There is a generalisation that older individuals and more competent individuals are best placed to enable the internalisation process of moving knowledge from what Vygotsky called the inter-mental social plane to the individual's intra-mental plane i.e. distributed cognition. Bruner talks in similar terms of *instructional scaffolding,* which 'refers to the steps taken to reduce the degrees of freedom in carrying out some task so that the child can concentrate on the difficult skill she is in the process of acquiring,' (Bruner, 1978: 19). Again it is an asymmetrical configuration in which an individual (i.e. a child) is identified as less able than another (i.e. a teacher or parent) and shared cognition is the goal. It could be argued that caution

in approach and precision of language are important when discussing such learning relationships. Collaboration as distributed cognition in the Vygotskian and Bruner tradition approximates the teacher-pupil dynamic, which is not the focus of this research, and as will be discussed, is a problematic simplification but frequently referenced in research literature. Discussing the cognitive and linguistic gains in the collaboration between adult learners with symmetrical ability is slightly different to the theory of pedagogy as it was originally intended for children with asymmetrical abilities compared to others in their environment. As Dobao summarises: 'The novice–expert relationship was originally described as a fixed and unidirectional relationship between a child and an adult' (2014b: 498).

Collaboration is an umbrella term. The literature reveals varying equivalents, all of which fall under the rubric of collaboration. Some are synonymous and some have subtle and important differences. 'There are many varieties of peer learning, but those most common to language classrooms are collaborative learning, cooperative learning, peer tutoring' (Philp et al, 2013: 3). These types of collaboration are identifiable in degrees of learner symmetry, or to use Philp's term, 'mutuality.' Collaborative learning might be seen to differ from cooperative learning in the sense that the former concerns itself with the cognitive gains in grouping learners whereas the latter concerns itself with the social organisation of grouping learners. In *cooperative learning* a pair or group of learners might have different tasks which contribute to a shared outcome. For example, learners may have similar or different proficiencies and each is tasked with collecting different information which is collated and shared to the rest of the class as a group activity. Working cooperatively to mutually benefit others is assumed to help with social cohesion, team-working and build confidence (Hillyard et al, 2010).

In many educational contexts it has become commonplace to group or pair learners purely for the purpose of socialising without necessarily planning pedagogically for cognitive gains through co-cognition of distributed knowledge. This is not to suggest that knowledge building does not happen through cooperative learning, rather, it is a matter of perspective. Just learning how to work together is considered by some researchers as knowledge in itself (Hammond et al, 2010; Johnson et al, 2007). Also, the tutor may have a pivotal role in facilitating the groupings, teaching how to 'cooperate' and managing behaviour (Gillies, 2008). From a computer-based perspective, there is a worry that learning *through* a computer in a 1:1 ratio can risk

students having a 'socially isolating experience' (Crook, 1994: 121). Pairing learners at a computer might sometimes be organised on linguistic and cognitive grounds to encourage distributed language and knowledge, but equally, the tutor rationale of pairing learners at a computer might be a simple concern with isolation; a worry borne out of the cooperative ethos in learning. This is anecdotal but many teachers encourage learners to socialise on sheer principle. There may not even be a pedagogical rationale.

Cooperative learning as collaboration is not a focus of this research; neither is peer-tutoring, but for clarification, *peer-tutoring* approximates the asymmetrical model discussed above in which a more proficient learner is paired with a less proficient learner. Primary schools depend heavily on this model, for example, through graded readers supporting others on lower level books. In L2 contexts learners might be paired in speaking exchanges where a more able learner is able to identify and change the mistakes of a fellow student. The 'expert peer' might gain something from the perspective of cooperative learning, being socially useful and functionally beneficial, and the 'novice peer' gains extra support which the teacher may not be able to offer if stretched across a large group of learners. Complications do arise however when proficiency is equated with status. Tensions can arise when peers are validated by the teacher as being more able, and by inference, more valuable (O'Donnell, 2006). There is less 'mutuality' in peer-tutoring because the learning relationship is top-down. Peer-interaction without an 'expert other' is the central focus of this research. Specifically in this paper, when discussing collaboration, it is within the context of second-language speakers where the learners are of an approximate ability. The teacher, more capable others and native speakers do not figure.

In *collaborative learning*, there is an assumed mutuality because learners are paired or grouped based on relatively similar proficiencies. They may self-group or be chosen by the tutor. There can be a spontaneous coming together with peers supporting peers. The preferred term in this research, when learners of relatively similar L2 ability collaborate, is *peer-interaction*. 'Peer' suggests mutuality. This pedagogical model is perhaps the most prevalent in language-learning classrooms because learners are often streamed according to group proficiency. The Common European Framework of Reference for Languages (CEFR) is a commonly used tool for identifying language ability against a set of descriptors; what people can and cannot do at a particular level. Similar scales include IELTS, Cambridge Main Suite and in the UK, ESOL Levels. Levelling of

language ability is common in language-learning settings because it allows for streaming. This is a slight generalisation, based on experience rather than any available research, but in classes of mixed-ability there is more likelihood of peer-tutoring being used for collaboration, compared to peer-interaction in classes where learners are of a similar ability.

A common rationale for using peer-interaction is to encourage learners to help each other with problem-solving in communicative tasks. The teacher for example might introduce a grammar or lexical issue and then pair learners to resolve the problem. The benefits (and potential pitfalls) have been well-documented, particularly by Swain & Lapkin in L2 contexts (e.g. Swain & Lapkin, 1998; Swain, 2000; Swain & Lapkin, 2001; Swain & Lapkin, 2002; Swain, 2006; Swain, 2010; Swain & Watanbe, 2013). Summarising the L2 literature on collaboration, Shehadeh finds that 'jointly performed tasks enabled learners to solve linguistic problems that lied beyond their individual abilities' (2011: 2). Frequent discussion is made in the literature, referencing Vygotskian social constructivism, that scaffolding to resolve linguistic problems is an outcome of collaboration between learner-expert and learner-novice when the learners under discussion are actually of a similar proficiency and are adults (e.g. Ajmi & Ali, 2014; Shehadeh, 2011). It is a marginal difference of perspective but learning relationships of an asymmetrical nature are perhaps better defined under the rubric of peer-tutoring, as described above, rather than the general umbrella term of collaboration. Scaffolding (or co-cognition to resolve a problem) can equally occur between learners of similar ability, as Crook explains, 'There is no reason why symmetrical peer pairings should not sometimes give rise to a socially defined cognitive system of the same sort: one that is comparable to that traditionally discussed for novices working with more expert partners' (1994: 134). This is also the finding of Storch: 'scaffolding can also occur among peers when working in group/pair work' (2005: 155) as well as other L2 research (e.g. Donato, 1994; Antón & DiCamilla, 1998; Swain, 2000; Swain & Lapkin, 1998).

Peer-interaction has been identified as the appropriate terminology to discuss the collaboration of learners as they work together on a computer to design images and write text. Consideration needs to be given to the classroom environment in which that interaction takes place. A framework for participation suggests a structure within which one can position embodied peer-interaction, what Goffman (1972) 'called an ecological

huddle, that creates a public, shared focus of visual and cognitive attention' (Goodwin, 2007a: 57). If learners are being asked to collaborate, we may want to understand the learners' kinaesthetic engagement with the semiotic resources of the classroom environment, including the *chronemic* arrangement of these in terms of synchronous and asynchronous actions alongside language. 'In the classroom, we will find student notebooks and class textbooks, but also many other meaning-inscribed material objects that afford heterochrony … students are designing and building … while they talk, and their activity spawns emergent practices and goals on several timescales' (Lemke, 2009: 281). Actions bleed into each other but every action also has a start and an end point and simultaneous actions overlap. There is an overarching structure within which shorter and longer actions, mediated by communicative modes and time, develop. Lemke explains further: 'Every process, action, social practice, or activity occurs on some timescale (in complex cases on more than one timescale). In a dynamical theory, an ecosocial system is a system of interdependent processes; an ecosocial or sociotechnical network is described by saying what's going on, what's participating and how, and how one going-on is interdependent with another' (2009: 275). In this research there is an interest in how semiotic resources are structured over time to realise learning outcomes, the negotiation and coordination of those in peer-interaction, and how learner actions unfold over longer or shorter scales of time within the process of one lesson.

The term 'participation framework' (Goodman, 2007a) is a term used to broadly name a scenario in which individuals organise themselves. This research is interested in the complex organisation of communicative modes, layout and learner alignment within temporal frameworks of participation/interaction as learners collaborate to produce images and text at a computer. Of specific interest is how those modes align, when, for how long and their significance to the unfolding collaborative actions of learners in the context of F2FCCW. Goffman (1981) first named 'participatory frameworks' to identify how people have shifting interactional roles in social situations. The speaker(s), hearer, over-hearer, etc, all have 'participant roles' in relation to spoken language: 'When a word is spoken, all those who happen to be in a perceptual range of the event will have some sort of participation status relative to it' (Goffman, 1981: 5). For Goffman, embodied interaction and actions, alongside other modes, are not given the same level of consideration as given to language. John Rae for example notes that 'despite its value, the idea of participation frameworks underemphasizes the importance of action;

in particular, how participants' actions make for unfolding contexts within which different actions become relevant' (2001: 255). Goodwin further elaborated on such frameworks by tagging on the adjective 'embodied participation framework' to draw attention to 'the interactive organization of participatory frameworks, including how they are structured and contested' (2007a: 53) within social situations. Goodwin's classification of embodied participation frameworks is used in this research to explain how learners collaborate. The research findings in section 8.3 provide commentary on learner alignment in relation to Goodwin's five stances of organisation. Goodwin states that 'The alignment of participants towards each other generates at least five different kinds of stance' (2007a: 70) including instrumental, epistemic, cooperative, moral and affective. These five types of alignment, and sometimes lack of alignment, are also evident in classroom collaboration so provide a theoretical framework in this research to identify and discuss how peer-interaction is structured.

Section 2.4 started with a call for more precise language use when discussing collaboration and scaffolding. Peer-tutoring is a better term for discussing contexts in which a more able peer supports another peer. Cooperative learning identifies the social organisation of collaboration, be it pairs or groups. Peer-interaction refers specifically to learners of similar ability working together in a shared problem space. In the current research that space is *at* a computer and the problem is writing together in a second language. Of significant interest then is *peer-scaffolding* to co-construct multiliteracies. The process of how learners do that, within a structured framework of participation, will be the main focus of this research.

2.5 peer-interaction and collaborative L2 writing

The previous section problematised the concept of Vygotskian ZPD in peer settings and identified peer-scaffolding as an appropriate term given the symmetry of the learners in the research. Reviewing the literature in collaborative L2 writing, many authors reference Vygotsky (1978) for rationalisation of the L2 writing pedagogy and then Donato (1994) to justify collective scaffolding: 'learners were individually novices, but collectively experts' (Lin & Maarof, 2013: 601); (see also Shehadeh, 2011; Dobao, 2013; Ajmi & Ali, 2014; Sajedi, 2014). From this we can presume with some confidence that there is general agreement in the field that a sociocultural approach to collaborative L2 writing is a legitimate pedagogy.

In this research, a theme arising from the pilot study was learner resistance and avoidance of writing. For example, in a pilot study learners were asked to produce multimodal texts incorporating image design and writing. When asked which was the most difficult, a common reply was the following:

RN: Writing! I hate writing.

NM: For me it's this, the writing.

BC: I think the writing because it involved a lot of difficult grammar and these kinds of things.

AM: The website is better. The writing is much more difficult.

TH: The writing is too difficult to understand second language. You have to know what is the meaning. That is too difficult.

As intermediate, second-language learners, there are a number of distinct concerns about the written word which merited further consideration as the research direction began to develop. Lin & Maroof worked with Malaysian ESL students and also found that the teaching of L2 writing 'led to negative perceptions among learners who view it as a skill they like the least' (2013: 599). Learners are 'afraid' of making grammar and spelling mistakes, demonstrate uncertainty around lexis, and there are further difficulties when thinking and talking and writing in different languages. 'Less advanced learners must often expose their shortcomings, for example by switching into the L1' (Gullberg, 2011: 138). In the pilot study, all learners said that when writing they composed mentally in their first-language before translating into English.

Of specific interest to this research is how off-screen conversation between two learners is peer-scaffolded and transposed to on-screen text, mediated by tools and other modes. Investigating the synchrony of talk and digital writing will give a unique insight into joint text-production within the context of face-to-face computer collaborative writing (F2FCCW); an area in which we know less compared to the usual fields of interest in CMC. 'Compared to research that examined the benefits of collaborative work for the spoken discourse, research investigating the benefits of collaborative work for the written discourse in L2, especially collaborative writing (CW), is scant' (Shehadeh, 2011: 2). Shehadeh is referencing traditional pen-and-paper collaboration. As suggested, CW as pedagogy is not as well researched as the literature on collaborative speaking as pedagogy.

There is very little comparable research where second-language speakers collaborate in real-time at a shared computer to produce writing; the closest research correlation is students using Wikis asynchronously and remote from each other but even this is limited research in L2 contexts (Kessler, 2009). Storch for example states that: 'the nature of the writing process and of the written text produced have received scant attention' (2005: 155). Not much has changed: 'the number of empirical studies that have investigated collaborative writing in L2 classes is relatively small (Storch, 2011: 277). Recent publications, though still a relatively small volume when compared to other papers in applied linguistics and multimodal text analysis, reveal a continuing interest in the process of L2 collaborative writing (e.g. Ajmi & Ali, 2014; Dobao, 2014a, 2015; McDonough et al, 2015). Given the focus of my research interest is predominantly on the coordination of talk and text (the process) rather than an evaluation of the writing itself (the outcome) this section will consider areas of interest to F2FCCW.

Reviewing previous research findings and activities in collaborative L2 writing, Storch finds that 'the use of small group/pair work in writing classes seems quite limited. It tends to be limited to the beginning stages (brainstorming), or more commonly, to the final stages of writing—the peer review stage … One of the drawbacks of peer reviews, however, is that the focus is often on the product of writing rather than the process of writing' (2005: 154). Peer-reviews in collaborative writing are seen to be beneficial (Ferris, 2003) but these tend to be at the end of individual, private writing rather than co-constructed sentences throughout the process. Peer-reviewing tends to pick out inaccuracies in form, mechanics, lexis (Nelson & Carson, 1998) when perhaps what we want to see more is learners mutually engaging with the linguistic and cognitive ingredients of making the cake rather than turning up at the end to review how good or bad it tastes! Storch & Wigglesworth have long been proponents of learners collaborating throughout the writing process, not just the beginning or the end (Storch, 2005; Storch & Wigglesworth, 2010; Wigglesworth & Storch, 2009).

In other controlled experiments where learners collaborated on the whole document, Sajedi compared the collaborative L2 writing of dyads and triads and found 'that students in pairs benefitted the most' (2014: 1650). However, Dobao also undertook comparative research into the L2 writing of individuals, pairs and groups and found that

‘the texts written by the groups were more accurate not only than those written individually, but also than those written in pairs,’ and she notes that collaboration ‘whether in pairs or in small groups, resulted in greater grammatical and lexical accuracy’ (Dobao, 2012: 55). Sajedi also found that overall there were still improvements in L2 writing amongst all groupings, though they both found slightly different results in the size of the groupings. There is agreement in the field that a wider linguistic pool to draw from has some correlation with increased accuracy, though clearly more research is needed here to find agreement.

A common tool for analysing collaborative L2 writing is to use language-related episodes (LREs) and to use language as a verb, to *language* (Swain and Lapkin, 1998; Swain, 2006). LREs are when individuals ‘talk about the language they are producing, question their language use, or correct themselves or others’ (Swain and Lapkin, 1998: 326). Later adapted by Swain, she states: ‘Languaging, as I am using the term, refers to the process of making meaning and shaping knowledge and experience through language. It is part of what constitutes learning. Languaging about language is one of the ways we learn language’ (2006: 98). Using these terms with translanguaging (Garćia, 2007: xii) gives researchers a terminology and a methodology for identifying and analysing how learners negotiate L2 with each other and sometimes through L1. ‘LREs are interpreted as segments of dialogue that illustrate the process through which learners use language to shape L2 knowledge by talking about, questioning, or reflecting on the linguistic properties of the second language’ (McDonough & Sunitham, 2009: 232). Translanguaging is when individuals perform the same strategy but use a first-language to metacognitively work with the second-language. Significant and recent studies which use LREs in collaborative L2 writing include McDonough & Sunitham, 2009; Dobao, 2012; Amirkhiz et al, 2013 and Ajmi & Ali, 2014.

From a student perspective, there is frequent reference in the literature to words such as student ‘perceptions’ and ‘views’ and ‘reflections’ where the focus is on how students feel retrospectively on sharing writing tasks, with most having a positive view (e.g. Shehadeh, 2011; Dobao, 2013; Lin & Maarof, 2013; Ajmi & Ali, 2014). However, Storch finds ‘a persistent reluctance on the part of learners to engage in co-authoring … and of teachers to implement such activities’ (2011: 285). Watanabe (2008) suggests that attitudes change positively after teachers and learners experience collaborative writing.

At the time of writing (July 2015) I emailed some lead authorities in the collaborative L2 writing field (e.g. Neomy Storch, Ana Dobao) to ask if they could recommend similar research to the current study. Dobao suggested McDonough & Sunitham (2009) as the closest similar research she could think of. The authors state: 'The nature of face-to-face oral interaction that occurs when learners do computer activities in pairs or small groups has been relatively neglected' (McDonough & Sunitham, 2009: 234). However, their pairing of Thai EFL learners was arranged to investigate how the students worked together through educational software, and whilst the learners languaged and translanguaged together to understand the English of the software, there was no writing collaboration. McDonough & Sunitham were able to show the frequency and type of language-related episodes.

A brief review of the field has identified a number of similar findings to suggest there are linguistic gains in the pedagogy of collaborative writing. Two gaps in the literature were (1.) very few researchers using audio methods to record the negotiated process of writing; some evidence for audio-recording (e.g. Dobao, 2012; Amirkhiz et al, 2013) but no evidence of videoing. Pre-testing and post-testing on fluency, complexity, accuracy, etc, were commonplace with learner surveys in a quantitative approach. (2.) No similar research was found on learners collaborating in real-time with a shared computer to compose their writing other than Gardner & Levy (2010) in a high-school native speaking context and Bhatt & de Roock (2014) who similarly record on-screen activity as digital literacy events, with the addition of recording off-screen actions, but do not discuss the second-language writing of the learners. From the L2 collaborative writing field I take primarily the idea of using language-related episodes to identify focal points in the collected data and languaging/translanguaging as learner strategies for collaboration.

2.6 peer-interaction and multimodal interactional analysis

This section will consider the different strands of multimodality to make clear the historical trajectory of multimodal interaction in the literature as it relates to this research. Clearly there is a rich history to the study of language in action and there are many fields of application in which one could position a multimodal approach, and out of which multimodality grew: (e.g. *communicative competence* as ethnography, Hymes, 1964, 1972; *conversation analysis* as ethnomethodology, Sacks et al, 1974 and

Garfinkel, 1967; *interactional sociolinguistics* as anthropology, Gumperz, 1982 and Schiffrin, 1997) to name but a few. The approach in the current research is to view language as an embodied, interactional process. A 'thick' description of spoken language in use, alongside nonverbal modes and the classroom environment, is the preferred grounded approach to the research methodology. 'The term "embodiment" is ambiguous, having taken on a number of philosophical and theoretical dimensions … for our purposes we want to embrace the general principle that as human beings we live our lives through embodied experience … and this has an impact on how we communicate as well as how we think' (McCafferty & Stam, 2008: 3). This approach to embodiment is applied similarly in this research. The following discussion will further clarify.

Multimodal analysis attempts to broaden our understanding of communicative modes and behaviours in addition to language because 'language and action are related dynamically' (Gardner & Levy, 2010: 2). Whilst this might always have been understood it has not necessarily been articulated in such a way. As Cook explains: 'While the complex interactions of language and paralanguage in speech are ancient and universal aspects of human communication … and in this sense the phenomenon is nothing new, the term 'multimodality' is mostly associated with written communication' (2011: 438). So as to position the theoretical research approach of this study, it is important to clarify the different strands of interest in multimodality. Cook alludes to two related but alternative approaches to multimodality. *Multimodal text analysis* is an approach to written communication, taken for example by Kress & van Leeuwen (2001, 2006) and Kress (2010); discussed previously in section 2.3. The spatial layout of writing alongside images and colour, the affordance and constraints of each, is typically given primary interest; borne out of social semiotics and the Hallidayan (1978) approach to the complexity of texts as materialisations of ideational, interpersonal and textual metafunctions. Jewitt (2009) uses the term 'social semiotic multimodal analysis' to refer to the approach of multimodal text analysis. Ideology and discourse are central. Jewitt (2009: 28-39) also offers an overview of the different approaches to multimodality and suggests an additional strand which is related to multimodal text analysis. *Multimodal discourse analysis* (MDA) is a theoretical framework which builds on Halliday's systemic functional grammar (1985). O'Halloran (2004) has done significant work here and as the title suggests, discourse is central. MDA is not a feature of this research and will not be discussed.

Multimodal text analysis was discussed in section 2.3 under the rubric of digital literacies because the learners in this study are collaborating to design their own multimodal texts. As stated, their texts will not be analysed in detail from a multimodal perspective. Cook's first reference to multimodality is concerned with language and interaction, rather than language as text. The central focus of the current research is *multimodal interactional analysis*, videoing and transcribing the learners' nonverbal and verbal language, and their actions, to understand the fluid structure of peer-interaction. This approach aligns more with the research of Norris (2004), Lemke (2009) and Goodwin (2000a, 2007a).

In what could be called an ethnographic approach to language in use, multimodal interactional analysis concerns itself with situated interaction and context. Scollon and Scollon (2003) for example offer a triadic approach to language in the material world, comprised of the interaction order, visual semiotics and place semiotics. 'We are calling this theoretical framework *geosemiotics* to make reference to the social meanings of the material placement of signs' (2003: 4). Jewitt (2009: 33) places the Scollon's approach under *multimodal interactional analysis* but this could be a mixing of multimodal approaches. Whilst the Scollons do engage with spoken language, the central focus of their influential *Discourses in Place: Language in the material world* (2003) is written communication as multimodal discourse and how people are positioned, and position themselves, their indexicality, in relation to visual signs and discourses in place. It is predominantly about people's interaction with multimodal written language rather than interaction via multimodal spoken and nonverbal language with other people. This is how I am applying *multimodal interactional analysis* in the current research, which is succinctly explained by Goodwin: 'a primordial site for the analysis of human language, cognition, and action consists of a situation in which multiple participants are attempting to carry out courses of action in concert with each other through talk, while attending to both the larger activities that their current actions are embedded within, and relevant phenomena in their surround' (2000a: 1492). Embodied language and action and distributed cognition are central.

Picking out some of the key terms in Goodwin's definition: the *primordial site* of this research is an L2 classroom; the *participants* are students; the *actions* are multimodal talk; the *larger activities* are designing and writing digital texts and *relevant phenomena*

are the tools, texts and spatial configuration of the computer they share. What the analysis seeks to understand is how 'language, cognition, and action' are coordinated in peer-interaction at a computer. Norris summaries the position of the would-be researcher: 'Multimodal interaction analysts set out to understand and describe what is going on in a given interaction. We analyse what individuals express and react to in specific situations, in which the ongoing interaction is always co-constructed' (2004: 4). In multimodal research there is a growing body of work considering how language is part of a larger multimodal ensemble in embodied environments. As discussed, there are different approaches though many features overlap. Nevile refers to 'the embodied turn' when he reviews and synthesises two decades and 400 papers on language and embodied interaction in the journal *Research on Language and Social Interaction* (ROLSI): 'I chart the rising interest in the involvement of the body—embodied conduct or "embodiment"—in research on language and social interaction' (2015: 121). Using a quantitative approach, Nevile charts the publication history and comes to an interesting conclusion: 'The year 2001 appears to mark a point of discernible change, the beginning of an established rising interest in embodiment, and so constituting the embodied turn for research on language and social interaction' (2015: 127).

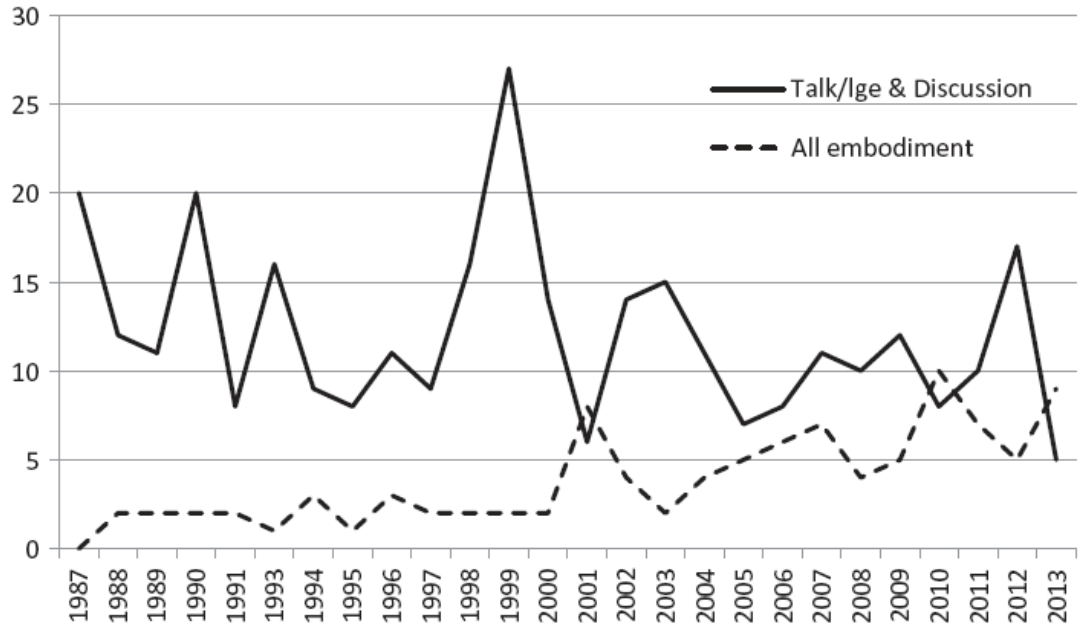


Figure 2-4 (embodiment papers in comparison to language, Nevile, 2015: 127)

The dotted line in the above chart shows the published research papers on language and embodiment in the journal ROLSI against the thick line representing published research articles in which the focus was just talk, language and discussion. It can be seen that in

2001, 2010 and 2013, published papers on embodiment with language slightly overtook language only papers, and has increased from 2003. It is perhaps no coincidence that the increasing interest in the field of embodied interaction with language mirrors that of the increasing interested in the field of social semiotics: 'From early 2000 there has been an explosion of interest in multimodality within research' (Jewitt, 2009: 19). Seminal texts in the embodied interaction literature include Goodwin's *Action and Embodiment* (2000a), McNeill's *Language and gesture* (2004) and Norris' *Analysing Multimodal Interaction* (2004). Seminal texts in the multimodal literature include Kress & van Leeuwen *Reading Images* (1996, 2nd edition 2006), Scollon & Scollons' *Discourses in Place* (2003) and the New London Group's multiliteracies pedagogy (Cope et al, 2000). The 'embodied turn' from 2000 might be said to mirror a 'multimodal turn' from roughly the same period. In both fields we see a flourishing interest in multimodal text analysis and multimodal interactional analysis. It is to the latter I now turn.

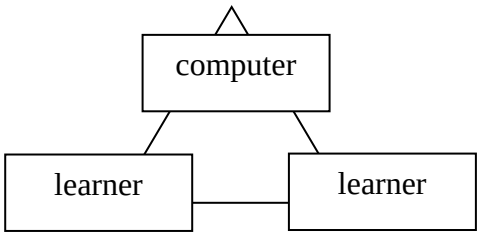
The book *Embodied Interaction, Language and Body in the Material World* (Streeck, Goodwin & LeBaron, 2011) is a handbook which engages with this very issue, be it embodied interaction in the home, the office, the classroom or a hospital. The title succinctly captures the increasing interest in embodiment in recent research. Norris summarises the importance of considering modes in addition to language: 'Previously, language has been viewed as constituting the central channel in interaction, and nonverbal channels as being subordinate … I believe that the view which unquestionably positions language at the centre limits our understanding of the complexity of interaction' (2004: 2). Language often is the primary communication channel, and sometimes it is not, but it is always part of a larger orchestral movement in which posture, gesture, proxemics, gaze, silence, and the layout of the immediate environment, to name but a few, all play their part in interaction between two or more people. The simultaneous organisation of such modes as learners collaborate at a shared computer raises questions about how that interaction is coordinated.

As discussed, research on multimodal peer-interaction in L2 contexts, and as collaborative digital writing, is limited and the current research will contribute to these. Gullberg explains: 'Despite popular convictions that L2 learners use all means at their disposal to communicate, their multimodal behaviour has received surprisingly little attention, both descriptively and in theorising about L2 acquisition and use. Moreover, the focus has largely been on the individual learner in isolation from the interactional

and multimodal context where the problems typically arise' (2011: 137). An individual who is less proficient with language may resort to other nonverbal modes if she or he wants to communicate. Writing and signing are obvious compensation strategies for someone who cannot speak but that is not the focus of this research. For individuals communicating in a second-language, with perceived limited fluency and inaccuracies in form, it is suggested that other modes have a role to play. 'It is a popular lay view that learners use their hands and feet to compensate for lexical shortcomings in their L2' (Gullberg, 2011: 138). This is perhaps a simplification because of our own lack of understanding in how other modes are being used. There are many types of nonverbal communication, including 'kinesics, proxemics, haptics, chronemics, physical appearance and paralanguage' (Quinlisk, 2008: 27). The following will briefly explain some of these modal categories as they relate to this research.

Proxemics (Hall, 1963), *vocalics* (Trager, 1958; Gumperz, 1982), *kinesics* (Birdwhistell, 1952) and *chronemics* (Lemke, 2009) are the most prevalent modes of interest, in addition to language, in this research. For instance: students sharing a single computer are arranged proxemically in an unusual configuration. A computer with two learners creates a triadic shoulder-to-shoulder arrangement with the computer screen as the third party, rather than the typical dyadic face-to-face interaction.

Figure 2-5 (triangular layout)



Consideration needs to be given to how peer-interaction and embodied modes are mediated by the proxemics of the environment, including the 'cognitive technologies' (Clark, 2001) of keyboard, screen and mouse, and what the impact of that arrangement might be on learner talk and action. Ulhirová (1994) for example audio-recorded two engineers sharing a computer and referred to the computer as a 'silent partner' – an interesting metaphor revealing something of the interpersonal relationship which people can have with technology. Gardner & Levy also reference an anthropomorphic third partner: 'a new form of collaboration which is possible with a computer – "collaborative personal computing"' (2010: 9).

Ulhirová states: ‘As a semiotic act, PC talk consists of (i) speech communication between the participants, and (ii) non-verbal interaction between the participants and the PC through the keyboard and screen. If we want to understand the speech in full, we must not leave the non-verbal interaction out of account’ (1994: 514). As discussed so far, embodied engagement with other people and objects could be factored into any analysis of language and peer-interaction. Unfortunately, Ulhirová did not use video technology to capture their interaction *with* and *at* the computer. Gardner & Levy also audio-recorded two individuals (high school students) collaborating with a shared computer, but did not use video technology to record off-screen action. To better understand proxemic engagement with technology and each other, the current research will video on-screen processes and off-screen actions, alongside talk, to identify and itemise the inter-semiotic, multimodal coordination of peer-interaction with technology and the environment. Referencing Gulson & Symes (2007), Luckin explains that they suggest ‘the treatment of space and place in educational settings is underexamined, undertheorized and underdeveloped’ (2010: 6). Whilst this study will not make much of a contribution here, it will however engage with the learners’ proxemic alignment in a classroom space around a computer.

How learners speak to each other, in addition to what they say, merits *vocalic* (or paralinguistic) consideration: speed of speaking, pausing, loudness, intonation, accent, etc, are all additional features of conversation. As a *silent partner*, how might a computer change some of the paralinguistic elements of learner talk? Gardner & Levy for example notice that ‘sometimes work is done to ensure coordination, either by slowing down the talk or pausing or stretching sounds mid-utterance’ (2010: 1). Audio-recording learners, with on-screen and off-screen videoing, might help to reveal what features of paralinguistic talk are evident and to what extent those are mediated by the cognitive technologies the learners are working with.

Chronemics refers to ‘how we perceive, structure, and react to time and ... the messages we interpret from such usage’ (Burgoon & Saine, 1978: 99). Section 4.3 will consider scales of time and the duration of actions. Chronemic behaviour is a mode and is referenced in research under different synonyms such as ‘time-lag’ (Yacci, 2000). If I send a text message to a close friend and do not receive a reply relatively quickly then I may become worried or feel I am being ignored. We expect reciprocal behaviour and have cultural expectations on the temporal duration of reactions to our actions, be that in

digital communication or physical. In conversation analysis we recognise turn-taking, not talking over people, wait time, lead time, etc, all of which are structured by an awareness of time and are considered in this research (Sacks et al, 1974). 'Despite the fact that time is handled differently between cultures, time messages in a communication event still convey meaning across multiple levels' (Walther & Tidwell, 1995: 361). Time is linear for some cultures and cyclic in others. For some cultures and individuals, actions are monochronic and for others polychronic. Any analysis of peer-interaction necessarily must take an interest in the chronemic behaviour of learners, the extent to which actions are sequenced, actions reciprocated and duration of actions over seconds, minutes and hours. Hall refers to *informal time* as the 'rules and expectations we learn from our culture' (1959: 3). Observing interaction includes an interest in how learners conform or deviate from the temporal expectations of others.

Kinesics (Birdwhistell, 1952) refers to the full range of bodily movements including posture, gaze, facial expressions, mirroring behaviour, etc. Broadly speaking there are four types of gesture which fall under the category of kinesics:

- deictic - pointing to objects or people often in synchrony with language: 'this' 'that' 'him' 'her'; from the Greek *deiktos* meaning 'able to show directly;'
- iconic – using the hands to visualise a concrete object such as writing a word on a table with a finger to try and spell it;
- metaphoric – visualising the abstract such as tapping on one's head to suggest 'craziness' in another;
- beat – (in/out) or (up/down) movements such as a tapping a foot to music.

A student rubbing their stomach as a *metaphoric* gesture for stomach pain, coupled with a pained facial expression, is one way in which kinesics can communicate when a student does not have the language to discuss their abdominal issues with their teacher. An *iconic* gesture might be a tutor raising their hand to their mouth as if drinking from a cup to communicate it is time for a break. A quickly tapping *beat* gesture on the table might be a signal to indicate the teacher wants the students to finish a task quickly. A good summary of the field is McCafferty & Stam, *Gesture: Second-language Acquisition and Classroom Research* (2008) in which research from several authors evidence how nonverbal communication is used alongside English as a second language. Typical configurations include teacher-student and student-native speaker

with the general consensus that in 'L2 contexts, the gesture-language connection can have important consequences' for vocabulary retention, explanations and improved engagement in the learning process (Quinlisk, 2009: 27).

Student-student gestures with language do not attract the same level of interest. In this research attention will be given to the kinesics of nonverbal communication as learners collaborate at a computer. It is a common classroom observation to see students point deictically at the screen to draw another student's attention to something they want to share or need help understanding, including pointing and tapping on the screen with phrases like 'click this' and 'what's that' or 'go there.' The mouse cursor might be used as a proxy for a finger to draw another learner's attention to something. Sometimes a circular movement with the mouse cursor on the screen might be used, like a circling finger on a map or newspaper, to draw someone's attention to the boundaries of the focal point under discussion. In such instances we can say there are different types of semiotic practices structured synchronously (language, gesture, screen) which the addresser and the addressee use as embodied cognition.

'Researchers have estimated that nonverbal communication constitutes more than 60 percent of encoded messages in adult communication' (Quinlisk, 2009: 29). For a linguist/analyst of communication that could be considered a lot of missing information if one was to focus only on the other 40% of spoken language. Admittedly, some of the nonverbal elements might be superfluous, but even if that is the case, it still may have some relevance. For this research, multimodal interactional analysis will seek to observe the totality of communication to understand the coordination of the verbal and the nonverbal, as embodied modes in a classroom where L2 learners collaborate at a computer. In this section it was shown how there are multiple approaches to multimodality and these were broadly categorised under three analytical headings: text, interactional and discourse. All have seen developing research interest in what could be called the 'embodied turn' and the 'multimodal turn'. Embodied peer-interaction was discussed as one of the central research interests. Understanding the coordination of off-screen talk, on-screen text and actions will be a primary outcome of this research; the findings of which are summarised in the discussion chapter at sections 8.2 and 8.3.

2.7 conclusion

Chapter 2 explored the relevant literature across the fields of new literacies, multimodality, L2 writing and peer-interaction. A concern was raised that an interdisciplinary approach such as this risks a lack of depth in any one particular field. However, in each of these disciplines it was demonstrated how these are fields to which the current study can contribute. Looking for example at collaborative L2 writing, but sharing *at* a computer rather than *through* a computer, as is common in CMC studies, revealed that there was not a great deal of relevant literature to discuss (Shehadeh, 2011). Multimodal collaboration at a computer, as embodied interaction, also demonstrated a lack of similar research. It was shown that there has been a profusion of research on multimodality and embodiment over the last fifteen years, in the fields of multimodal text analysis and interactional analysis, but very little research into peer-interaction at a computer (e.g. Ulhirová, 1994; Gardner & Levy, 2010). In addition, L2 research demonstrated a concern with multimodality in SLA. Significant research has been undertaken with language but less so on how L2 coordinates with other modes in interaction, whether that is to enrich communication or to compensate for deficiencies (Gullberg, 2011). The idea of an 'embodied participation framework' was introduced as the organising structure within which modes and actions unfold across different scales of time (Norris, 2004; Lemke, 2009; Goodwin, 2000a, 2007a). It is not clear what such a framework might look like with two learners collaborating at a shared computer and writing in a second language; nor is it clear what other modes are evident, and are dependent on, to transpose off-screen talk to on-screen text.

Taking into consideration this discussion on potential research areas for contribution, and considering the learners' voiced concerns with second-language writing, the following two research questions are raised:

- *How are off-screen talk and on-screen text coordinated?*
- *How are peer-interaction frameworks for learning structured?*

In relation to the first question, there will be significant linguistic-scaffolding between the learners revealed through *language-related episodes* (Swain, 2006) where the learners *language* and *translanguage* on form, mechanics, lexis. Learners also widen the linguistic pool beyond themselves to access other linguistic portals: paper, electronic, and live (Anstey & Bull, 2006) such as websites, dictionaries and other

people. Additional modes will be evident and these will be identified through their coordination with language and action, including proxemic, kinesic, vocalic and chronemic configurations. All of these sit within a peer-interaction framework with participant roles and modal alignments. Answering the first question will lead to an understanding of the second question, discussed in sections 8.2 and 8.3. How one goes about collecting and analysing the data to answer these questions is the subject of the next two chapters.

3 Methodology

3.1 introduction

This chapter describes the activities and tools used to collect the data. Each section will elaborate on the methodology and provide a rationale for the activities and tools used, including a discussion of problems that occurred during the process. 3.2 identifies the research participants and the educational context. Section 3.3 considers earlier instruments and trial data collection in the piloting phase and the epistemological changes that were an outcome of this experience. 3.4 discusses the experience of interviewing learners in the study and considers issues around reliability and validity. 3.5 locates the research within a practitioner-researcher case study using an ethnomethodological approach. 3.6 describes the tools and methods used to collect the data in a single lesson and across paired participants. Observation as video analysis has four broad sequential strands: *collection*, *preparation*, *transcription*, *analysis*. Section 3.6 discusses the *collection* of the video data; the situated, videographic challenge of deciding who and what to record. Section 3.7 considers the *preparation* and pre-analysis task of video editing and synchronising multiple video streams. *Transcription* and *analysis* have their own sections in chapter 4. Section 3.8 concludes the chapter on the research methodology.

3.2 the research participants and educational context

This section provides biographic and demographic information on the learners in the case study. There is an ethnographic consideration of the situated practice in which the case study takes place, patterned as it is by a number of forces: migration, deprivation, funding, employability, educational progression and especially language and literacy. In brief, the research participants are asylum seekers, refugees and migrant workers studying in a community college in Leeds. They each attend the same ESOL class where the focus is on developing language/literacy but with the use of technology as a multiliteracy strategy. Leeds City College is one of the largest further education organisations in the UK. This study is based in a community centre belonging to the college. The immediate neighbourhoods of Gipton and Harehills are multicultural and home to a significant number of asylum seekers, migrant workers and refugees.

The Indices of Multiple Deprivation are produced by the Department for Communities and Local Government based on a range of poverty indicators: income, work, crime,

health, education, environment and housing. Statistical information is gathered by the Neighbourhood Statistics Service, from the 2001 and 2011 national Censuses, which in turn are used to inform the Indices. The information is used to inform policy making in the Government's National Strategy for Neighbourhood Renewal. The Indices are able to show where the most deprived areas in the country are and thus inform policy making and prioritise funding. The immediate neighbourhood in which the case study is located, and where most of the learners live, ranks highly in the indices of deprivation (Indices, 2010: Figure 3-1).

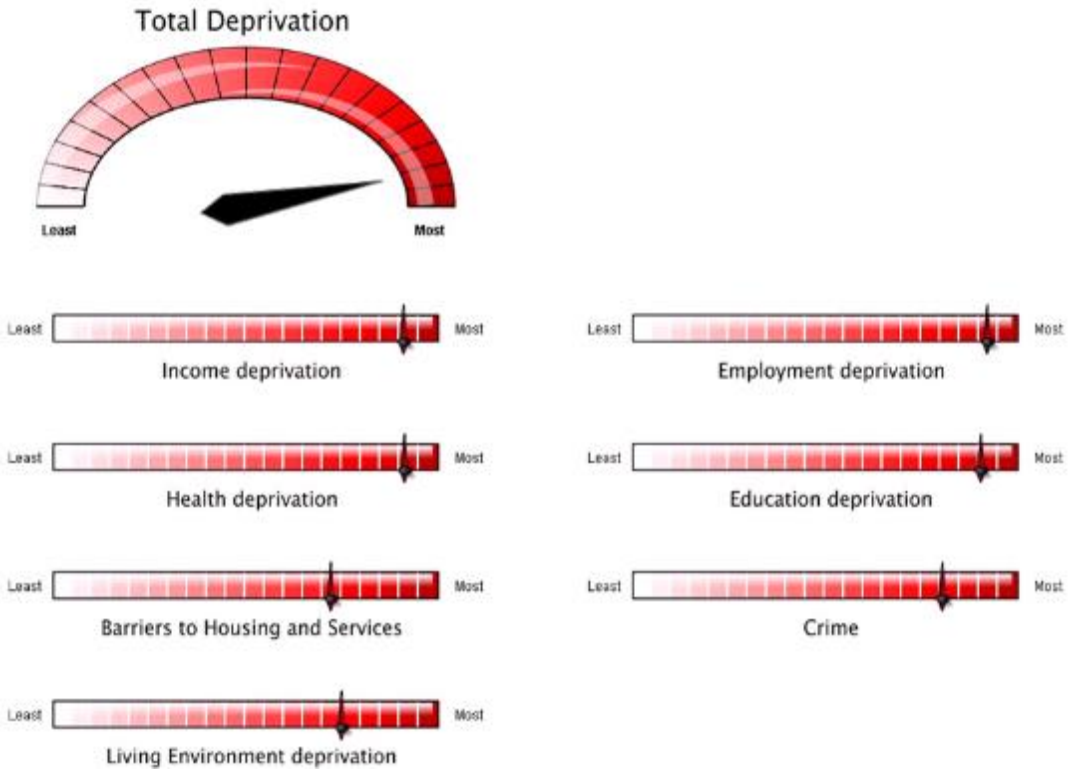


Figure 3-1 (indices of deprivation, Harehills, Leeds)

From the most deprived on the right, to least deprived on the left, the Indices evidence that in the neighbourhood in which the case study is situated there are many people out of work, with low educational attainment, limited income and crime is high.

The Harehills ESOL Needs Neighbourhood Audit (HENNA) project was completed in 2011 and produced a comprehensive demographic study of the individuals and their needs in the same area. Significantly, 'migrants bring with them valuable abilities, qualifications and experience which can lie untapped unless they have the chance to learn English to an appropriate level. The way that this is achieved is through English language provision known as ESOL' (Simpson et al, 2011b: 1). Homogenisation should

be avoided but there are common features across ESOL learners who come to the UK in the local area. There is often very little English and usually a limited educational background. Stabilisation of L2 proficiency at Entry 3 is commonplace. The domino effect of this is clear, though greatly simplified here: ESOL adults lack the language and literacy skills to pass upper-intermediate qualifications; are then unable to gain access to vocational or Higher Education learning; and are thus unable to gain the qualifications they need to work in the sectors they aspire to. In a national survey, ESOL learners were found to be 'over-represented amongst the unemployed and low paid' (Baynham et al, 2007: 12). The HENNA project found similar results for the local neighbourhood: 'Compared to Leeds as a whole, Harehills has much higher levels of out-of-work claimants' (Simpson et al, 2011b: 23). Asked why they are presently in a class, most will talk about progression to university and potential career opportunities. In the pilot study the learners were asked what they thought would prevent them from realising their goals:

MT: Language! Because I've already got my certificates in German but over here the language people want you to have.

NM: I need to improve my English. Like next year I need to do an English class, definitely.

OD: The only thing that can stop this at the moment is the state of this country. I am still an asylum seeker. It is not easy.

RS: My English maybe.

ST: Yes. My status first which is the basis of everything. That is also the reason why I sometimes ask myself 'why am I struggling?' I'm still fighting to shift it. I wouldn't give up.

What the pilot study showed is that temporal uncertainty is an occurring feature in these learners' lives. As well as in the classroom where some are stuck in the bilingual space between their existing L1 competence and L2 languages/literacies, most voiced concerns about a limbo space in their lives, caught between their first-language and needing to discuss their needs with public figures using a second language, between education and work, between their status as an asylum-seeker and settled status in the UK. In the pilot study, learners were asked what their career outcomes might be:

DH: With the help of my tutor and with the help of the college I want to find job. If I can't apply what I have learned in a job I consider myself that I have lost two years. For

example if I employed in cleaning. I have spent two years learning computer if I am not employed using computers then I have lost.

AM: Really, in the future, I'm going to use all these things I have learned here. If you're going to do a job like an accountant you need to be able to use the computer, to use all these kinds of software.

AH: I like it. And I told you that I want to work with my son. My son also learn web design skills and he's going to open an office and we're going to work together.

RS: To work in an office.

ST: To be honest, I'm still in a dilemma.

Many of the career objectives shared by the learners are ill-defined or generalised. There is a temporal sense of studying towards something, of becoming something else, yet their identity is still first language, first culture. 'Identities are about negotiating new subject positions at the crossroads of the past, present and future' (Block, 2007: 27). Second language proficiency and qualifications are perceived opportunities out of this crossroad: a transitional stage between a number of conditions: languages, official status, cultures, education, homes, etc.

Given this brief ethnographic and demographic consideration of ESOL learners, the participants in the present study could be considered to be representative of a type of migrant learner in the UK. They live and learn in a dense inner-city suburb which is characterised by a number of deprivation features. This is equally the finding of the HENNA project based on the area in which the study is located: 'Harehills is a superdiverse neighbourhood with a large multilingual population, many of whom have English language needs. In many ways the neighbourhood is emblematic of the bigger picture of superdiversity in Britain's cities' (Simpson et al, 2011b: 31). Language, literacy and technology are not *apolitical*, *asemiotic* resources, but cultural tools to progress into work and education: '- what Bourdieu (1986) termed *cultural capital*. This cultural capital includes previous education, language and literacy, a range of qualifications, skills, knowledge and prior experience' (Simpson, 2016, forthcoming). This sense of cultural capital to progress was voiced by all the learners in this research. The following is a biography of the individuals who are the main focus in the case study.

Sakia – pre-task interview (Appendix D)

Sakia comes from Poland. She is 26 years old and educated to university level in her country of birth. Like many migrants, she is over-qualified for the work she is currently doing. Sakia trained as a teacher of the Polish language but is working in Leeds as a receptionist. Asked why, she says teachers are low paid in Poland. She is confident with technology and uses it in all aspects of her social life and work role, including an iPhone, computer at work and home and in her studies. There is less confidence with her English; she talks of one day being able to be able 'to speak well, to write essays, to have normal English.' Competency in language and literacy are essential to her. Sakia is isolated from the Polish community but does not see this as a negative. Her only regular Polish contact in the UK is with Gamda. She uses technology to stay in regular contact with family in Poland. When asked why she uses computers, she states: 'To communicate with my family, friends, check emails, Skype. At work I have to use the computer at work for typing.' In general, Sakia is very confident and the most competent member of the group. In the classroom she sits besides Gamda, who is also a friend outside class, and often supports Gamda with her class work.

Gamda – pre-task interview (Appendix E)

Gamda comes from Poland. She is 30 years old and educated to university level in her country of birth. She was training to be an accountant in Poland but now works as a waitress in a restaurant. 'The money is better because you have tips.' English competency (spoken and written) is a concern:' I hope to be able to communicate without any problems. I try to do what I can to learn English. That will help me here and in my country.' She communicates with family in Poland using technology such as Skype. She also has an iPhone. There is a strong Polish community and identity in Leeds but Sakia and Gamda are isolated from this community 'I don't meet many Polish people.' Gamda is aware of non-linguistic modes brought about by technology: 'Most people are communicating through pictures through the internet.' Gamda is not confident with technology or her English so often draws on a number of resources to help her: electronic translator, online dictionaries, and her friend Sakia.

Darras – pre-task interview (Appendix F)

Darras comes from Iran. His first-language is Kurdish. He is 20 years old and educated to college level in his country of birth. He is not allowed to work. He arrived in England as an asylum seeker, a minor with virtually no English language and poor L1 literacy. He refers to himself as the 'farmer boy,' suggesting how far he has come from humble

beginnings to a city in a different country. Of all the research participants Darras lacks the most confidence and is very deferential, or perhaps, as he says of himself: 'very lazy.' His status has been unsettled for a number of years as he waits for a decision from the Home Office – if he will be granted 'leave to remain' or if he will be deported. He has a social worker and has been forced to move home to at least three council owned properties. The uncertainty of his position is clear, as he says: 'Believe me, I have many problems ... I still don't know my future. What I should do.' He has concerns about his English competency and technology but uses social networking sites in his social life to communicate with people around the world, including PalTalk, chat rooms and Skype, and understands the importance of technology for work. 'I like to improve my typing, get information and get a good job.' Darras first started using computers when he came to England. In the future he wants to work as a translator but recognises the challenge of a second language: 'That is my dream. I must improve my English.'

Shourok – pre-task interview (Appendix G)

Shourok comes from Iran. His first-language is Kurdish. He is not allowed to work. He is 25 years old and started university in his country of birth, studying business management for two years, but had to leave his country and come to the UK as an asylum seeker. When asked about his English competence his sense of frustration is clear: 'I get angry. It makes me nervous. And people they will not try to understand you.' The challenge of language for Shourok is usually with 'official' others: 'I know what I want to say but I can't say it in English sometimes. I have some problems with my home. Telephone bills and things like that.' When asked about a sense of community where he lives, Shourok identified multiculturalism as a problem, particularly because he lives in a dense area in Gipton: 'The problem there is many multicultural people, for example, your neighbours are from other countries. It's too hard to talk to them.' However, technology is the tool he uses to connect with a wider Kurdish community, as well as other people, using Messenger: 'I've got nearly 50 friends. Some of them are Iranian. And others in different countries that are a long way from here. Australia. Canada.' The affordance of different modal resources, mediated by technology, was identified by the New London Group (Cope et al, 2000). Shourok equally recognises the potential of different communicative modes because his English is not as competent as he would like it to be: 'If you look at my website its name is iran4all. It gives me a different way to 'talk' to others. If you can't say it in one way you can say in a different way. It's like Rimi said, if you can't speak in good English you

can say it in another way, for example, designing a protest image and saying it that way. If we don't like something and we can't say it we can communicate through images or file; something like that.' When asked about the future and what was important to him, there is a clear sense of trying to pick up where he abruptly left off: 'I really want to go to university and I need English.'

The remaining learners in the case study will not feature in any discussion or analysis. The rationale for this is raised in section 4.6.

The concern of the New London Group for a pedagogy of multiliteracies was in response to increasing cultural and linguistic diversity at community level as a result of globalisation (Cope et al, 2000: 9-10). We see these same issues being raised by the research participants as they discuss the three realms of (1) their working lives (2) public lives and (3) personal lives (Cope et al, 2000: 10-17). Language, literacy and technology cuts across all these dimensions.

The opening chapter of the multiliteracies strategy provides a rationale for the pedagogy. This chapter originally appeared in the Harvard Educational Review (Cope & Kalantzis, 1996). It is a dense piece of text defining terms and processes. Figure 3-2 overleaf is an attempt to simplify its main features. The focus is very much on 'social futures' and what can be done at classroom level to help diverse groups of individuals become literate in a range of text types (paper, electronic, live) so that they can consume and produce meanings in and across different semiotic resources with one eye on the future so that individuals leave the classroom with the ability to respond and adapt to evolving communicative changes. To be multiliterate then is to engage communicatively in the multiple social spaces which populate the three domains of work, public citizenship and personal lifeworlds.

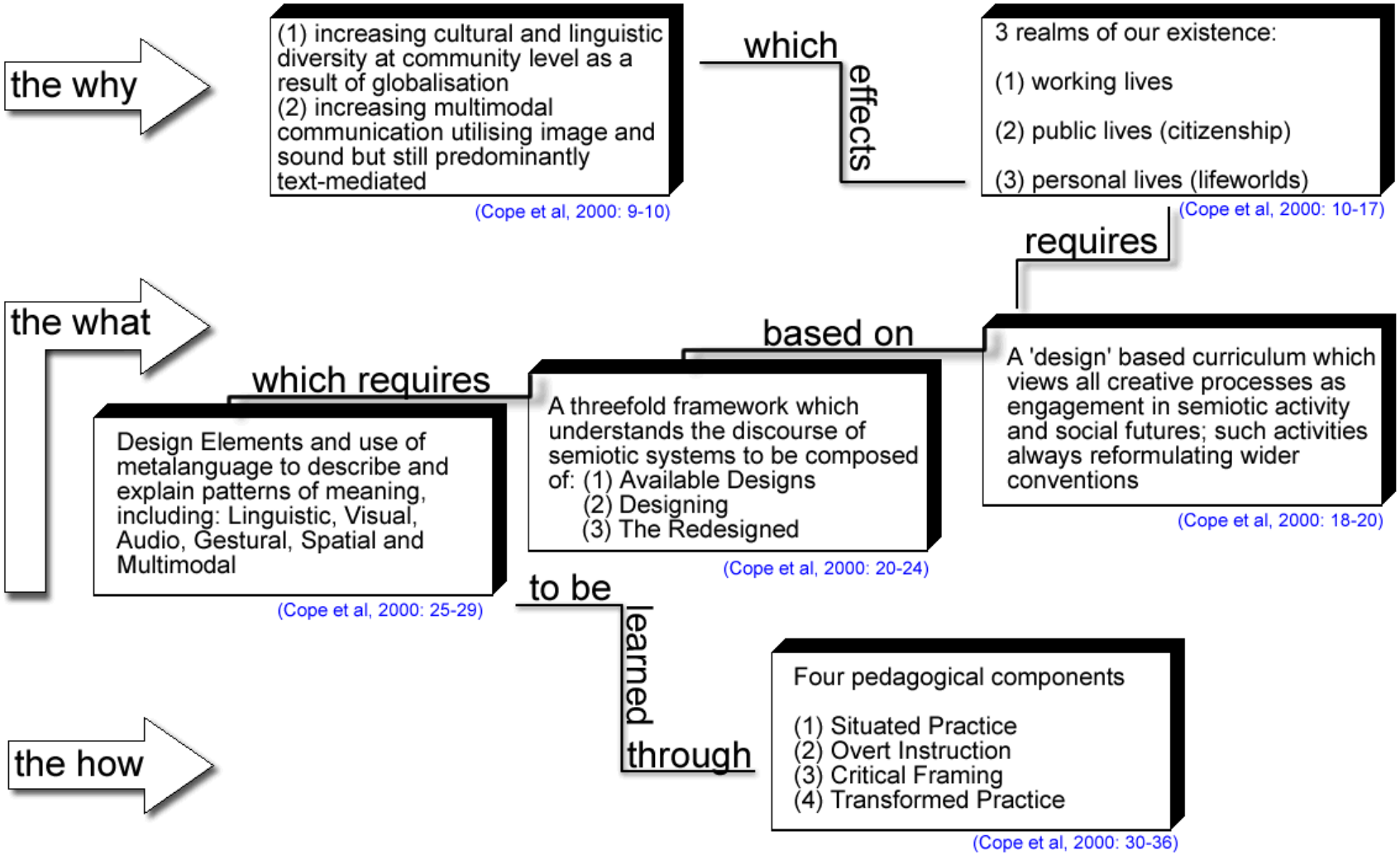


Figure 3-2 (Multiliteracies diagram

The HENNA project explains that some 'ESOL students have little or no literacy, either in their expert languages or in English' (Simpson et al, 2011b: 10). A point of concern in the multiliteracies strategy is, yes, a call for curriculum change to meet cultural and linguistic diversity and the emergence of multimodal communication, but a problem is dominant semiotic systems biased as they inevitably are in their hegemonic cultural and linguistic configurations. Research with minority children in Australia suggests that the very groups trying to be included in this pedagogy are those who struggle because the dominant culture and language is incongruent with minority languages: the distance between 'experiencing the known and the new' was too great without scaffolding from the teacher (Mills, 2006: 26).

Another point of concern is how many years of support an ESOL adult requires to truly be multiliterate considering that a significant number of ESOL learners come to the UK with no competence in English (Simpson, 2007) and with limited educational experience in their L1, typically with no experience of technology, in a funding system which limits the learning hours per week and the levels they can aspire to before a cost to the individual is applied. Even if an ESOL learner is willing to pay, for example, to study on a vocational course at Level 3 or at university at Level 4, they risk linguistic penalisation through mandatory language tests prior to enrolment; what has been termed elsewhere as 'linguistic gate-keeping processes, either implicitly or explicitly designed to control access' to work or course related opportunities (Roberts et al, 2007: 22). In many ESOL classrooms, students quickly arrive at the end of their formal learning journey. Government funding and a dominant language system at institutional levels hinders further progression. One would hope that individuals, after leaving formal education, are able to develop their own skills in dominant multiliteracies in the domains of work and citizenship. To what extent that happens is unclear.

3.3 piloting phase

The pilot study trialled different methods of data collection to gather student information on their 'actual' and 'perceived' digital and literacy competency. The intention was to identify themes relevant to the group. The research methodology might be described as *grounded* or *cyclical*: 'cyclical … research often follows an inductive path that begins with few perceived notions, followed by a gradual fine-tuning and narrowing of focus' (Mackey & Gass, 2005: 163).

In the piloting phase, a mixture of qualitative methods (interviews) and quantitative methods (questionnaires) were used. Analysis of the data intended to look for patterns, or themes, to allow 'a gradual fine-tuning.' This revealed problems with both the research tools and the research approach. The research *had* perceived notions, particularly about competency. Unintentionally, the epistemological position was found to be reductionist because it positioned the learners as deficit, lacking in cultural, visual, digital and literacy competency. Research into the demographic and biographical profile of the learners, as out of work, struggling to find work, finding language barriers, etc, inadvertently fuelled a reductionist orientation to the research. As discussed in section 2.2, functional literacy perspectives position the learner as deficient, lacking in skills, and makes a correlation between a lack of competency and poverty, prior education and employability (e.g. Denny, 2000; Crowther et al, 2001: 23-41; DfEE, 2001; Papen, 2005: 120). The focus is on what the individual cannot do, the reasons for this and the impact for the individual and the potential burden on wider society. The early stages of the research bought into the 'deficit myth' (Gregory and Williams, 2000) by accepting a competency model; what a 'standard' or 'native' individual is able to do, and then seeking to measure a lack of competency on behalf of the 'non-standard' or 'non-native' individual, i.e. the ESOL learner. The following discussion explains the path of logic that led to this epistemological position and research approach which subsequently changed following the pilot study.

Within second-language research, *communicative competence* has been regarded as a primary objective of second-language teaching (Hymes, 1972; Canale and Swain, 1980; Bachman, 1990; Cook, 2003). The linguistic system has been a dominant research focus in attempts to operationalise *communicative competence* but tutors and learners now consume and produce meanings across a range of communicative modes. Multimodal forms of digital communication call for an improved understanding of what is meant by *English* and *Literacy* and how such subjects might be taught, what has been called *multimodal communicative competence* as an extension of communicative competence (Royce, 2002; Royce & Bowcher 2007; Rassool, 1999).

> The conventional notion of what it is to be communicatively competent is predicated on a near equation of communication with face-to-face interaction that has become undermined by the growth in mediated communication of the past decades (Kenning, 2006: 364).

Kenning finds traditional communicative competence in need of updating as technological forms of mediation grow more prevalent. *Multimodal communicative competence* as an overarching objective might be seen as dependent on two sets of skills, digital and semiotic competence, each of which have their own sub-components.

multimodal communicative competence	
digital competency	**semiotic competency**
access	spatial
manage	gestural
integrate	auditory
evaluate	visual
create	linguistic

Table 3-1 (multimodal communicative competence)

These two competencies (digital and semiotic) were misguided 'assumptions' on my part. First, the sub-components of *digital* competence 'represent a continuum of skills and knowledge and are presented in a sequence suggesting increasing cognitive complexity' (ETS, 2002: 17). I began to ask through the research methods: can second-language migrants access information in digital format, which also happens to be in a second language, and can they evaluate the quality and relevance of that information? Finally, at the more complex end of the continuum, can learners reciprocate in the act of digital communication by creating their own information? The simple answer, retrospectively, is yes, they can. What the pilot study discovered is that the ability to access and manage information was not of significant interest. The learners in the pilot study were all competent users of technology, with a mixture of experiences and preferences. The information gained was peripheral to the interest in embodied peer-interaction with technology. The research approach changed to prioritise the real-time act of *creation*, how the learners used technology with literacy and language, irrespective of a perceived digital competency. The challenge, discussed later, became one of how to capture the process of peer-interaction in real-time.

Second, the original research approach believed that for migrant learners of English there is the potential for reduced visual competency. The increasing production of multimodal communication in digital content requires a pluralistic notion of semiotic resources. 'From a multimodal perspective *all* modes contribute to learning' (Jewitt, 2006: 27). If this is true, then I believed there was a requirement on educators to ensure

all modes *do* contribute to learning and that *some* modes do not reduce learning. Typically text-image configurations are most prevalent in multimodal communication and receive most discussion in the literature, but rarely does such discussion engage with second-language migrants and how they interact with bimodal configurations. 'Additional research to develop tools that measure an individual's degree of L2 visuality, including skills of creating and interpreting visual language, is important in evaluating the overall impact on student learning' (Stokes, 2006: 17). I believed L2 learners lacked visual competency and so I attempted to measure that.

This two-fold 'deficit' approach was indeed a myth. It is self-evident that individuals in a new culture and using a new language will have problems decoding meaning. As Kress states: 'the semiotic reach of modes … is always specific and partial in any one culture, though differently specific and partial' (2009: 57). The 'reach' of modes across cultures is a universal problem. Problems in decoding meaning became the central focus of the research and this was a wrong approach. Adults are 'active designers of meaning' (Cope & Kalantzis, 2000: 7). Taking this approach, that learners are 'active' *meaning-makers*, rather than attempting to measure how they are 'getting it wrong,' was pivotal in changing my epistemological values and ultimately the research approach.

The pilot study used the data collection methods below. Table 3-2 identifies the 'method' used, the 'rationale' for using that method and what the 'findings' were.

	method	**rationale**	**findings**
1	A biographical questionnaire to identify previous education, languages and exposure to English.	The two biographical questionnaires were used to identify variables which might account for differences in digital and semiotic competency; for example, do learners with more sophisticated second-language skills develop digital competence more quickly? How much of an indicator is previous educational experience, etc?	Evidence indicates that second-language competence has little correlation with digital competence. Some learners with limited English showed greater digital competence compared to more advanced language learners but more competent language learners developed better IT language specialisation. Prior educational attainment from a first country was not a guarantor of communicative competence in a second language.
2	A biographical questionnaire to identify existing IT use.		

3	Learners were given tasks to design an image, a leaflet, a website and free writing.	Completed designs are evidence of learning and a measure of digital and semiotic competence. Do the learners understand a concept? Can they articulate that concept multimodally though text, image, colour, etc? What are the challenges and affordances of communicating through different modes and programs? Do the learners have modal preferences?	The learners believed they had greater difficulty with text-based forms of communication compared to visual. There was often an avoidance of text in preference of image but later discussion with the learners evidenced confusion with what some of the images actually meant.
4	A Likert scale survey measuring the learners' attitudes towards the different programs used.	Identify learners' preferences for one design software over another and to identify any challenges for the learners across the range of design software.	Evidence shows learners preferred image based media and had greater confidence in web based technologies but many had less confidence and some anxiety about literacy based communication.
5	Semi-structured interviews to talk about the students' experiences, modal preferences and understanding of the different modes and programs used.	Identify any challenges for the learners in creating the designs. Identify some of the colour, word, image, layout choices made by the learners. Identify preferences for one design over another.	Language anxiety and concerns about writing were the main findings. In addition, when asked, many of the learners couldn't explain certain aspects of their images.
6	Comprehension questions in relation to a cartoon image. Two groups were used: so called 'native' speakers compared to second-language speakers.	Using a culturally biased image, could second-language adults successfully answer comprehension questions revealing visual competency compared to 'native' speakers?	The ESOL adults had enormous difficulties in interpreting what to the 'native' speakers was a very simple cartoon image.

Table 3-2 (pilot study data collection methods)

These methods were used because the piloting phase bought into the *narrative* of the 'deficit myth' by accepting a competency model of individual ability. This erroneous essentialising of the learner in the piloting phase helped to refocus the research and reorientate my epistemological position in relation to what was most important to know. How the learners manage their own learning, not how accurate are they, became the focus.

As a consequence of the piloting experience, the research shifted significantly from a product based approach to a process based approach; it became less interested in *what* the learners were producing and more interested in *how* they were producing with a focus on second-language writing as this was a recurring theme raised by the learners. In the revised research, methods two, four, five and six were dropped and a semi-structured interview used instead, asking five basic questions of each learner to generate biographical information (Appendices D to G). Concerns with interviewing in the research are raised in the next section. Method three was kept, but instead of looking at the designs and asking students to discuss their design choices, the research sought to understand the journey not the destination by providing the learners with tasks relevant to what they were studying (Appendix J). A post-task interview asked them about their experience of shared collaboration with technology (Appendices H and I). The new focus on understanding the learning journey brought with it significant data-management and analytical challenges, discussed in later sections of this chapter.

3.4 interviewing

Whilst videoing was the primary method of the data-collection, interviewing was a secondary method, in the piloting stage and the research proper (Appendices D – I). This section will reflexively discuss those experiences with reference to the literature on interviewing and with consideration of reliability and validity. There is some consensus that validity and reliability are not relevant in a qualitative research paradigm: 'Reliability and validity are tools of an essentially positivist epistemology' (Winter, 2000: 10). And further comments such as reliability and validity 'defined in quantitative terms may not apply to the qualitative research paradigm' (Golafshani, 2003: 600; see also Glesne & Peshkin, 1992; Stenbacka, 2001; Altheide & Johnson, 1998; Leininger, 1994). I lack confidence in the certainty of other qualitative researchers who claim that reliability and validity are not relevant. Regardless of the methodological paradigm in which one is positioned, the research process is ultimately about interpretation and explanation so 'the goal of finding plausible and credible outcome explanations is central to all research' (Morse et al, 2002: 3).

Videography (Knoblauch, 2012) was the primary data-collection method, discussed in sections 3.6 and 3.7, and interviewing was secondary. The experience of interviewing raised many questions about the accuracy of my interpretations so reliability and validity were considered to be important dimensions. The conventional perception of

the interviewer as neutral was found to be misleading. Rather than passivity, the research literature now recognises the 'fact that interactional, interpretive activity is a hallmark of all interviews' (Holstein & Gubrium, 2004: 140). There are many pitfalls, as Watson succinctly captures: 'we are all – researchers and researched – unreliable narrators' (2006: 367). Both parties are potentially liable in distorting the process, for example, because of: (1.) 'hierarchical power distributions between interviewer and interviewee' (Glesne & Peshkin, 1992); and (2.) 'bias, error, misunderstanding, or misdirection' (Holstein & Gubrium, 2004: 141); and (3.) unskilled 'novice interviewers' (Roulston et al, 2003).

These, and many more potential *contamination* sources, should give concern. For some there is agreement that interview as methodology has moved away from a model of neutrality: 'This reframing of the interview process has significant implications for such issues as "objectivity", "validity" and "reliability"' (Watson, 2006: 368). The three potential impacts on reliability highlighted above (power, bias, skill) are raised here because they were evident in my interviewing. It was not until I transcribed the data that I realised how inconsistent and how useless I was as an interviewer. As a quick example, I had thirteen questions in a structured interview but for some reason I did not ask all the interviewees the same questions. Some I asked all, some only nine or ten, some even less; and most of them I paraphrased. I do not know why this happened but it is perhaps an indicator of the interactional, very spontaneous nature of interviewing. Decisions are made in an instant and the path from question one to the last question is not rigid but fluid. I believe I was *seduced* away from the map of questions on the paper because 'interactions and meanings are a shifting carnival of ambiguous complexity' (Scheurich, 1995: 243). Perhaps the process of meaning-making is what dictates the direction of the interview path taken because the reciprocal to-and-fro of information is subtly negotiated whenever there is interaction between two parties: 'all interviews are active interviews' (Holstein & Gubrium, 2004: 140). I tried to be rigid and 'quantitative' in my interview approach and it did not work.

Another reason why there was such fluidity is because of interviewer inexperience. In this example, one of many, I am asking what I think is an obvious question. When the learner answers incorrectly I skip forward and drop further questions:

SW: Would it make a difference if the instructions were in Vietnamese?

> TH: I think in my country is important because they don't learn like Dreamweaver. They just learn to type.

How has TH interpreted what I am asking? Is the question ambiguous? In other interviews I ask this question much more clearly: "Would it make a difference **in the classroom** if the instructions were in your first language?" Here I have quantified what is meant by 'instructions' with 'classroom' but not in the question asked of TH. She has supplied her own context. Another example from the same learner:

> SW: What do you think the purpose of your website might be?
> TH: I think the purpose of this is when you eating the apple, it already miss a bit here.

Her answer here has no correlation with the question. Again, what knowledge am I assuming TH has? Given there is a lack of clarity in the phrasing of the questions and that there is no parity in the amount of questions asked across the respondents, it did raise the question of reliability. Clearly I am not a neutral participant in the interview process. If anything, I am as active as the interviewee; maybe more so.

It might be however that the question above is not ambiguous, but instead, the language level I am using is not appropriate to the language ability of the individual. Complications can arise in second-language interviews when the interviewer has full competence in the dominant language and the interviewee has less competence. This raises issues with *linguistic equality* and *cultural bias*. For example: 'the rules of the interview as a language game serve to construct' the social reality of the interviewee (Watson, 2006: 368). Watson is alluding to the fact that interviews are not a natural situation; they are an artificial setting, even ritualistic with expected roles, modes of behaviour and language phrases. If I am attempting to capture an *approximation of accuracy* by talking to respondents, there should be concern with how the interview as a ritualised event can skew the reliability of the research instruments. An example of reduced language capability and associated frustrations is provided by Woodrow, who found that 'the most frequent source of anxiety was interacting with native speakers' (2006: 308). I was forced to question the reliability of my learner interviewing and subsequently the validity of my interpretations based on their answers. To try and provide some validity, member-checking was used as a method of iterative transaction between *researched* and *researcher*.

One way of trying to ensure reliability is to incorporate into the data collection a 'recursive, process-oriented view of validity' (Cho & Trent, 2006: 319). Figure 3-3 is an illustration of a process approach to data collection with iterative interaction between researcher and the researched. Transactional validity as process 'consists of techniques or methods by which misunderstandings can be adjusted and thus fixed. In most cases informants are engaged in making sure their realities correspond with the interpretations brought forth by the researchers' (Cho & Trent, 2006: 322).

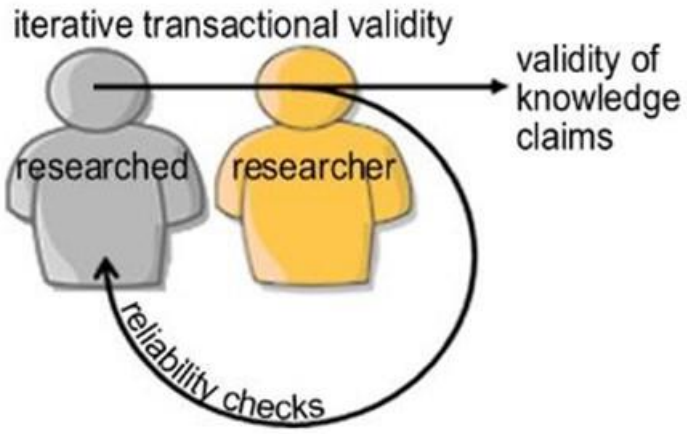


Figure 3-3 (member-checking of interview data)

Figure 3-3 implements cyclical reliability checks to ensure the 'techniques or methods' used are accurately capturing the 'correct' knowledge. The *researched* shares information with the *researcher* and at some point the shared information is verified with the *researched*. Iterative methods of data collection might therefore be regarded as a scaffold for successive approximations of accuracy. We are here looking for formative verification that the data we have captured is accurate and thus valid. Methods of reliability checks discussed in the research literature include member-checking (Morse et al, 2002), triangulation (Bryman, 2007), self-reflexivity (Richardson, 1997), process notes (Campbell, 1996), decision trail (Sandelowski, 1986). An iterative methodology is commensurate with the qualitative paradigm because 'qualitative research is particularly useful for studying phenomenon or event(s) about which little is known' (Appleton, 1995: 993). When we do not fully understand a given situation we need a research methodology which is flexible rather than fixed. All the *reliability* methods itemised above are not without their problems but member-checking, reflexivity and triangulation were used in this research.

With a transcription of each interview I used member-checking as a validity tool. This was a print-out of the interview transcript for each respondent where they could talk about what they had said and I discussed issues to do with language ability and language anxiety. When asked to comment on their experience of the interview I received the following replies:

> OD: I felt happy and enjoyable.
> RN: It was great because it give me more experience about interview.
> AM: The interview I have had is really good and satisfying.
> BO: It was normal.
> AH: I was relaxed and gave me confidence about myself.
> DH: It was excellent. I feel happy.
> MT: I felt a little bit nervous.
> RS: I was a bit of nervous but it was a good experience.
> OD: The bad experience is the interview recorded. You don't know where will be taken and who will listen it! Make you scared, afraid also.

As can be seen, most were happy with the experience and did not express any anxiety. When asked about their language use during the interview there was not the same level of confidence:

> RN: I feel little bit sad because I did lots mistakes about my word when I speak.
> AM: I need to improve my grammar and learn new words.
> BO: These are my ideas but I would prefer some changes to the grammar.
> DH: My English was not as good as required.

An interesting redistribution of power was evident in the interview, indicating the traditional model of interviewees as 'passive vessels of answers' is untenable (Holstein & Gubrium, 2004: 144):

> SW: What do you see yourself doing in the future?
>
> DH: This is the most important question. Since September I have studied. **It's my question as well to you. Is it helpful for job, what I'm doing in class?**
>
> SW: Yes, a lot of the programs we're using, any office environment will use all the programs you've been learning.
>
> DH: **So you answer my question now let me answer.** In the future I would like to apply what I am learning at job.

The sentences in bold are of most interest; DH is not an empty vessel, he is an 'active agent' in the interview: listening, responding and probing the interviewer by asking his

own questions. He even reverses the power distribution by confirming he will now answer my question because I answered his: **So you answer my question now let me answer.** It is perhaps a mistake to assume that in the interviewer/interviewee role that one is passive and the other is active: 'the interview can be thought of as a collaborative construction' (Watson, 2006: 369). All of the interviewees were given the opportunity to 'member-check' their responses.

On the whole, in terms of reliability and validity, some concerns have been raised about my interviewing skills. I do not believe this makes the data less valid, particularly as cross-referencing with member-checks showed agreement with the accuracy, but it does mean that self-reflexivity and rejection of neutrality is important: 'full disclosure of the researcher self' (Pole 2003: 19) is fundamental to an ethnomethodological approach to data-collection. Whilst I have advocated member-checking as a transactional process to support reliability and ultimately validity, an alternative view is suggested by Nunkoosing: 'the intellectual rigor and validity of our interpretations have to meet with the requirements of the research community rather than the agreement of the people we interview' (2005: 699). In response to seeing my interview questionnaire sheet, it was recommended by one of my supervisors that I adopt a more themed approach. A CD on ethnographic interviews/conversations was loaned with a guidance sheet that can only be described as commandments for novice researchers:

- 'start from ignorance'
- 'start without hypotheses'
- 'ask as few questions as possible'
- 'listen and wait'.

This became my revised approach to interviewing. In trying to answer what is meant by reliability and validity I find I have to engage with fundamental questions about the epistemological status of myself and my research community. Depending on how one positions oneself in quantitative, qualitative or mixed-methods, each will influence how reliability and validity are interpreted. The discussion in this section has been very much a reflexive negotiation of researcher-self. What is at stake is 'methodological identity' (Gored et al, 2004: 383) - discussed in the next section.

3.5 ethnomethodological case study

This research sits within a practitioner-researcher case study model. The study developed from an observation of my own learners and my own uncertainty with the complexity of how they were collaborating at a computer and orchestrating the learning between them. Whilst this could be considered a 'convenient sample' for research purposes, it is however a very relevant 'sample' as it was these 'types' of learners and their contexts which stimulated the enquiry. This positions the group very neatly as a 'case.' With four pairs of learners, and a focus on embodied peer-interaction, this provides four potential cases for exploration. This section will contextualise and problematise a case study model, including strengths and weaknesses.

In one sense the methodology and analysis might be considered transdisciplinary because the current study applies methods and orientation from a range of related research disciplines: i.e. case study, ethnography and grounded theory. A *case study* approach is provided because there is a clearly defined group in a contextualised setting 'which involves an empirical investigation of a particular contemporary phenomenon within its real life context' (as Robson defines case study, 2002: 178). *Ethnographic* in the sense that depth rather than breadth is more important so fewer cases are more feasible to allow understanding of a new area; including ethnographic methods of participant observation and interviewing with a demographic appreciation of the particpants. *Grounded* because theorising of peer-interaction and collaborative writing are outcomes obtained from the research participants; grounded theory is an analytical approach and the microanalysis of rich detail is the method used in this research. An *a priori* theoretical position, as much as it is feasible to not have pre-existing theories and assumptions, has been the methodological orientation in the data-collection and is an ontological position shared by all these research approaches. Robson explains that whilst 'ethnography is a distinctive approach, it can be linked with either the case study or grounded theory approaches. A case study can be approached ethnographically; or an ethnographic study can be approached by means of grounded theory' 2002: 190). In the sense of interdisciplinary research methods, *ethnomethodological* might be a more appropriate overarching term to describing the ontological and epistemological position in the current study. Ethnomethodology is not so much a methodology, contends Rawls, but any series of research methods used 'to discover the things that persons in particular situations *do*, the methods *they* use, to create the patterned orderliness of social life' (2002: 4). This rationale chimes very well with the current study looking at a particular

group of learners, what they do and how they do it within a framework of peer-interaction. Because of this evident overlap in research approaches, the term *ethnomethodological case study* is suggested with recognition that it includes research and analytical methods commensurate with elements of ethnography and grounded theory. Common features across the methodologies include:

- a located, specific setting
- access to insider views
- participant observation and interview as primary research methods
- particularisation of the data rather than generalisation
- holistic description of interaction within the location
- reflexive substantiation of the researcher's interpretations and processes

These characteristics (Pole 2002: 3) belong to an emic perspective. An *ethnomethodological case study* prioritises qualitative methods within a broader methodology of interpretivism. The nature of the knowledge to be learned is particular and brokered between the researcher and the researched. However, gaining access to insider views is a process not without consequences: 'Confinement to experience-near (i.e. emic) concepts leaves an ethnographer awash in immediacies as well as entangled in vernacular. Confinement to experience-distant (i.e. etic) ones leaves him stranded in abstractions and smothered in jargon' (Geertz, 1976: 223). Being too close can be overpowering for the researcher. Being too far can be de-powering for the subjects, as discussed in many literacy studies; Street (1993, 1995); Barton et al (2000); Papen (2005).

Methodological identity is one formed by the experiences and affinities of the researcher. Sections 3.2 to 3.3 detailed the biographic and demographic contexts of the learners, which within an *ethnomethodological case study* approach towards a single group, require a research approach appropriate to the types of learners and the types of questions being asked. A more formal quantitative approach, experimented with in the pilot study, was deemed to be inappropriate. '… qualitative methods of research may be seen to equate to methodologies based on an epistemological tradition of interpretivism, more quantitative methods may be seen in broad terms to be applied to positivist traditions' (Pole, 2003: 7). Does a researcher believe that social reality is structured, predictable and that individuals are of a type, hence behaviours and knowledge can be generalised and patterned given enough sample data? Here a positivist methodology would seek data from as wide a cross-section of people as possible to provide a sample

from which uniformity can be recognised and generalisability made. Structured interviews, surveys, closed questionnaires, etc, are methods of data collection which are commensurate with a positivist methodology and a deterministic epistemology. Explanation is more important than description. Predictability and constancy are key thoughts. Some would counter-argue however that in a world where 'knowledge is socially constructed, empirical methods and "objectivity" have only marginal epistemological status' (Ramanathan 1999: 45). Table 3-3 categorises 'qualitative-quantitative dimensions' (Tomlinson 2005: 1). Each polarisation quickly illuminates positivist/interpretive perspectives which might be applied to an *ethnomethodological case study* approach in a qualitative paradigm.

quantitative	qualitative
Numeric dimensions	Non-numeric categories
Replications	Single case
Descriptive statistics	Verbal accounts
Inferential statistics	Non-generalisation
Breadth	Depth
Artificial	Naturalistic
Structured /Pre-defined	Open/grounded
Hypothesis-testing	Exploratory
Positivist	Interpretivist
Absolutist	Relativist
Realist	Phenomenalist / idealist

Table 3-3 (qualitative-quantitative dimensions)

A qualitative approach 'emphasises the role of people as active agents in the construction and negotiation of social reality, it privileges research methods such as interviews and participant observation over structural analysis' (Papen, 2005: 61). Features of structural analysis are evident in the quantitative column. An interpretive epistemology recognises the complexity of human experience and the difficulty of generalisability. Observing and talking to people can sometimes be the only way to understand individual cases. To seek a common causality in behaviour is it to look for homogeneity where it may not exist. Positivist methods look for similarity but some forms of behaviour have complex causal relationships and require heterogeneous explanations. An *ethnomethodological case study* approach is applicable in such contexts.

Yin explains: a case study is 'an empirical enquiry that investigates a contemporary phenomenon in depth and within its real-life context' (2009: 14). The setting for the current research is a classroom session and is considered to be naturalistic and 'real-life' because the learners are attending a normal classroom session and engaging in learning activities which are a regular weekly occurrence. The learners were not asked to do anything they would not normally do. The recording devices were background media as the learners carried on in their normal learning patterns. In this sense the approach has ethnographic characteristics (participant observation through video, a naturalistic setting, talking to the participants) though clearly without the temporal element of a longitudinal approach and the more natural environment of their immediate homes and communities. *Natural* social interactions with learning though is highly relevant to the current research and the use of 'video analysis, in this sense, is *ethnographic* … By natural settings it is meant that, typically … researchers do not try to create the situations they study but attempt to record interactions where and how they are assumed to happen' (Knoblauch, 2012: 252). Arriving at a 'thick description' (Geertz, 1973) of how pairs of learners collaborate is the intended outcome of the research.

Hodkinson & Hodkinson state that a strength of a case study approach is that it can be grounded in *lived reality*. Although 'social research simplifies the phenomena investigated … case studies can do this in ways that strongly relate to the experiences of individuals, small groups, or organizations. They retain more of the "noise" of real life than many other types of research' (2001: 3). Sometimes the 'noise' might be part of the thick description which makes detailed findings more valid. As will be discussed in sections 3.6 and 3.7, capturing every action, reaction, every gesture, every word, of every individual, on-screen and off-screen across a three hour classroom session, generates a large amount of data and context; as Geertz suggests, leaving a researcher 'awash in immediacies.' An inductive approach to the observable phenomena is desirable to provide a holistic account of the cases. With the right methodological and analytical approach, it is possible to simplify the noise of the phenomena, not in a negative sense, but by identifying and naming patterns of embodied peer-interaction and social order.

The findings of an ethnomethodological case study are not easily generalisable because of the very situated number of small cases examined. This might be considered a

weakness. ‘For many researchers and others, this renders any case study findings as of little value’ (Hodkinson & Hodkinson, 2001: 10). Findings are non-numerical, and based on small numbers, but this can also provide a high level of detail which may lead to additional benefits: ‘case studies can facilitate rich conceptual/theoretical development’ (Hodkinson & Hodkinson, 2001: 8). Focussing on thick detail in underexplored areas can be fruitful ground for new/improved ways of thinking and doing. In the current research this has meant rethinking how to capture complex audio-visual data where there are significant chronemic and spatial complexities: the space of a full classroom, the embodied space of each learner, the space on-screen as a manifestation of what happens in the space off-screen, with actions unfolding across seconds, minutes and hours and across multiple participants. Successfully collecting (section 3.6) and beginning to work with that data (section 3.7) has meant rethinking how to manage and analyse such data (chapter 4). The methodology, or videography, is a contributory outcome from this research and this is because a very detailed case study approach has allowed for experimentation. Valsiner for example states that ‘the study of individual cases has always been the major … strategy in the advancement of knowledge about human beings’ (1986: 11).

Maoz suggests that the ‘case study absolves the author from any kind of methodological considerations. Case studies have become in many cases a synonym for freeform research where anything goes’ (2002: 164-165). In this section it has been shown that methodology is a very serious concern, including reliability and validity. An *ethnomethodological case study* rationale, in a qualitative framework, was provided as an amalgamation of overlapping methods and approaches from interdisciplinary research methods including the case study, ethnography, ethnomethodology and grounded theory. It was explained how the methods of such approaches are commensurate with the research participants and the research focus of the current study. The following section explores the collection of the video data.

3.6 data collection

The previous sections discussed the methodological and philosophical journey of the research process. This section will explain *how* the revised research was conducted and *what* data was collected. It will include a detailed description of all the steps taken and the methodological challenges of the different data collection methods. Capturing audio-visual data in ‘naturalistic’ settings as people go about their everyday normal

interaction is very challenging. For example, if people are static around a table, and know they are being recorded, it is possible to have a single audio-recorder (or camcorder) placed equidistant between all parties. In a normal classroom people are moving bodies so capturing every movement and every vocalisation of every individual is a significant logistical challenge.

Traditional 'pen and paper' methods of observation involved researchers observing actions in real-time without video technology. Coding methods were developed (or 'ethograms') in an attempt to objectively capture a range and frequency of behaviours: e.g. 'sits down' and 'looks left' or 'touches face.' Used predominantly in ethology (study of animal behaviour: 'licks paw' and 'bites') it has however been used extensively in social settings to describe the behaviour of groups and individuals in groups. The findings are frequently codified in specialist software such as Observer, though as Ice explains: such methods are 'not appropriate for collecting qualitative ethnographic data' (2004: 354).

The observational approach in this research is very much qualitative, and that adds a number of complexities. Knowing what to video, where to video, how to video, edit video and analyse video are new concerns in a relatively new methodology of analysing social interaction in natural settings. 'There is … a paucity of literature around a number of areas such as the links between research, video and software' (Woods & Dempster, 2011: 3). Mondada for example proposes a *praxeological approach to video practice* because '*analytical* studies focussing on a video as a timed accomplishment and as a social practice are still very few … focussing on the way in which videos are locally and contingently produced by social scientists' (2009: 51). As discussed by Woods & Dempster, and Mondada, research using video-as-observation requires discussion across a number of interrelated strands. 'Gradually a methodology is emerging that has come to be called succinctly video analysis' (Knoblauch, 2012: 251). And yet the term is too broad. Video analysis is prolific in many areas such as the social sciences, social media, film studies, multimodal discourse, etc. What is needed is the naming and defining of practice to identify what video analysis *is* and *is not,* in the field of specialism in which it is being applied. Knoblauch provides this:

> By natural settings it is meant that, typically (with some exceptions in applied research) researchers do not try to create the situations they study but attempt to record interactions where and how they are assumed to happen also without the

> researchers' interventions. This does not exclude the study of experiments – if experiments themselves are the subject matter of the study (and not the method). Audiovisual recordings and analyses of 'natural' social interactions demand that researchers go to 'where the action' is. That is to say that researchers of this type of video analysis regularly enter certain fields in an ethnographic way. Video analysis, in this sense, is *ethnographic*. In order to avoid the misleading notion of 'natural settings' and in order to stress the relevance of doing ethnography and in order not to confuse it with quantitative, standardized, and experimental forms of video analysis, it seems to me useful to call it videography. The preference for such a specification of the methodology seems to me even more urgent in the face of the recent video revolution (2012: 252).

Videography as research methodology can be defined as one which gives precedent to recording *natural social interactions* in an ethnographic manner with the intention of later analysis which itself adopts an a priori position to observable phenomena. The researcher as videographer, transcriber and analyser is common. Videography as praxeology demonstrates an interest in 'locally and contingently produced' videos created by researchers. There are practical and theoretical considerations. In the introduction to this chapter it was suggested that videography has four broad sequential strands: *collection, preparation, transcription, analysis*. First, *collecting* the video data is a praxeological concern discussed by Mondada (2009) - the situated, videographic challenge of deciding who, how and what to record. The practical dimension of situated videoing is the focus of this section. Second, *preparing* the completed video for analysis is a software challenge which inevitably has pre-analysis predilections: what to cut, what to trim, what to leave in, what is interesting and relevant, what is not; all of these decisions sit outside the theoretical framing of the later analysis but they have theoretical relevance. The challenge and practicalities of preparing data is discussed in section 3.7. *Transcription* and *analysis* are the third and fourth strands of videography and these have their own sections in chapter 4.

The first attempt at collecting the data did not work well. The learners were paired with someone who shared a first language. Participants were asked to speak in English, but if when supporting each other they could not express themselves in English, they could use their first language. Each learner was given a task: to design a website of their own interest and this was relevant as the learners were at the conclusion of a series of lessons on web design. The task was left deliberately loose, with limited instructions, to enable the learners to draw on their own resources and modal preferences. Figure 3-4 illustrates the configuration of learners and collection tools used. There were five data collection methods, explained below.

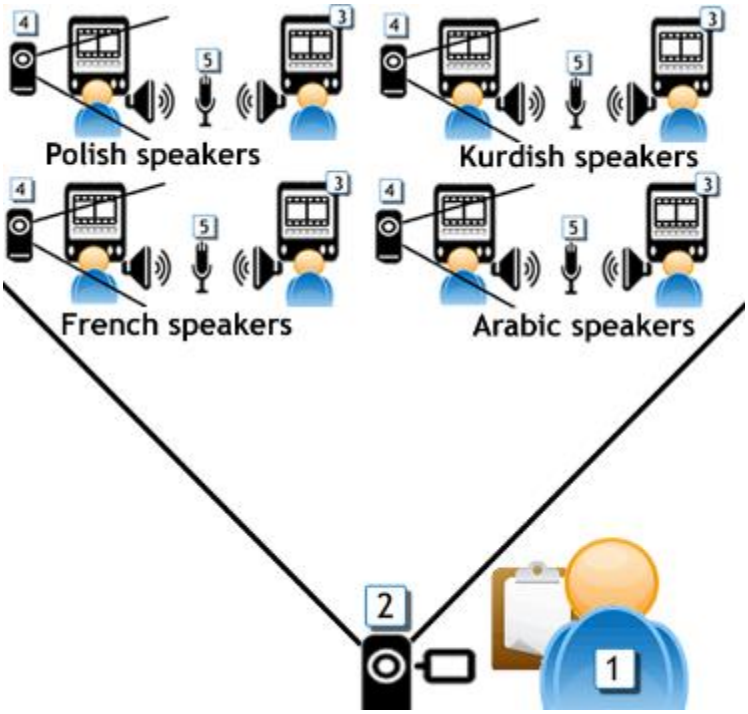


Figure 3-4 (first attempt at data collection)

1. the practitioner-researcher making notes,
2. a video camera to capture interaction in the classroom,
3. on-screen video recording software (Camtasia) to record mouse movement and everything the learner was doing, typing, clicking, websites and software used, etc,
4. a smaller video camera to capture action around the computer,
5. a voice recorder between the learners to capture any dialogue.

Each of these five methods revealed problems. Table 3-4 explains the problems and solutions following this attempt.

Problem	Solution
1. Because the researcher was also the practitioner it was very difficult to make notes. Whenever there was an issue a student would call for support and I was obliged to help.	Taking notes was not crucial to the research process. Better configuration of the recording media would capture nearly all the audio-visual interaction. These media could be played back later and thus freed me to monitor the recording devices.
2. The videoing at classroom level did not work well because there were two cameras which created blind spots in the middle of the classroom.	Use a fish eye lens and position the camera high in one corner of the room.
3. The Camtasia software records on-screen activity at a rate of fifteen frames per second, i.e. it records 15	Change the frame rate so that the Camtasia software records less frames per second. Five frames per second were found to be

	on-screen images per second which are later played back at normal speed. The computers in the classroom were older models. They struggled to process both the memory intensive videoing as well as the learner interactivity with the computer. Students were frustrated by the slowness of the computers.	optimal for the machines to process without slowing it down as the learner engaged with other activities.
4.	The smaller keyboard level video cameras were good but because the learners were working separately on their computers there was very little collaboration and interaction.	Pair the learners around one computer so that they collaborate to complete the task together.
5.	The single voice recorder between the two learners was intended to record any dialogue but there was little dialogue because the learners were working on their own tasks on a computer each. In addition, the recorder captured background humming from the computers, so much so that any dialogue could not be heard.	Give each learner their own voice recorder to hang around their neck with a lanyard. Create a joint task so that learners share a computer in peer-interaction.

Table 3-4 (problems and solutions with the first data collection)

A revised attempt at data collection was made, incorporating lessons learned from the first attempt. The changes made can be seen at Figure 3-5 and a discussion of these changes follows.

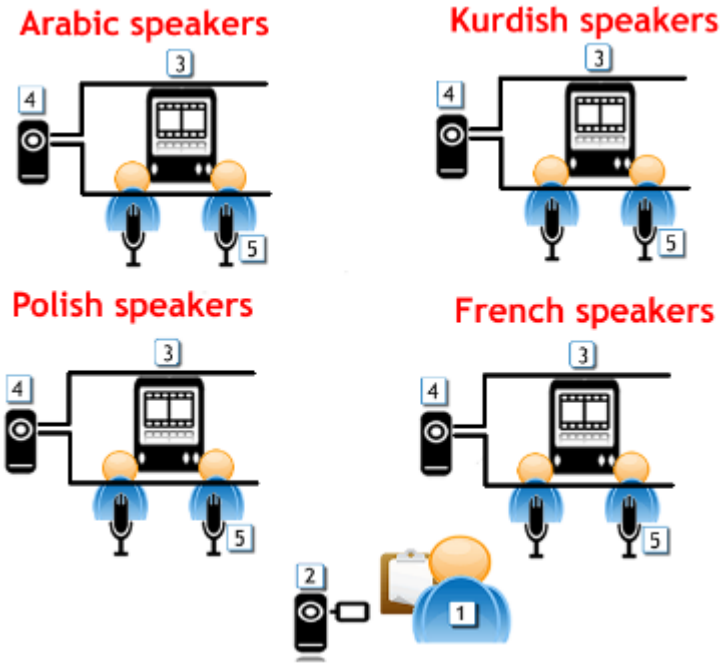


Figure 3-5 (second attempt at data collection)

1. researcher/practitioner

Originally, I thought I would observe and make notes during the data collection process in an 'ethogram' style. However, the positioning of multiple recording devices meant I did not need to observe in real-time as I could synchronise the recording devices later and playback learner interaction. All the recording equipment was configured prior to the beginning of the class. My role then as researcher/teacher was to introduce the topic for the lesson and brief the learners on what they needed to do. I set up the recording devices at methods 2, 3 and 4. I handed out one voice recorder per learner (method 5) and asked each learner to press record. I ensured throughout the session that all the devices were recording and then collected them at the end of the session, moving in and out of practitioner/researcher role.

The level of visual detail might be considered on three analytical levels: micro, meso and macro. The prior video attempts revealed gaps between camera angles when a learner was 'lost' on screen. It was uncertain what a tri-level detail of videoed interaction might reveal but it was intended that no individuals nor actions would be missed if every angle and corner was covered. This scale of detail, like the zoom lens on a camera, allows for *interaction-magnification* – to visually zoom in and out of the data to see interaction throughout the classroom (macro) and to see interaction at each computer (meso) and to see on-screen the gradual unfolding realisations (micro) of these higher-level actions. At some level this approach answers Goodwin's observation that 'In practice no record is completely adequate. Every camera position excludes other views of what is happening. The choice of where to place the camera is but the first in a long series of crucial analytical decisions' (2000b: 6). Multiple cameras is inclusionary. The praxeology of videographic practice is also discussed by Mondada (2009).

2. macro level of video data

A fish eye lens was added to a single camcorder and positioned high in a corner of the classroom. The aerial video was able to capture a macro level of interaction in the classroom. It recorded interaction beyond the meso level of video detail captured around the computer at method 4.

Figure 3-6 (macro level of visual data)

The image shown here is a thumbnail depiction. At full-size the video reveals much detail. Before analysis it was uncertain what a macro level of detail might provide. It was thought for example that people might unexpectedly enter the room and change the dynamics of interaction. Before the macro video, it was uncertain what the walk-patterns of the Learning Support Assitant (LSA) and tutor were as they moved around the classroom supporting the learners. If one of the paired learners left the proximity of the computer, recorded at method 4, it was unknown where that learner had gone. In the later analysis, the macro-level video proved valuable in a number of key moments such as when the LSA was stood behind a pair of learners, supporting them, but wasn't seen on the meso video at method 4. It also showed the extent to which some learners were favoured in walk patterns with the LSA or tutor spending more time with some learners than others. It showed how many times learners looked around them or raised a hand for support, and how many times no support was offered because the LSA or tutor were busy elsewhere. Given the focus of the research was on peer-interaction, discussed in section 2.5, this was not a concern. In these instances, the learners were able to solve the problem themselves.

3. on-screen video capturing the design process

Figure 3-7 shows a micro level of detail. Screen capture software has become very popular in training and education where an individual can record activity on the computer screen, provide an audio explanation, and save to a video format. Often these on-screen videos are produced to show other users how to navigate and use popular programs. These videos (or *screencasts*) are later downloaded or viewed online. In my own teaching practice I had been using Camtasia for a number of years to create screencasts for learners, showing and explaining how to use technology and programs. The software seemed ideal for capturing what the learners were doing on-screen, the real-time process of multimodal design. Geisler & Slattery for example note that 'video

screen capture has thus made visible phenomena that might otherwise have gone unnoticed in digital writing' (2007: 187).

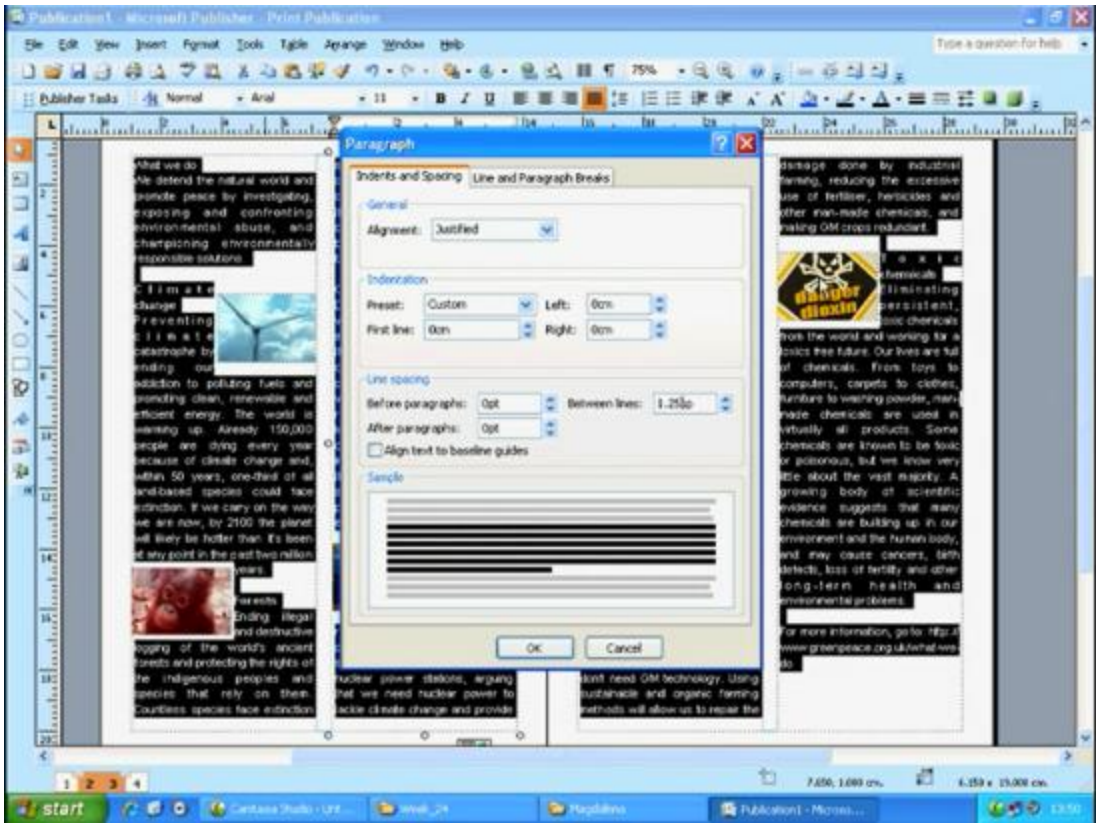


Figure 3-7 (micro level of visual data)

Previous research has used think-aloud protocols (TAPS) to try and capture in real-time what learners are doing (e.g. in L2 contexts see McDonough, 1995). This method can be useful for encouraging learners to meta-cognitively articulate what they are doing and why. Also called verbal-protocol analysis, Kuusela and Paul compare *concurrent* and *retrospective* data from individuals verbalising in-task and post-task on what they are doing (2000). Both have advantages and disadvantages *Concurrent* TAPS for example can be intrusive as it interrupts learners in real-time, breaking into a learner's stream of action, interaction, thought processes and can risk putting them 'off-task.' A *retrospective* method was loosely used in this research, asking the learners post-task about their experiences; (e.g. Appendices H and I).

Other research papers were found which similarly used screen-capture software to record learner actions on-screen:

> Process data were collected through recording on-screen activity with video screen-capture software … as the students collaborated on the online task. This

> enables us to capture screen activity and audio of the voices of the two students working at the computer simultaneously (Gardner & Levy, 2010: 2).

A significant absence from this data collection process was a meso or macro level of visual data, remedied here by methods 2 and 4; as the authors themselves admit: 'From an analytic perspective, the study would have benefitted from videoing of the participants' (2010: 2). Geisler & Slattery (2007) also used screen capture to explore the process of digital writing but did not include audio-visual methods for observing actions 'outside' the screen. Bhatt & de Roock (2014) and Bhatt (2012) used screen capture technology which included a webcam to record the 'granularity' of interaction on and around the screen with a focus on digital literacies.

Camtasia software was installed on every computer in the classroom. This is licensed, purchased software and an additional cost to consider. Free alternative versions are available but were found to be less functional and less stable. As a process-tracing methodology, screen capture software can record every movement on the screen and evidences the very fluid and organic nature of design: what internet searches are done; what images are found, chosen, selected, deselected, etc; and also, what impact other people have on the on-screen design process captured at 2 and 4.

4. interaction off-screen at the computer

Figure 3-8 (meso level of visual data)

When asking how learners collaborate in real-time there is a need to consider the physical domain and the symbiotic relationship between on-screen activity and off-screen activity. Small Disgo video devices were used to capture movement and discussion around the screens. These were used to evidence the learners' sociomaterial engagement with the computer and each other in the design process. Sometimes there

are significant modal configurations at the computer such as gesture with comments and tools. Layout can be a dominant mode in interaction. The proxemics of the learner interaction (with each other and with the computer) is mediated by a triangular layout; they both face the computer, shoulder-to-shoulder rather than face-to-face, so how does this layout mediate language as a communicative mode, if at all?

It might be that in some cases this triadic configuration constrains language use because it destabilises typical face-to-face interaction and its associated subordinate modes, e.g. eye gaze and gesture. Many classes will pair learners around one computer to encourage collaboration and talk. There is often a presumption that the participation frameworks structured by learners will be of benefit, educationally, conversationally and socially. The frameworks of participation might however be contested. This research sought to problematise such collaborative task-based methods. Inevitably, the configuration of modes will fluctuate throughout the interaction. The learners could at any point turn to face each other to talk. This might then demote layout as primary mediator and promote gaze. The learners might keyboard-switch and mouse-switch several times. Only a detailed level of videoed interaction could capture these real-time actions. Videoing can be intrusive but these recorders are very small, quiet and were positioned adjacent to each pair of learners.

5. voice recorders

A small voice-recorder was worn by each learner. It is possibly the most intrusive device used as it hung around their neck by a lanyard. Though audio was recorded at methods 2 and 4, there was a high risk of reduced sound-quality from these devices so a small voice-recorder was essential to capture any discussion. Eight students wore one each and also the tutor and LSA. Here the research was interested in how the subjects verbally negotiated the design process and how they supported each other. The learners were encouraged to use English (because of the later transcription and potential challenge of translating four different languages) but told to switch to their first-language if they needed to. Code-switching as a resource for learning is a strand of interest in this research: to what extent the learners *language* (Swain, 2006) and *translanguage* (Garćia, 2007.) A monolingual approach to communication in the classroom is a feature of many ESOL contexts, but as Simpson contends, 'ESOL learners are often surrounded by many languages, use several languages themselves, and move between them (*translanguage*) as a matter of course, or use English as a

lingua franca with speakers from diverse backgrounds' (forthcoming, 2016). A question to ask then of language diversity in the classroom, is for what reasons do learners switch to an alternative language? At critical points in the design process? To make social conversation? To make suggestions on improvements? To clarify? To request help? A conversational analysis approach to the audio data will be combined with an analysis of off-screen physical interaction, from the video streams at two and four, with the unfolding actions on-screen at video stream three, and thus provide a cross-modal microanalysis of embodied interaction at a number of close and distant levels, from spoken and print language(s) to embodied actions.

This section explained *how* the data-collection process was organised, including problems that occurred, and the different tools used to collect varying levels of audio-visual detail, from the macro to the meso to the micro. This scale and variety of visual data also supports triangulation. A change in interaction at keyboard level can sometimes be explained by an influence somewhere else in the classroom; for example, cross-fertilisation of ideas where one pair of learners embedded a YouTube video into their website. This was heard by another pair of learners and ultimately changed what they were doing. Four methods for data-collection have been described in this section: (1) classroom level video camera (2) keyboard level video camera (3) on-screen video (4) voice recorder per learner. A significant step in the research process was synchronising all this data in preparation for analysis. This synchronisation process is the topic of the next section.

3.7 preparing the data pre-analysis

This section will briefly explain how the different strands of audio-visual data were prepared for analysis and the difficulties which were found. Transcribing audio from a voice-recorder, whilst still challenging, presents one stream of data to work with. Using multiple streams of information from video and audio devices creates methodological and synchronisation problems. Woods & Dempster for example state that while 'analyzing text data can be complicated, the analysis of visual data adds a further level of complexity to qualitative research … how-to textbooks talk about variations on the theme of "drowning in data"' (2011: 2-3). The methodology of video analysis is the topic of the next chapter.

The management of the data in preparation for analysis proved to be very time-consuming. Multiple streams of audio-visual data from multiple recording devices require synchronisation. Each stream has a different start point. For every pair of learners there were five different devices collecting data. I chose to use video editing software to synchronise these; Corel VideoStudio 12 (X2 Pro) as shown at Figure 3-9. Compiling all the video and audio files into a single viewing space (picture-in-picture) was considered the best method to synchronise them for analysis. Most video editing software allows for multiple timelines and cutting and trimming of audio-visual footage.

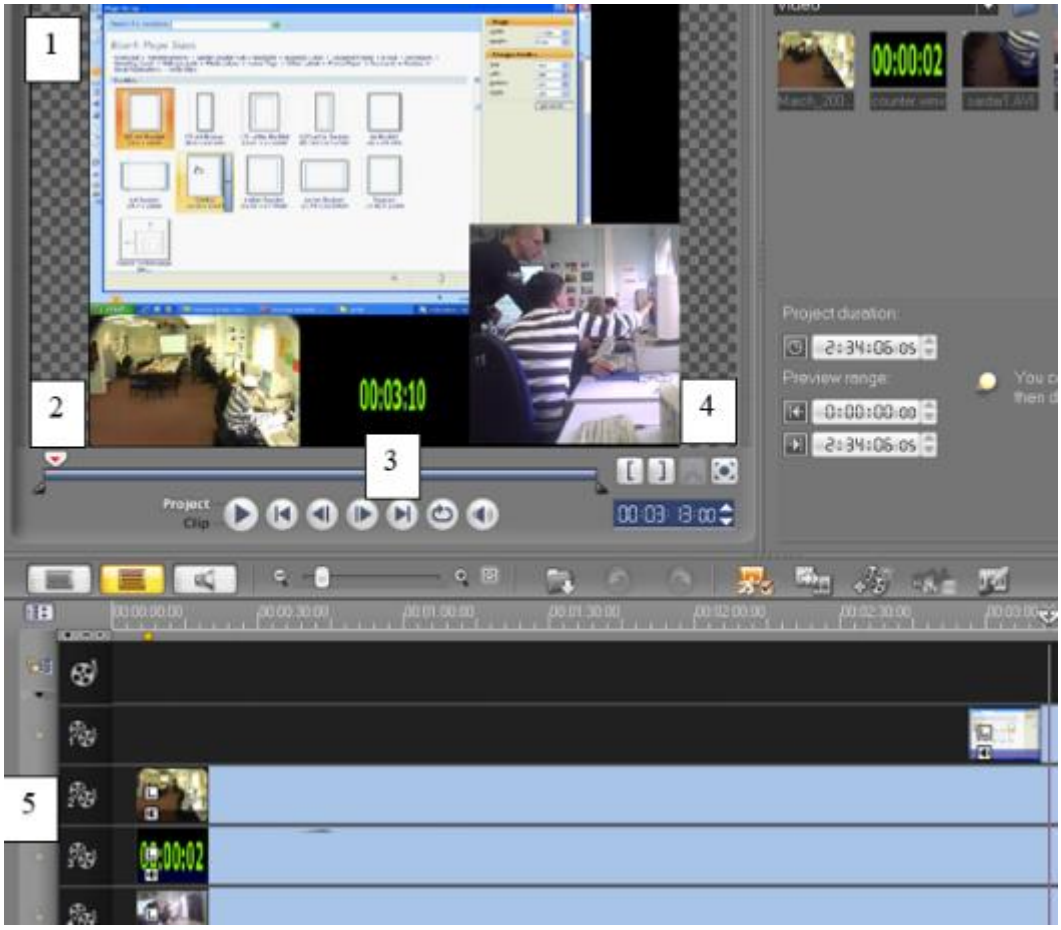


Figure 3-9 (Corel VideoStudio 12)

Item one in Figure 3-9 is the *micro* video stream showing the unfolding actions on-screen. Item two is the *macro* video stream providing a global view of classroom actions. Item three was a timer to show at what point actions started and ended across the totality of the classroom session. References to time sequences in the analysis used this timer rather than the built-in timer of playback media such as Windows Player, VLC, RealPlayer, etc. It was later found that each of these programs resulted in time discrepancies depending on the player being used. For analysis and post-analysis, even

the loss of half a second between what is being referenced to in discussion and what is being displayed on-screen can be problematic. A central time-stamp was thus added on an additional timeline. Item four is the *meso* video stream showing actions around the screen. Item five is the multiple timelines one can add when using video editing software. The audio from the voice-recorder was added to another timeline. The number of timelines available varies with the software being used but in most modern software it is multiple.

Each video on each timeline could then be cut, trimmed and moved backwards and forwards (left and right) on the timeline to synchronise them. Because I was the only person setting up all the devices they each had different start and end points when I moved around the room pressing record on each device. This meant I then had to trim the start points of the different videos and audio from each of the devices. In film production a clapperboard is used to support the later synchronising of audio and images in editing. The sharp sound of the 'clap' provides an audio cue-point across multiple recording devices; this makes it easier for an editor to then synchronise sound and movement when multiple streams are combined. In retrospect, I should have used a clapperboard or similar.

With a sharp sound it would have been easy to trim all the video streams on the software editing timelines to that start point. As there was no sharp identifiable sound on each of the video and audio streams I had to laboriously listen and watch the opening sequences of every device. Given there were four pairs of learners and each pair had three video streams (macro, meso, micro) and two audio streams (a voice recorder each) this required months of data-editing synchronisation just to manage the data ready for analysis. An alternative sound common across the recording devices was eventually identified and it was myself saying 'press F9' so I was able to cut the audio and video data on each timeline at that point which then enabled synchronisation.

Qualitative analytic software such as Transana can handle multiple audio-visual files for analysis directly within the software but not this many files. In addition, given the intention was to eventually present the combined video data to a wider audience, this required compiling all the different media into a central viewing space and exporting to a popular file type such as MPEG and AVI which are compatible with popular playback media such as Windows Player, VLC, RealPlayer, etc. For this reason then of

playability and compatibility, the different audio and videos files were synchronised in video editing software first and then imported as a single file into Transana for analysis.

The biggest challenge proved to be audio. Because there were two voice-recorders I merged the audio files from each device onto a single timeline in Audacity. Exporting merged audio tracks from Audacity as a single audio file, and importing into Corel VideoStudio to place on a timeline, reduced the audio quality because of compression and created problems with synchronisation. The effect was like a badly dubbed film, when what people say lags behind their actions and lip movements. Additionally, when an LSA or I interacted with a pair of learners, this meant adding a third audio timeline from the third audio-recorder to ensure no communication was missed. Even a tenth of a second out on multiple audio timelines can create a cacophony of unintelligible talk. Leaving the audio files on their own timelines in Audacity proved to be the best method (1.) for synchronisation and transcribing because it retained the best quality and (2.) allowed for triangulation to improve accuracy. These two points are explained below.

(1) Synchronisation issues: the soundwave peaks and valleys in Audacity provide a visual cue to synchronise multiple streams. At Figure 3-10 there are two audio tracks from two learners. Each track is displayed on two timeline channels because the file format is stereo. If the file was mono there would be only one timeline channel per audio stream.

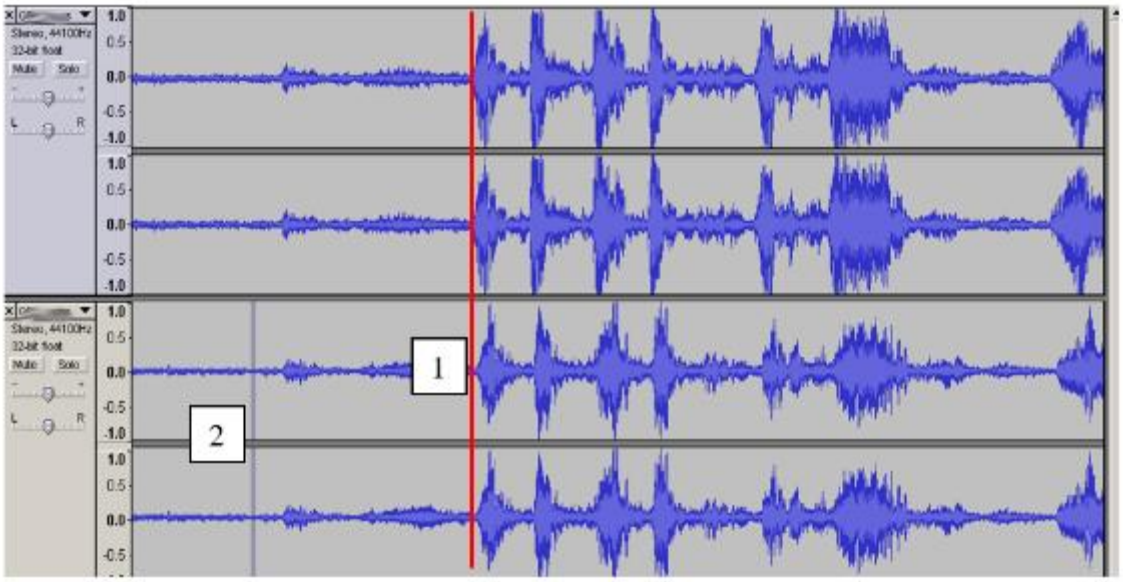


Figure 3-10 (synchronisation problems in Audacity)

A red line is presented at item one in Figure 3-10 to show the gap between the audio track on the top two channels and the audio track on the bottom two channels. It can be seen that the learner's first word (represented by the first large waveform on the top two channels) is slightly ahead of the second learner's audio track on the bottom two channels. This is the same word spoken by the first learner but picked up on the second learner's audio device; hence the waveforms are slightly smaller on the second device (bottom two channels) because that device was further away from the first learner. The gap is only two-tenths of a second but the effect is that of an echo, making transcribing very difficult. However, using the waveforms as visual representations of words, it then becomes easier to match patterns of waveforms with similar peaks and valleys and thus aid synchronisation of multiple audio devices. Item two at Figure 3-10 is a highlighted section within Audacity about to be cut from the timeline channel. This thereby moves the remaining audio, to the right of the section to be cut, down two-tenths of a second to the left and thus synchronising the waveforms across all four channels so that in playback all waveforms on all channels play at the same time.

However, it was later found that some of the learners had pressed pause on their recording devices (for example when they went for a break) and then pressed re-record when they came back into the classroom. Consequently, when I thought I had successfully synchronised multiple audio files by matching the soundwaves of a first word (for example myself saying 'press F9') the audio playback later became jumbled because one learner had pressed pause for ten minutes. This meant using Audacity again to visually find matching soundwaves and either adding ten minutes of silence to the audio channel of that learner or removing ten minutes from the second learner's audio channel; and then resynchronising with the visuals of the video in Corel VideoStudio. Ten audio recording devices in a classroom session of three hours provided a minefield of audio synchronisation issues; complicated even further by so many camcorders.

For example, one pair of learners went to the college toilet halfway through the lesson, realised they still had the audio recorder around their neck, so pressed stop on their devices. This was not discovered until the editing stage when in playback the learners could be heard entering the toilet and realising what they had done. Only one learner pressed the record button again when re-joining the classroom. This significantly hampered the quality of their recorded conversation in the second half of the lesson as only one device was recording.

In this example I had synchronised three audio tracks onto three mono channels in Audacity; Figure 3-11. These were the LSA, Darras and Shourok. One hour into transcribing, the audio became very confusing. Looking at the soundwaves in Audacity revealed that the third audio channel at the bottom (the LSA) looked significantly different to the two audio channels of the students with matching soundwaves at the top.

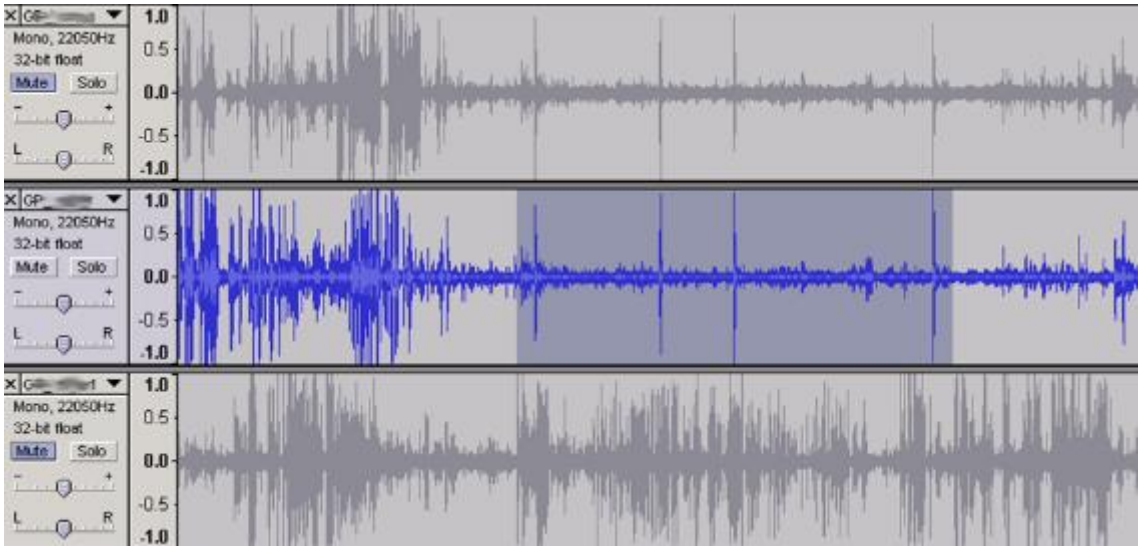


Figure 3-11 (three audio tracks synchronised)

At first I thought perhaps these were door slams hidden under other louder sounds the LSA was exposed to. I could no longer hear the students speaking, just chaotic noise from the LSA's audio channel. After much further listening, this proved to be a wild goose chase as the door slams were not hidden under other sounds. The Kurdish learners had left the room for a coffee break and placed their recorders on the desk, hence they picked up the door sounds of people leaving and entering the room. The LSA left for a coffee break ten seconds after the students but kept her recorder around her neck, hence didn't pick up the door sounds but did record some very confusing sounds from the canteen. The ability to mute channels in Audacity (top and bottom channels greyed out at Figure 3-11) and look at visual soundwaves, enabled me to resolve this audio dilemma and other similar instances.

As discussed, there were significant issues with data editing to synchronise all the devices; and several months of work. These were overcome using software tools to cut and trim audio-visual data but it was painfully laborious. Central to avoiding these issues of synchronisation is (A) use a clapperboard for a common anchor start point; (B) insist that learners do not tamper with any of the recording devices, and (C) establish

common rules, for example, if people need to leave the room then the audio device stays in the room.

(2) Audio triangulation: a positive outcome of synchronising multiple devices in Audacity is the ability to triangulate sounds to improve accuracy of transcription. Figure 3-11 for example shows three channels from the LSA, Darras and Shourok. The bottom and top two channels are muted. The ability to mute multiple channels can be very useful for screening out confusing sounds and focus on particular words. As the conversation is included on three devices, it is possible to cross-check from each device. If it is unclear what one person says on one timeline channel in Audacity, that channel can be muted and listened to on another channel. Table 3-5 provides a brief example of how interpretation can change and how easy it is to miss large sections of conversation from people who are distant from a voice-recorder.

	Before	After
1.	LSA has sat behind them. She says, 'You missed one,' referring to an image missing from the text.	LSA has sat behind them. She says, 'You missed one,' referring to an image missing from the text.
2.	LSA still sits behind Shourok and Darras. She says, 'Remember to do something. You finding mistake only?' LSA seems to be referring to that fact that Darras is doing nothing, Shourok is doing all the work and Darras is just looking for mistakes.	LSA still sits behind Shourok and Darras. She says, 'Try to remember to say something. Don't talk too quiet.' LSA seems to be encouraging them to talk; and perhaps referencing the audio recorders by saying 'don't talk too quiet.'
3.	'Mmh?' says Shourok, not turning around.	'Mmh?' says Shourok, not turning around.
4.	Darras turns to his left and addresses LSA. He says, 'He doesn't understand something.' Shourok laughs. LSA laughs.	Darras turns to his left and addresses LSA. He says, 'He doesn't want to say something.' Shourok laughs. LSA laughs. 'He's keeping it all to himself.'
5.	Shourok turns his head briefly to Darras and says, 'Cheers.'	Shourok turns his head briefly to Darras and says, 'Cheers.'

Table 3-5 (transcript of before and after triangulation of audio channels)

On row two it was difficult to understand what the LSA said. From the context of Darras not interacting, and him pointing out a mistake at row one, I transcribed what she

said incorrectly. With that came an interpretation that she was being critical of Darras's lack of participation. Muting the audio timelines of the two students in Audacity meant I could listen more carefully and I heard a better version of what she said which could then be corrected. This resulted in a different interpretation of the LSA being encouraging rather than critical.

At row four I transcribed Darras's response to the LSA as 'he doesn't understand something' and did not hear him say 'he's keeping it all to himself.' Muting the audio channels of Shourok and the LSA allowed me to focus in on Darras's audio only and get a better transcription. In this scenario it was pivotal because Darras's criticism of Shourok is what explains Shourok's actions later when he pushes himself away from the computer space and refused to interact with Darras. Prior to the accurate transcription it was difficult to understand why Shourok behaved the way he did. There were many similar instances of uncertainty in everyone's conversations but triangulation across audio timelines in Audacity by muting allowed most uncertainties to be resolved. To ensure accuracy of transcription this meant Transana and Audacity were best used in conjunction with each other rather than depending solely on the audio track in Transana; Figure 3-12.

Figure 3-12 (using Transana in conjunction with Audacity)

Here is an example of when gesture and unclear audio coerced me into an incorrect transcription of what was actually said. Triangulation was achieved across the programs.

Before	After
‘Under,’ says Gamda and her hands do a sweeping upwards movement. ‘Under,’ her arms then raise above her head and move frantically about, like a child doing an impression of tree branches in the wind, 'wings.'	‘On the,’ says Gamda and her hands do a sweeping upwards movement. ‘On the,’ her arms then raise above her head and move frantically about, like a child doing an impression of tree branches in the wind, ‘rubbish bins.’

Table 3-6 (before and after of transcribing from gesture)

Because the audio was unclear in Transana I used gesture as a supplementary cue to make meaning. Her arms high and waving in the air were suggestive of something being high and flapping; e.g. trees, wings. The gesture was understood to be iconic, representing a tree or a bird, when in fact it was metaphoric; her hands were physically scrambling in the air to cognitively find the English words she was looking for. This is confirmed by the repeated ‘on the’ as she tries to think of the compound noun ‘rubbish bins.’ Using Audacity and muting one audio channel to hear clearer a different audio channel proved successful in most cases.

Other difficulties in transcription arose when the learners switched from English to their first language. Sometimes it was obvious when the learners had switched. At other times it was uncertain if they were speaking English, or their first language, because the words were difficult to hear. Language-switching between second language speakers can be full sentences but often might be a single word, for example, in Kurdish or Polish but interspersed within English conversation. In each instance of uncertainty and clear switching I flagged these on the timeline to return to. As I work within a large ESOL department I was able to sit with a member of teaching staff, first a native Polish speaker and then a native Kurdish speaker, and together we dual transcribed these instances of language-switching. Collectively this contributed to greater overall transcription accuracy of the learners’ conversations.

3.8 conclusion

In this chapter the overall methodology for data-collection was explained. Demographic and biographic information on the learners and their locality was provided. Significant issues around employability, deprivation and language complications were found with most learners experiencing challenges in these areas. A practitioner-researcher approach was introduced with consideration of an *ethnomethodological case study* because of the similar interests and methods across case studies, ethnographies and grounded analysis. This included an explanation of interviewing and observation as the tools and methods used to collect the data in a single lesson and across paired participants. Observation as video was defined as videography with four broad sequential strands: *collection*, *preparation*, *transcription*, *analysis* and the first two were discussed in this chapter. The next chapter will consider the other two, discussing the analytical approach with an explanation on how the multimodal data from the videos was transcribed.

4 Analytical approach

4.1 introduction

The purpose of this chapter is to explain the multimodal transcription of the collected data and the relevance of the analytical approach. A micro-analytical approach is adopted to analyse the interaction of the learners, introduced in this chapter and applied later in chapters 5 and 6. Chapter 3 discussed the data-collection and methodology. This chapter introduces a methodological framework for analysing the learners' interaction (Norris, 2004) alongside a discussion of scales of time (Lemke, 2009) and analysable units. Consideration is given to what is meant by an action, where it starts, ends and bleeds into other actions. Analysing the moment-to-moment interaction of learners to understand the sequential and parallel actions in collaboration is one reason why the totality of the lesson was recorded and transcribed. The structuring of lower and higher-level actions in Norris' methodological framework (2004) is used to enable this.

Section 4.2 explains the learning outcomes for the lesson; what the learners were asked to do and why. It then attempts to define what an action is in the context of this research and from an ethnomethodological approach. Section 4.3 discusses the problem of time scales (Lemke, 2009) and analysable units. A methodological framework for analysing interaction (Norris, 2004) is introduced and higher-level actions explained as an outcome of preliminary video analysis. Section 4.4 discusses lower-level actions and modal density as a term for describing the interplay of several modes. Section 4.5 explores the problem of transcribing multimodal video data and introduces the method created for analysis in chapters 5 and 6. Section 4.6 introduces *modal density tables* as a transcription method for displaying modal configuration in actions. 4.7 explains the secondary transcription method which are *multimodal extracts* for displaying the sequential synchronous and asynchronous appearance of modes. Section 4.8 considers the focal events, how these were identified, and which participants were chosen for detailed analysis. 4.9 concludes this chapter.

4.2 defining actions

This section contextualises what the learners were asked to do and why. Consideration is then given to actions because when analysing peer-interaction, and how learners negotiate and complete their learning outcomes, actions are performed and these are the observable phenomena for analysis. Identifying the organisation of actions within

learner collaboration is what will ultimately lead to a peer-interaction framework in this research. Actions though are complex so this section will problematise the term to clarify how it is being applied.

The learners in the study were given a general aim for the lesson which was to 'create different types of texts providing information about Greenpeace or environmentalism.' The three specific outcomes they were asked to achieve were: (1) create a booklet and (2) a protest image and (3) a website. Brief instructions were printed and given to each learner; (Appendix J). Digital competency skills had been previously taught using the programs Microsoft Publisher, Adobe Dreamweaver and Adobe Fireworks. From a pedagogical perspective, the learners were being asked to apply and transfer those skills independently. The learners were given the freedom to achieve these three measurable outcomes without additional tutor input and to do so through peer-interaction. The tasks were designed to exploit the learners' multimodal, multiliteracy skills, as discussed in section 2.3, encouraging them to work with language, literacy and technology (Anstey & Bull, 2006). From a practitioner-researcher perspective, I was interested in the intersection of these as a multimodal ensemble in peer-interaction; see Figure 2-2. Their embodied interaction at a shared computer and their negotiation of the design and writing process were of particular interest. Their digital skills were not of interest.

For educational purposes it is important to understand the learning constraints and affordances which are generated by the intersection of language, literacy and technology; for example, asking how multimodal designs are *digitally-realised* on-screen through the off-screen negotiation and collaboration of learners. To work towards this requires an approach which identifies action as the core unit of analysis. Norris for example states that 'in multimodal interactional analysis, the *mediated action* is the unit of analysis, and since every action *is* mediated, I will simply speak of the action as the unit of analysis' (2004: 13).

The research collected audiovisual data for the duration it took the pairs of learners to complete the three outcomes; every word, every gesture, resource used, action and interaction, off-screen and on-screen. The justification for this whole lesson approach to analysis was based on uncertainty around actions. How the actions of learners are sequentially organised across modes and temporal timescales was something I wanted to better understand. If actions are then to be rendered into analysable units, it is not

easy to divide actions and inter(actions) into research friendly chunks; actions bleed into each other with sequential impact on other actions and people in the wider environment. *Sequentiality* is a key concept in conversation analysis for identifying the sequential organisation of utterances between speakers to maintain coherence and sustain the interaction. Multimodal interactional analysis has the same concern but broadens the communication cues. Prior to starting the research, and identifying specific research-related questions, I had several questions in my head from casual observations of learners interacting. I began to think about:

- What is an action and how is it different or similar to an interaction?
- When does an action begin and end across different scales of time as learners work in real-time to achieve learning outcomes?
- How does an action cause reactions and subsequently series of interactions?
- What modes are significant, and less so, for learners in the coordination of action?
- What additional resources do learners draw on to help them with their actions?
- And finally, when we think we have *defined* and understood 'action' in an educational context, how do we *identify*, *capture* and *analyse* those inter(actions) in real-time?

In terms of *defining* actions, there are phenomenological, behavioural and philosophical interpretations in these fields, particularly as developments within early psychology when the field began to grapple with the concept of action; (e.g. writings of Edmund Husserl, B.F. Skinner and John B. Watson). It has been a multidisciplinary concern: 'Different disciplines have taken very different kinds of phenomena, ranging from the mental intentions of individual actors to large, historically shaped social structures' (Streeck, Goodwin & LeBaron, 2011: 1).

Action as I am applying it in an educational context is understood to have purpose, intentionality; there is agency and cognitive engagement with goal-orientated outcomes which are realised through physical means. In this sense actions are observable phenomena, the embodied manifestations of intention. Heath, et al argue for a 'methodological framework … that prioritises the situated and interactional accomplishments of practical action' (2010: 1). Using video as an observational method might be considered ethnomethodological because the video analysis concerns itself with empirical knowledge of communicative interaction, based on the observations of actions performed by individuals, to make sense of social order.

Tariq for example is sat in class wondering how to spell a word. He is thinking about asking the teacher for help, or another student, or standing up to collect a dictionary from the cupboard. He may be cognitively engaged with learning but until he physically *engages* with the environment it is not possible to see actions. There may be inaction, which in itself might be meaningful. However, from a researcher perspective in multimodal interactional analysis, observable actions are central and not what students might be thinking or feeling. For the purpose of this research, the primary interest is with what people say and do and their interactional awareness of other students in terms of how they react. Tariq might collect a dictionary to look up a word, so it can be surmised from his actions what he is doing and he reveals preference through his choice of action, but he is not interacting; other than with a book. This would suggest that interaction is a social process between people, comprised of actions and reactions. There are individual actions and social interactions. The focus of this study is educational events 'in which multiple parties are carrying out endogenous courses of action in concert with each other within face-to-face human interaction' (Streeck, Goodwin & LeBaron, 2011: 1). If Tariq stands up and walks across the room to ask another student how to spell the word, then from an analytical viewpoint, I am interested in what he does and how he does it, and equally, how the other student responds. Multiple chains of actions and reactions, mediated by verbal and nonverbal means, contribute to the sequentiality of interaction.

It is commonplace in many educational settings for the teaching and learning to be designed around achievable outcomes which are time-bound by a single lesson; for example: the teacher starts the class with "by the end of the class today you should be able to …" and at the end of the lesson reviews with the learners if they have achieved the objectives for that session. The journey taken by the learners through that process is one of multiple actions sequenced in coordination with the actions of others. A single lesson, bracketed by a clear start and end point, seemed to be an appropriate scale of time to start addressing my questions about actions. I decided to video a full lesson of four pairs interacting at a shared computer. At the time I did not realise the complexity of the questions I was asking myself when thinking about actions and peer-interaction. The data was so rich I could only micro-analyse minutes rather than hours. This methodological approach of video analysis is similar to that of Bhatt & de Roock who explain that 'our research methodology is … guided by ethnomethodological (e.g. Garfinkel 1967) approaches that prioritise seeking an emic/insider perspective … In

other words, the phenomena being examined, and their constitutive practices, guide theorising and data gathering. We focus on empirically observable events and interactions as they happen' (2014: 2-3).

Grounded knowledge of actions and peer-interactions with technology was my main priority so tasks were given to the learners, and videoed, for later analysis. Upon later viewing it was evident that a full lesson contains hundreds of interactions comprised of thousands of actions including utterances and gestures. What I needed next was a methodology to *identify* and name unfolding actions, the catalysts and completions, the failed actions, and the embodied cross-modal coordination of these. A method for identifying the overall interaction of learners towards the completion of their learning (the sequential actions, reactions and even inactions) is the subject of the next section.

4.3 scales of time

This section will introduce Sigrid Norris' methodological framework (2004) for analysing interaction. Actions are identified as operating on different scales of time. This section will focus particularly on higher-level actions and the next section will focus on lower-level actions. The three learning outcomes were (1.) design a booklet about the environment with writing; (2.) design a protest image, and (3.) design a website. Four pairs of learners (Kurdish, Polish, Arabic and French) were videoed for the entire lesson. Completing the three outcomes required the successful sequencing of multiple *higher-level processes.* Lemke explains the emergent processes and patterning of actions on shorter and longer timescales.

> … there is always … a higher level process already in place, already running on its own longer timescale, and this sets the context that constrains what is likely and what is socially appropriate at the next scale below. A student's answer to a teacher's question is also meaningful for the participants as part of an exchange, not just as an utterance in its own right, and is judged as appropriate or not to the ongoing exchange and to the episode, the lesson, the unit, the curriculum … and many higher-level contexts. These contexts, however, are not static; they are themselves processes un-folding in time. Very slow processes function like constant, static backgrounds on the timescale of much faster processes (Lemke, 2009: 276).

The curriculum year, with a Scheme of Work mapping out the learning journey over several months, might be considered the longest scale of time which runs slowly in the background but also constrains what happens at the next scale of time below it: the term

and then the unit, the week, the lesson, the learning objectives, the actions, etc. In this research a single lesson was the top-level scale of time, from when the learners started their tasks to when they completed them. This timescale defines the temporal parameters of the study at the highest level, however, as Lemke suggests, there are always scales below that. If the research is interested in the totality of the learning process, and all the sequential steps the students perform and negotiate to complete their learning in a single lesson, then a method is needed for capturing the whole and deconstructing the totality to understand the underlying parts of the whole. Sigrid Norris' methodological framework is useful for analysing the multimodal actions of individuals and the interaction of multiple learners. It details the communicative modes prevalent in interaction and breaks these down to materiality, density, complexity and levels of participant awareness. Each gesture, each utterance, could be considered a *lower-level action* (LLA) which contributes towards realising the *higher-level action* (HLA). Norris explains this methodology of actions which can be understood as higher and lower:

> We can think of lower-level actions as the actions that are fluidly performed by an individual in interaction. Each lower-level action is mediated by a system of representation (which includes body parts such as the lips, etc. for spoken language; or hands, arms, and fingers for manual gestures). Higher-level actions develop from a sum of fluidly performed chains of lower-level actions, so that the higher-level actions are also fluid and develop in real-time. Every higher-level action is bracketed by social openings and closings that are at least in part ritualized' (2004: 14).

Lower-level actions are the embodied communicative modes discussed in section 2.6 under the headings of proxemics, vocalics, kinesics, chronemics and linguistics. For Norris, separating the communicative modes is an analytical procedure to understand how they work together in interaction. This de-contexturing (or unweaving) of interaction to its constituent actions is also recognised by Streeck, Goodwin & LeBaron: 'when joined together in local contextures of action, diverse semiotic resources mutually elaborate each other to create a whole that is both greater than, and different from, any of its constituent parts' (2011: 2). Identifying actions could be considered a twofold process: (1) *higher-level actions* which are goal-orientated intentions and (2) *lower-level actions* (semiotic resources such as gesture and talk) which structure the higher-level action through their materiality: utterances, touches, movement, etc.

Observing the videos allowed for multiple viewings, pausing, and adjusting the speed of play. The video was played back in Transana which allowed for timestamping of when actions began and ended. A narrative transcript was written as a descriptive commentary on the video, prior to a detailed transcription once the focal points had been identified. For example, the video of the Polish learners lasted nearly 3.5 hours. The narrative transcript for the video of the Polish learners took 23,000 words to write and approximately two months to complete. By way of example, Table 4-1 provides a short sample of the narrative transcript written using Transana, with the software's time-stamps on the left, so that for further analysis, one can quickly jump backwards and forwards through the video and to researcher commentary or ethograms.

1.1 higher-level action begins: create a booklet template		
1.	0:01:23.3	"Okay. Open Publisher," says Sakia. Gamda holds the paper instructions with her left hand, uses the mouse with her right-hand and stares at the screen.
2.	0:01:31.4	Gamda opens Publisher.
3.	0:01:33.1	Gamda looks to her right, smiles and whispers to Sakia. They appear to be laughing at something the other students are doing. Gamda seems to say: "We should beat the mouses," referring to the male students to her right.
4.	0:01:40.3	"Okay, booklet," says Sakia as Gamda uses her mouse to roll over the 'publication types' in Publisher.
5.	0:01:42.1	"Where is booklet?" says Gamda, and waves the worksheets in her left hand.
6.	0:01:43.2	Sakia waves her hand vaguely towards the screen.
7.	0:01:46.3	Sakia says, "Go to file."
8.	0:01:57.1	Sakia picks up her worksheets and reads briefly. Gamda uses the mouse to find the booklet type.
9.	0:02:03.7	Sakia quickly leans forward to tell Gamda the correct booklet type. She says, "14.5 cm."
10.	0:02:05.5	She places her finger on the screen to indicate the booklet type. "That's the booklet. Okay."
11.	0:02:17.5	"Erm ..." says Gamda.
12.	0:02:27.9	A pop-up box appears asking them to confirm 'more pages' and Sakia points at the screen at the 'yes' button.
1.1 higher-level action ends: create a booklet template		

Table 4-1 (narrative transcript of the Polish learners)

It was discussed in section 4.2 that actions are here defined by 'intentionality; there is agency and cognitive engagement with goal-orientated outcomes which are realised through physical means.' In this sense actions are observable phenomena, the embodied

manifestations of intention. Watching the video from the perspective of higher-level actions, by asking what the learners are trying to achieve and when it is achieved, enabled the segmenting of actions with clear start and end points. Table 4-1 is presented as an example of how one higher-level action was identified and narrated. The narrative is an attempt to give a factual description of actions, avoiding interpretation and leaving theorising till later. In this instance the learners want to create a 'booklet template' and the action begins when the learners open the Publisher program at 1, 0:01:23.3 and ends when Sakia points at the screen to show Gamda where to click at 12, 0:02:27.9; the template has been created so that particular action is considered to be complete. The lower-level actions are the modes the learners use to complete the higher-level action. At this primary stage of video narrating it was enough to simply describe what the learners were doing. There is evidence of spoken language as they communicate to complete the action, using questions to clarify, imperative verbs to instruct, whispering, laughter. There is evidence of them using paper instructions to help each other. There is gesture as they point at the screen and spatial positioning as one learner moves into the screen and away from it. It is therefore possible to say that the learners achieve the higher-level action through the modal configuration of lower-level actions performed across linguistic, proxemic, kinesic and vocalic modes. One can even talk about their chronemic behaviour in terms of how quickly or slowly they respond to each other. The analysis of lower-level actions, at the level of second-by-second peer-interaction, is used in the analysis chapters at chapters 5 and 6.

Writing a narrative transcript like this for a video of a full lesson was a significant undertaking but also revelatory. The inverted pyramid at Figure 4-1 visualises the top level scale of time within this study and the constituent timescales of interest under that.

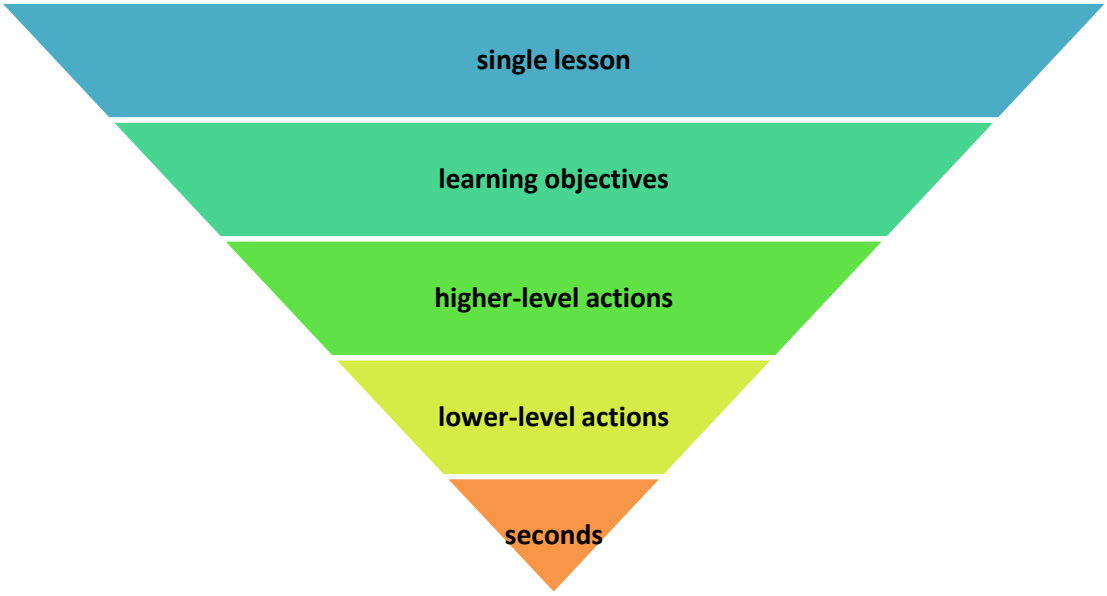


Figure 4-1 (inverted pyramid showing actions across different scales of time)

Watching the video of the entire lesson revealed the start and end points of each completed objective: (1.) booklet (2.) image and (3.) website and the start and end points of each higher-level action within those objectives. The primary stage of narrative transcription in the video analysis was important to identify all the higher-level actions (e.g. Appendix B for the Polish learners and Appendix C for the Kurdish learners). The lower-level actions were described in the narrative transcripts (e.g. Table 4-1) and later micro-analysed second-by-second when the focal points of interest had been identified.

Observing and narrating on the totality of interaction in a lesson, with time-stamps, revealed the duration of each interaction and that some actions performed over longer scale of time have shorter actions nested within them; see Appendix B. For example, higher-level action 1 starts at 0:00:00.0 but does not end until 1:48:55.7. Within that action there are multiple sequences of other higher-level actions and within those HLAs multiple sequences of lower-level actions. To be clear, the performance of higher-level actions is not alluding to multitasking. Rather, most actions run sequentially but some actions were found to run parallel to each other. This would suggest that some actions are foregrounded and backgrounded in terms of learner prioritising.

At Appendix B, the start of making the booklet at higher-level action 1.0 is temporarily suspended at 1.5, 0:24:47.7 as the learners start the second learning outcome 2.0, 0:24:52.2, and then they return to the first higher-level action at 1.6, 1:32:10.1. The

higher-level actions from 1.1 to 1.5 run sequentially so that the completion of each goal-orientated action enables scaffolding towards the next action. Some actions run parallel to each other on different scales of time. Completion of the booklet for example ends at 1:48:55.7 but is on a longer scale of time so is backgrounded at 0:24:52.2 when the learners start to design an image, until the booklet is continued at 1:32:10.1. Other actions are nestled within those higher actions.

Thinking of cognitive-load, there may be differing levels of awareness and attention during interaction. Norris explains this: 'attention/awareness levels have to be considered in human interaction for two reasons: 1. without differing attention/ awareness levels, simultaneous interaction would not be possible; and 2. expressing and perceiving (as displayed through expressing) are both performed at least partially with the mind of the social actor'(2004: 150). This would suggest that individuals are capable of having fluctuating levels of awareness and attention. In *awareness* a learner can be conscious of something without reacting to it whereas *attention* would suggest a level of *focussed* concentration. A learner can engage in simultaneous higher-level actions with differing levels of attention/awareness. This would suggest (at a cognitive level) that individuals have to foreground their attention to some actions and modes whilst others are midgrounded and even backgrounded; the learners might still be aware but it is not primary in their attention. From an analytical approach this is important to consider: observing actions and ascribing intentions needs to be handled carefully. Learners may be inactive but this does not mean they are unaware or inattentive to other actions and learning in the classroom environment.

Norris (2004: chapter 5) uses what she calls a modal density foreground-background continuum as a methodological tool to display how attention/awareness levels fluctuate across simultaneous higher-level actions, dependent on the modal density and individual priority at any given point in an interaction. It is a visual depiction only to show how an individual prioritises actions with decreasing levels of awareness.

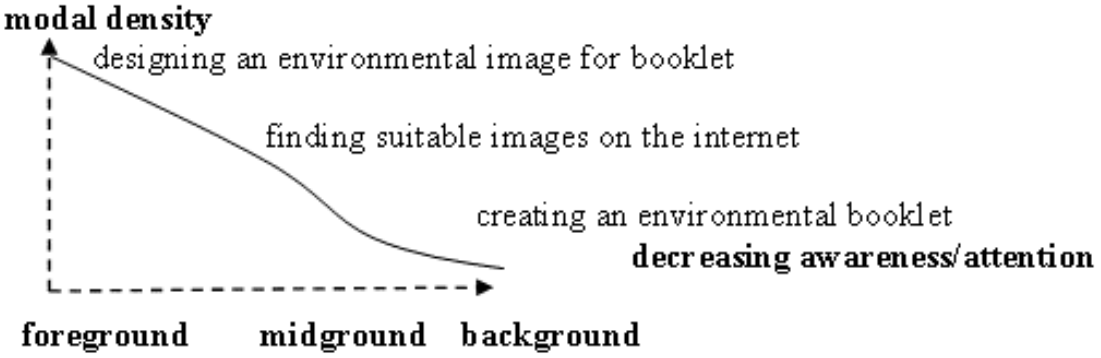


Figure 4-2 (foreground-background continuum)

Modal density is discussed in the next section. In the example actions discussed above from Table 4-1, it can be seen in Figure 4-2 that what was a foregrounded action of creating a booklet with sophisticated modal density of language, gesture, posture, etc, was slowly backgrounded as other actions were foregrounded. The three higher-level actions run parallel to each other with different start and end points. These are noticeably different to actions which run sequentially. The x-axis reveals the decreasing level of attention against the y-axis of modal density. 'The curve in the graph indicates that a decrease of modal density equates to a decrease in expressed attention/awareness of the individual. This curve is only an approximation' (Norris, 2004: 98-99).

The preliminary analysis of the video resulted in the identification of actions across different scales of time (e.g. Appendices B and C). Most actions were sequential as they built on each other whilst other actions ran parallel with decreasing and increasing levels of attention. Actions were identified as lower-level and higher-level and Norris' methodological framework was helpful for identifying actions. This section introduced higher-level actions across different timescales. The next section will look in more detail at lower-level actions.

4.4 modes and modal configurations

Modes are itemised to help with analysing them independently. This can help to identify how they are interdependent. The importance of this is voiced by Flewitt et al, calling for 'levels of description and interpretation of data … describing the ways in which different modes interact together … but also how individual modes are constituted' (2009: 42). This section will explore modes further to prepare the ground for the 'description and interpretation' element of transcribing which is in the next section.

Deconstruction of interaction at modal level can help to hierarchise modes and understand their interconnection for individuals and for the interaction: 'Only when focusing on the real-time interactions without preconceived notions of hierarchical structures of modes, will we be able to discern the true value of each mode in a specific interaction' (Norris, 2004: 53). Placing *value* on a mode is problematic, as modes shift in relation to unfolding actions, so it needs to be clear what is meant by this: not that one mode is always more important to an action than another but that in multimodal ensembles there are modal affordances across different modes. Norris uses the term *modal configuration* to refer to the 'hierarchical, equal, or connected relationships among the modes that are at play in a given higher-level action … Modal configuration is a concept that investigates how modes in interaction are structured in relation to one another' (2009: 78).

Modes are understood as systems of representation and communication. In the Oxford American Dictionary, the term *modal* is described as 'a way or manner in which something occurs or is experienced, expressed, or done.' Norris offers a slightly more helpful definition: 'modes are not bounded units. A mode is a loose concept of a grouping of signs that have acquired meaning in our historical development. We need to keep an open mind about communicative modes and think of them as loosely bounded units rather than distinct entities' (2004: 152). To talk about the mediation of modes, and the mediated action, requires an understanding of the rules and regularities which govern how a mode is used, shaped by sociocultural, material and historical forces. Kress for example dedicates a chapter to try and answer 'What is mode?' He begins: 'Mode is a socially shaped and culturally given resource for making meaning. *Image, writing, layout, music, gesture* … are examples of modes used in representation and communication' (2009: 54). Actions are mediated by communicative modes in the sense that actions are performed and performance is *governed* by the systems of representation they draw on with different affordances and constraints; (e.g. the materiality and temporality of modes). Modal density and configuration is one way of looking at this, explained below, as this is a core analytical approach in the later chapters.

Modal density is a term for naming the collective occurrence of modes in interaction, as coined by Sigrid Norris (2004). It is through the density of inter-semiotic relations that actions are structured and communication enabled. The categorisation of

communicative modes is potentially huge as explained by Poyatos: 'taking place through "*kinesic, kinetic, vocal, chemical, thermal, and dermal*" channels for "emission," and "*visual, auditory, olfactory, dermal, and kinaesthetic*" channels for "perception," (1980: 114, cited in Stam & McCafferty, 2008: 3). Noticeably these are embodied, sensory modes for perceiving and emitting information. Other static disembodied modes such as space and layout can also be significant as people position themselves, or are positioned, by the layout. There are culturally learned ways of behaving with space and layout; for example, sitting next to someone you have never met before in an almost empty train would be considered inappropriate, and even threatening, but tolerated if the train was becoming full. Sitting next to someone you have never before met in an almost empty classroom might be perceived as more socially appropriate as an extension of friendship.

Poyatos, with Stam & McCafferty, reference embodied communicative modes. Norris however makes a distinction: 'all modes are both embodied and disembodied – depending on whether an action is performed or read and interpreted – and interaction is a constant shifting back and forward … the boundary between embodied and disembodied is fuzzy … we do not need to worry about the actual boundary too much' (2014: 46). Norris provides the mode of print and writing a shopping list as an example. Writing a shopping list is an embodied higher-level action structured by the lower-level actions of writing individual letters, with pen and paper, and through the individual's physical and cognitive engagement with the tools and the activity. For the purpose of an aide memoire, the modal affordance of print has higher value than speaking because it has lasting temporal materiality, compared to speaking with limited materiality. Once the list is finished it becomes a disembodied mode for later use. For Norris, the fluid actions of writing are *frozen* in the list. Another person using the list at a later time would use the list in a disembodied way. The modal density of the action is configured by several modes (pen, paper, letters. layout) but it is the letters, or print, which has most value as it contributes most to meaning.

In this research, the modal density of the learners' interaction at a computer is configured predominantly by the following array of communicative categories:

- print (e.g. worksheets, website, writing, translation device)
- languages (e.g. English, Polish, Kurdish, Azerbaijani)
- kinesics (e.g. gaze, posture, gesture, artefacts, object handling)

- proxemics (e.g. layout, personal space, objects)
- vocalics (e.g. prosody, pausing, loudness, intonation)
- chronemics (e.g. time-lags, temporality, speed of speaking, responsiveness)
- visual (e.g. colour, shapes, image choices)

For some gaze would fall under oculesics but here is referenced under kinesics in the sense that it is an embodied mode. Other categories of nonverbal communication, such as haptics, olfactics, gustorics, were not considered relevant in the current study. The pre-analysis of writing a narrative transcript revealed these seven modal categories as most prevalent. Peer-interaction is a fluid and complex arrangement of the above categorised modes. Understanding the shifting modal density of interaction is an important part of identifying how talk and action are configured and peer-interaction frameworks structured. Each mode has potential value in interaction, at a communicative and practical level, to complete an action. To understand the modal density of a passage of interaction it is helpful to deconstruct actions at a modal level but only if the intention is to demonstrate the inter-semiotic relationship of the whole as a modal configuration. As van Leeuwen explains: 'describing semiotic resources by themselves, or analysing the work of the modes separately, does not begin to show what happens when they are put together' (2011: 675). The next section will demonstrate how such modal configurations might be delineated as a form of transcription for the purpose of analysis and representation.

4.5 video analysis and transcription

In this study there are three types of multimodal transcription used: (1.) modal density tables show the configuration of modes to suggest dominant and subordinate modes; (2.) multimodal extracts show the unfolding synchronisation of modes and actions on their own timelines; (3.) sequence of stills with spoken language and arrows overlaid. This section will introduce and explain how these multimodal transcripts are used in Chapters 5, 6 and 7. The research literature on multimodal transcription with video analysis reveals a common voice of uncertainty across two particular concerns: transcribing multimodal data and analysing it. 'When it comes to the transcription of visual phenomena we are at the very beginning of such a process' (Goodwin, 2000b: 10). Fifteen years later, whilst there have been developments, not a great deal has changed: 'Unlike transcribing talk/language, for which authors often acknowledge the foundational transcription conventions of Gail Jefferson … no commonly shared format for representing and transcribing embodiment has yet emerged' (Nevile, 2015: 133).

There is greater confidence in the analysis of embodied interaction but less so in the transcribing. Flewitt et al provide a chapter giving an overview of multimodal transcription. What that chapter reveals is varied practice as researchers bespoke systems of representation for multimodal text-based content and videoed social interaction: 'matrices with columns have been used by an increasing number of researchers to add a spatial dimension' (2009: 47). The challenge is representing multiple modes mediated by space and time in a print-based format.

In section 4.3 it was discussed how identifying actions could be considered a twofold analytical process: (1.) higher-level actions which are goal-orientated intentions and (2.) lower-level actions which structure the higher-level action through their materiality: utterances, touches, movement, etc. Observing the videos allowed for multiple viewings, pausing, and adjusting the speed of playback. A number of passes were used to identify the higher-level actions. The video was played back in Transana which allowed for timestamping of when actions started and ended. A narrative transcript was written in Transana as a descriptive commentary on the video, prior to a detailed transcription once the focal points had been identified (e.g. see Table 4-1) and a structure of actions was identified (e.g. see Appendices B and C). The secondary stage of analysis required detailed transcription of the lower-level actions. Three systems of representation were bespoked and these are explained next.

4.5.1 modal density tables

Tables of modal density like below are used in this research to visualise the choreography of modes when discussing learner interaction. The larger the square the greater value, or contribution, the mode has in the interaction. 'Modal density refers to the modal intensity and/or the modal complexity through which a higher-level action is structured' (Norris, 2004: 79). Norris uses circles to illustrate this but squares in tables are preferred in this research.

Sakia

language		gaze	
		print	
gesture	objects	layout	proxemics

Gamda

print		gaze	
		smiling	
gesture	objects	layout	proxemics

Table 4-2 (modal density 1)

Modal complexity refers to the configuration of several modes at once where none has greater intensity, or value, than another. Each mode has relatively equal importance in the interaction. In the table above we can see that for Gamda, gaze, print and smiling have slightly more value to the interaction as she looks at Sakia and smiles in agreement as she reads her worksheets. For her, gesture, objects, layout and proxemics have less contributory value to the interaction, as though they are present, these modes are not significant for her at this given moment. No one mode is more dominant than another so for Gamda the modal density of her action is said to be complex.

Modal intensity refers to the increased weighting of a mode in an interaction, where there is a more prevalent mode, because it structures the other modes and is primary to the completion of an action. In the table above we can see that 'language' has the greatest significance for Sakia as she is talking and typing on the screen; therefore, though there is modal complexity in her interaction because there is a prevalence of modes, we can also say there is modal intensity in Sakia's actions as she is dependent more on language than the other modes. Their participatory framework then is structured by differing levels of modal density but speech for Sakia has the greatest communicative intensity as this is structuring their interaction. I will provide other examples from Table 4-1 to explain further.

timing	actions	commentary
0:01:42.1	"Where is booklet?" says Gamda, and waves the worksheets in her left hand.	Language has the highest value.
0:02:05.5	She places her finger on the screen to indicate the booklet type. "That's the booklet. Okay."	Gesture and language have equal value.
0:02:27.9	A pop-up box appears asking them to confirm 'more pages' and Sakia points at the screen at the 'yes' button.	Gesture has highest value.

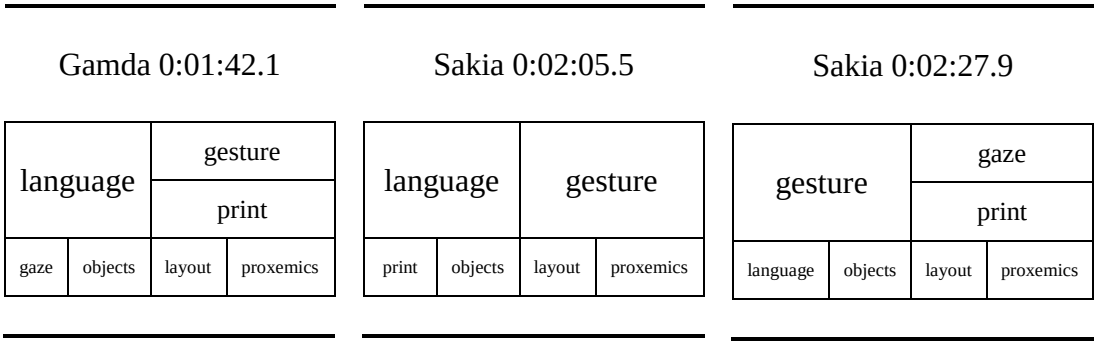


Table 4-3 (modal density 2)

At 0:01:42.1, Table 4-3, Gamda asks ‘Where is booklet?’ and waves the worksheets in her left hand. These were provided previously by the tutor and she understands from the instructions that she needs to create a booklet. The disembodied mode of print is evident and is used as a form of gesture, waving the worksheets as a metaphor for her cognitive uncertainty, which is also articulated by the question. Gaze is evident as she is looking at the screen to understand where to click, as well as other lesser modes, and these are presented in smaller boxes. Language has superior value to the action, denoted by a larger box, because this is the mode with most communicative meaning. What Gamda wants to communicate can be understood by speech alone, without the gesture; so the gesture is subordinate as it cannot be understood without the utterance. The gesture with the worksheets still adds value to the interaction, more so than the relatively static modes of gaze, layout, etc, so these are placed in comparatively larger boxes. In this instance, the easiest representational mode for Gamda to communicate is language. The modal configuration in this interaction has *modal intensity* because one mode is more prevalent than the others.

At 0:02:05.5, Table 4-3, Sakia places her finger on the screen to indicate the booklet type and says. ‘That's the booklet. Okay.’ In this instance language and gesture have equal value and are placed into equally sized boxes at Table 4-3. If Sakia was only to use language (‘that's the booklet okay’) it would not be possible for Gamda to understand what she meant. ‘That’ is a demonstrative pronoun which indexes the students in relation to the computer and the interaction. It is not possible for Gamda to understand Sakia’s utterance without understanding what the referent of ‘that’ is so Sakia uses a gesture to point at the screen to signifiy the referent of ‘that’ (e.g. see Scollon & Scollon, 2003: 31, for the indexicality of language). *Environmentally coupled gestures* like these broaden our understanding of the communicative landscape as they can evidence an indiviudal’s cognitive engagement with the environment through multiple modes (Goodwin, 2007a: 55). In this instance we can say that the modal density in their interaction is complex but the modes of language and gesture have the greatest intensity because of their superior affordance to the action.

The final example reveals how gesture can have the most superior value to an action. At 0:02:27.9, Table 4-3, Sakia points at the screen at the 'yes' button to indicate where

Gamda needs to click so they can complete the higher-level action of creating a booklet template. The students are not speaking. The deictic of the gesture has enough communicative meaning. It is the easiest representational mode for Sakia to use. The modal density in this interaction has *modal intensity* because the mode of gesture is more prevalent than any other; deictic gesture structures their interaction at this given moment of time.

The visual hierarchy of modes into box sizes should not be considered as definitive; rather, it is an analytical means for presenting a range of fluctuating modes. This is the same for displaying attention/awareness levels on a foreground-background continuum (i.e. Figure 4-2). Interaction is comprised of actions and reactions and these are orchestrated by modes with shifting degrees of value to the interaction. They are not fixed states but fluid arrangements; as Norris states: they allow 'us to analyse large aspects of what has traditionally been termed *context* in discourse analysis' (Norris, 2004: 149). Modal deconstruction for the purpose of analysis can help to 'make explicit how different modes work together to create meanings' (Flewitt, 2009: 52).

4.5.2 multimodal extracts and sequence of stills

Modal density tables can help to visually represent the configuration of modes but not the sequential synchronous and asynchronous appearance of modes and their temporal timescales for the action. The multimodal extract below (Extract 4-1) was developed to show the temporal sequence of actions, using the top row to show the timing, and the subsequent rows the appearance of the most dominant modes with a description. A simplified Jefferson notation is used with visual symbols and this is provided at Appendix A. When deemed helpful, a sequence of stills is used in conjunction with the extracts to help illustrate a visual dimension to the transcription. This section is used to introduce and explain how the multimodal extracts are used in the later chapters.

In the analysis chapters a conventional transcript is provided first (e.g. Transcript 4-1) and this is followed by the multimodal extract (e.g. Extract 4-1). In this example it can be seen that Sakia asks 'what next' Gamda replies with rising intonation 'mhm↑' and after a one second pause says 'click.'

Transcript 4-1 (00:03 – 00:04)

00:03	Sakia	what next

00:04	Gamda	mhm↑ (1) click

The multimodal extracts provide the additional layers of 'context' which accumulatively build a fluid profile of what peer-interaction looks like second-by-second with multiple opening and closing sequences of modes; most of which merge synchronously with other modes with fluctuating temporal durability. For discussion purposes a numbered reference point is used in the left-hand column and that is used in conjunction with the timing on the top row. For example, 'at Ref: 1(3s)' refers to Sakia talking at reference point 1 down the left-hand column and 3 seconds across the timing row.

It thus becomes easier to visually illustrate multiple sequences of actions for representation and for analysis. Transcribing like this is a matter of watching the combined videos at macro, meso and micro levels, playing, pausing, rewinding, and playing back in slow motion in a constant iterative manner. Ten seconds of interaction can take sixty minutes of multimodal transcribing. The process is one of factually describing second-by-second all the observable phenomena, and then, analysing what is happening. This extract is used as a short introduction to understand how to 'read' the later extracts in conjunction with the transcription symbols at Appendix A.

Ref:		Seconds:	1s	2s	3s	4s	5s	6s
1.	Sakia	talk	what next					
2.		actions						
3.		keyboard			------google------			
4.		screen			---------- opens explorer			
5.		mouse			----------			
6.		gaze	----------		---------		-------	----
Ref:		Seconds:	1s	2s	3s	4s	5s	6s
7.	Gamda	talk					mhm↑	click
8.		actions			-------			
9.		keyboard						
10.		screen						
11.		mouse						
12.		gaze	----------		----------		-------	----

Extract 4-1 (00:01 – 00:06)

In Extract 4-1, Ref: 5(1s) Sakia holds the mouse. She can be seen looking at the keyboard and monitor at Ref: 6(1s). The mouse cursor moves to the top left of the screen at Ref: 4(1s) and she opens Internet Explorer. Gamda is seen watching the screen from Ref: 12(1s) but then she writes something on her worksheets at Ref: 8(3s). At Ref: 3(3s) Sakia types 'google.co.uk' in the search box and simultaneously turns to Gamda at

Ref: 1&6(3s) to ask ‘what next.’ Gamda turns briefly to face her at Ref: 12(5s) and says ‘mhm↑’ and then says ‘click’ as she points at the screen at Ref: 8(6s). Sakia turns to face Gamda at Ref: 6(6s)

Some elements of this description can also be displayed visually with a sequence of stills as in Figure 4-3, though screenshots on their own provide only a limited account of what is happening. Sakia’s speech is presented in red and Gamda’s in black. At 00:03 Sakia is looking at the screen shortly before she turns to Gamda and says ‘what next.’ At 00:06 Sakia turns to Gamda and gesturing at the screen says ‘click.’ The image stills in themselves do not provide much communicative detail but they can help to visually contextualise the interaction when used in conjunction with the extracts and commentary.

Figure 4-3 (two learners interact)

To conclude, this section was used to briefly explain how the multimodal transcriptions are presented in the analysis and discussion chapters. Analysis has not been presented here, only description and explanation of the interaction to familiarise the reader. The challenge is how to present highly-detailed information in ‘readable ways’ (Flewitt, 2009: 47) without saturating the reader with detail. Inevitably there are still selections and deselections of data, so whilst multimodal transcribing attempts to capture more detail than is provided in traditional spoken transcripts, it must still reduce social elements of interaction for the purpose of presentation and discussion. The videos of the focal events are presented on a CD accompanying this thesis and will be referenced at the appropriate time. The focal events are described in the next section.

4.6 the focal events

This section is brief to explain the focal events chosen, how these were identified and which participants were chosen for detailed analysis. Chapters 5 and 6 are the analysis chapters of two focal events. The first stage of analysis was a narrative description in Transana to describe observable phenomena and identify higher-level actions; (e.g. Appendices B and C). An example of a narrative transcript was provided and discussed at Table 4-1. The narrating of the Polish learners' interaction over an entire lesson comprised 23,000 words, identifying every gesture, every spoken word, mouse-switch, keyboard-switch, etc. The same was done for the Kurdish learners. The preliminary transcribing of these two pairs of learners took nearly four months to complete. Switching between Transana for visual analysis and Audacity for audio analysis enabled increasing accuracy through triangulation of multiple audio sources. Multiple passes through the visual data was required to identify all actions. In section 4.2 an action was defined as goal-orientated outcomes realised through physical means. Actions are understood as observable phenomena and it is possible to bracket the start and end points of achieved actions; for example, finding pictures for a document on the internet starts with opening a browser, typing keywords, downloading the required images and inserting them into the documents. It is clear when that action starts and ends and what the sequence of lower-level actions are between the two. These multiple higher-level actions present themselves as potential focal points.

Having completed a full narrative transcription for the two pairs of learners it was evident that I was 'drowning in data.' I needed to prioritise my time-management and points of interest because there were so many potential focal points. A full narrative transcription of the French and Arabic learners was not completed because in preparing the data for analysis it became clear from the audio and video streams that there were complications. The French learners were male and female and clearly did not like each other. They argued constantly about their own design choices and mostly in French, creating significant problems for translated transcription. The Arabic learners relied heavily on the LSA so did not qualify as well as the others as a model of peer-interaction. In addition, one of the Arabic learners turned off her voice recorder when she went to use the toilet and did not press record when she re-entered the classroom so a great deal of her conversation was lost, though she could be heard in some sequences on her peer's recorder. The quality of their audio-data, and lack of independent peer-interaction, were thereby problematic.

The Kurdish and Polish learners showed the greatest level of peer-interaction independent of the tutor and the LSA. Both pairs were sat adjacent to each other so the videoed interaction of each pair was also caught on the meso video recorders of the other pair. This proved helpful for example when Darras's back was turned to the camera, his face could be seen on the Polish meso camera. Being adjacent also provided cross-fertilisation of ideas and competitiveness between the two pairs. The types and number of higher-level actions (Appendices B and C) were broadly similar. The pairings were comprised of females and males; both of whom socialised outside the classroom. They both wrote similar lengths of texts and of comparable accuracy. However, their frameworks of participation (including conversation, tool-use, spatial positioning, etc) looked different even though they both completed the same outcomes. Initial impressions suggested positive and contested frameworks of collaboration so these would present interesting comparative discussion. So whilst there was plenty of commonality across the two pairs to merit dual-analysis, there were also differences in how they interacted but achieved the same outcomes.

Throughout the piloting phase and the interviews a recurring theme had been a concern with second-language writing. The preliminary analysis of narrating the video revealed a sophisticated modal configuration of semiotic resources across all the actions. However, the coordination of off-screen talk and on-screen text appeared to be particularly complex, enough to merit microanalysis. A review of the literature, discussed at section 2.6, indicated there was very little comparable research where second-language speakers collaborate in real-time at a shared computer to produce writing: 'the number of empirical studies that have investigated collaborative writing in L2 classes is relatively small (Storch, 2011: 277). Recent publications in L2 collaborative writing suggested this was still a relevant, if under-researched discipline (e.g. Ajmi & Ali, 2014; Dobao, 2014a, 2015; McDonough et al, 2015).

In addition, the literature on embodied peer-interaction in second-language contexts was also minimal, as Gullberg explains: 'Despite popular convictions that L2 learners use all means at their disposal to communicate, their multimodal behaviour has received surprisingly little attention, both descriptively and in theorising about L2 acquisition and use. Moreover, the focus has largely been on the individual learner in isolation from the interactional and multimodal context where the problems typically arise' (2011:

137). There was some similar research to my own, where screen-capture software had been used to provide an audio-visual perspective on literacy and interaction, such as Gardner & Levy (2010) in a high-school and Bhatt & de Roock (2014) in a second-language classroom. Gardner & Levy offered some tentative findings on the coordination of talk and writing but Bhatt & de Roock did not.

My general focus of research interest towards literacy process over product created an inclination towards a more detailed analysis of real-time collaborative writing. This was reinforced by the three reasons discussed above: (1.) the quality of data resulted in a natural deselection of some video footage; (2.) the grounded relevance of L2 writing raised by the learners; (3.) gaps in the L2 writing and multimodal interaction literature. For all these reasons, the higher-level action of collaborative writing across two pairs of learners was the focal point chosen. For both pairs of learners this is action 1.7 in the overall classroom sequences shown at Appendices B and C.

4.7 conclusion

This chapter discussed the analytical approach to actions, interaction and modes of communication. This was provided in its own chapter to explain the grounded approach to video analysis and to introduce the types of multimodal transcription used to represent embodied peer-interaction. Three types of transcription were introduced and explained. These will be used extensively in Chapters 5, 6 and 7. (1.) *modal density tables* show the configuration of modes with smaller or larger squares indicating the importance of the modes to the completion of the action; (2.) *multimodal extracts* show the unfolding synchronisation of modes and actions on their own timelines; a simplified Jefferson notation is used with visual symbols; (3.) *sequence of stills* with spoken language and overlaid arrows indicating direction of gaze and gesture.

Norris's methodological framework was presented as a tool-kit for approaching interactional analysis of video data. As she explains, 'the reason that I developed this framework was my belief that we can only truly understand interaction if we investigate the visual channels of communication as well as the audible channels … we are really only at the beginning of multimodal interactional analysis' (2004: 148). Key terminology for multimodal interactional analysis, presented in this chapter, includes:

- scales of time
- higher-level actions

- lower-level actions
- sequential and parallel actions
- awareness/attention of actions on a foreground-background continuum
- modal density (structured by modal intensity or/and complexity)
- modal configuration
- modal value (superior/subordinate)
- modal affordance and constraints
- inter-semiotic relations
- semiotic resources
- interactional awareness
- environmentally coupled gestures
- embodied modes
- frozen actions within disembodied modes.

A preliminary analysis of interaction was provided to illustrate the application of the terms. Peer-interaction frameworks are here constituted on two levels of interaction. Preliminary analysis of the video data identified (1.) higher-level actions (what the learners are trying to achieve, over different scales of time, and the start and end points of these actions) and (2.) the lower-level actions (what the learners are doing, on a shorter scale of time, through the configuration of resources and communicative modes) to help them complete the higher-level action. Actions become the units of analysis for re-framing past activity. And whilst most actions were sequential, some actions ran parallel to each other; as can be seen at Appendices B and C.

The methodological discussion in Chapters 3 and 4 has identified the complexity of interaction which might be understood as a hierarchy of sophistication, from peer-interaction framework at a top level (as the organisational structure within which interaction occurs) down to the configuration of higher-level actions and lower-level actions which occur synchronously, asynchronously and sometimes too quickly for the human eye. Multiple videoing and audio recordings were used to capture on-screen and off-screen interaction to witness real-time linguistic and non-linguistic configurations. Combining the multiple videos and synchronising them, and slowing down the playback, was considered to be the best method to identify actions and the configuration of modes within peer-interaction. Chapters 5 and 6 will begin the analysis of the two focal events. The microanalysis from these two chapters will inform the discussion at chapter 7.

5 Analysis 1

5.1 introduction

This chapter describes the interaction between Sakia and Gamda and the resources and modes they use to complete the learning outcomes of the lesson. Chapter 6 will provide similar analysis between Darras and Shourok, with comparisons and contrasts made between the two pairs of learners in the discussion at chapter 7. In this chapter close analysis will pay attention to the cross-modal configurations in interaction and the relevance of the different modes to the production of two sentences of on-screen text. The focal point chosen provides an insight into the real-time design process where language, literacy and technology come together. It is a similar focal event in chapter 6 where the collaboration of the Kurdish learners offers interesting contrast to the Polish learners. The video of their interaction is available on the accompanying CD called **Polish_1.7**. The totality of language for analysis in this chapter, between sections 5.1 and 5.6, is as follows. The transcript conventions are at Appendix A.

00:01 Sakia	<°environmentalism°> (2) zaraz sprawdzimy	{we're going to check}
00:07 Gamda	mhm↑	
00:13 Sakia	en: vi:	
00:15 Gamda	ron >mental<	
00:24 Sakia	ekolog (6) ekologia↑ (14)	{activist (6) ecology}
00:40 Gamda	we can speak about recy [cling	
00:43 Sakia	[aha	
00:44 Gamda	and on website	
00:46 Sakia	yea (11) like we should more care about recycling (2) if we (4) we want	
01.10 Gamda	want to keep the	
01.11 Sakia	want to keep our planet	
01.14 Gamda	as [long	
01.14 Sakia	[nice and clean	
01.15 Gamda	mhm	
01.25 Sakia	haha	
01.25 Sakia	haha	
01.28 Gamda	best fit	
01.35 Sakia	maybe in the middle lepiej	{that's better}
01.57 Gamda	a few sentences ( ) more than one	{they both laugh}
02.05 Gamda	if we want to keep our planet	{reads from the screen}
02.09 Sakia	clean and healthy↑	
02.10 Gamda	mhm	
02:12 Sakia	( )	{unclear}
02:14 Gamda	aha	
02.24 Sakia	or we should <u>more</u> care about recycling↑ (6) I never know if it should be comma or not	
02.36 Gamda	ah	
02.39 Sakia	no [I don't think	

02.40 Gamda	[mhm	
02:42 Gamda	but (1) we can write it as well (1) like	
02.45 Sakia	mhm	
02:50 Gamda	it like (1) only takes (1) a few seconds to sort them out separately	
02.56 Sakia	aha	
02:59 Gamda	and	
03:14 Sakia	s e pa	
03:15 Gamda	ra sepa <separately>	
03:18 Sakia	separately nie wiem	{I don't know}
	°it takes only few seconds to put the litters separately°	
	{reads from the screen}	
	{at this point there are over four minutes of inactivity when nothing is written}	
07:56 Sakia	it takes only few minutes	
07:58 Gamda	to put the litters in order	
08:08 Sakia	and [prevent	
08:08 Gamda	[prevent	
08:10 Sakia	the earth from degradation (08) from what↑	
08:20 Gamda	from (1) zaśmiecenie	{littering}
08:27 Sakia	from pollution (08) jeden odcinek	{that's one part}
08:42 Sakia	zaśmiecenie to ok a degradacja	
	{littering is ok but what about degradation}	
08:43 Gamda	mhm	
08:44 Sakia	degradacja	{degradation}
08:49 Gamda	myśle że to (pollution)	{I think that's it}

Transcript 5-1 (full transcript)

The following sections will take extracts from the total language exchange above and offer multimodal transcription to analyse what else is happening in addition to talk. Section 5.2 analyses the learners' interaction and their switching of languages, considering multimodal alignment, as the learners try to understand the word 'environmentalism.' There is evidence of *languaging* (Swain, 2006) and *translanguaging* (Garćia, 2007) as they negotiate meaning through L1 and L2. Section 5.3 will analyse a more extended piece of dialogue, through a conversation analysis approach, as the learners think out loud what they should write for their first sentence. 5.4 considers the transmodal process of moving off-screen talk to on-screen text and the temporal mapping required to do this as they write their first sentence which appears on-screen as: '*We should more care about recycling if we want to keep our planet clean and healthy'*. All on-screen text typed by the learners will be presented in the analysis using *italics*. Section 5.5 discusses the first half of their second sentence: '*It takes only few minutes to put the litters in order.*' Section 5.6 examines the second half of their second sentence: '*and prevent the Earth from pollution.*' Sections 5.5 and 5.6 will analyse the off-screen and on-screen production of the second sentence and consider the

features of talk required to sustain this process. The need to interrogate the polyvocality of linguistic practice and its sources is a central finding here. 5.7 will summarise the main findings from the microanalysis of the learners' interaction with each other and with their writing.

5.2 understanding the task

Sakia and Gamda collaborate to write text for the fourth page of their booklet. They are using Microsoft Publisher. There is an identifiable beginning (a blank page on the computer) and end (completed writing). Within the overall classroom sequence this focal event is higher-level action 1.7 which sits in the longer first action sequence of creating a booklet about environmentalism; (see Appendix B). Immediately preceding the task of writing, for page one of their booklet, they designed an image on environmentalism at Figure 5-2; 'recycling' being their chosen topic. This image helps to contextualise the following interaction and their language choices as they decide on what to write.

Figure 5-1 (learning outcome 2 - create a protest image)

At the beginning of the 1.7 higher-level action, the learners write a title for the page called 'My Thoughts' and insert a recycling image they found on the internet. They then

add a textbox under the image in which to write their text. The completed text is two sentences long and when finished reads as: '*We should more care about recycling if we want to keep our planet clean and healthy. It takes only few minutes to put the litters in order and prevent the Earth from pollution*'; see Figure 5-2. The sequence of actions in focal event 1.7 is complete when the Polish learners finish the writing.

Figure 5-2 (page four of recycling booklet)

The Polish learners need to agree on what text to write under the recycling image. They use the classroom instructions to guide them. They use a Polish website and an electronic translator. The instructions state: 'For page 4 of your booklet create a header called 'My Thoughts' and add a text box. Write a few sentences about Greenpeace or

environmentalism. What do you think?' Analysis begins from the creation of the textbox under the recycling image when Sakia questions the word 'environmentalism.'

The first three extracts for analysis last thirty seconds. There is a chain of lower-level actions cumulatively building towards the completion of the higher-level action of writing the sentences for their booklet. In the first sequence of lower-level actions the learners do not fully understand the word 'environmentalism' in the teaching instructions. The opening action begins at Transcript 5-2 and Extract 5-1 and their uncertainty when first reading the word and ends with Extract 5-3 when they understand what the word means. The outcome is completed and this in turn guides them successfully onto the next sequence of actions. Over a thirty second period the verbal output of both learners is as below.

Transcript 5-2 (00:01 – 00:24)

00:01 Sakia	<°environmentalism°> (2) zaraz sprawdzimy	{we're going to check}
00:07 Gamda	mhm↑	
00:13 Sakia	en: vi:	
00:15 Gamda	ron >mental<	
00:24 Sakia	ekolog (6) ekologia↑ (14)	{activist (6) ecology}

Of interest is the inter-semiotic configuration of their actions, including translanguaging, environmentally coupled gestures, alignment and technology. Extract 5-1 to Extract 5-3 will delineate this transcript through multimodal transcription and microanalysis.

Extract 5-1 (00:01 – 00:09)

Ref:		Seconds:	1s	2s	3s	4s	5s	6s	7s	8s	9s
1.	Sakia	talk	<ºenvironmentalismº>		**zaraz sprawdzimy**						
2.		actions									
3.		keyboard									
4.		screen						ling.pl			
5.		mouse									
6.		gaze									
Ref:		Seconds:	1s	2s	3s	4s	5s	6s	7s	8s	9s
7.	Gamda	talk							**mhm↑**		
8.		actions									
9.		keyboard									
10.		screen									
11.		mouse									
12.		gaze									

The mirroring behaviour of Gamda is made visually evident by the staggered colour coding. The staggered start times of the yellow, green and blue actions along the talk and gaze timelines evidence the time-lag between Sakia and Gamda. In the gaze timelines at Ref: 6 and Ref: 12, Gamda lags approximately a second behind Sakia as Sakia leads the interaction. Gamda's time-lagged actions mirror Sakia, similarly turning to the worksheets to look at the instructions and then to face Sakia and then to talk; all of which are imitating behaviour. Sakia's focus of attention at Ref: 6(4s) shifts to the screen as she looks for an online translation of 'environmentalism.' Gamda feigns a proactive response, matching Sakia's behaviour one second later, but turning her focus of attention instead to her electronic translator. We see proxemic and linguistic alignment.

The mouse was available up to this point for Gamda to use but, in dance terms, Sakia leads and Gamda follows, Gamda using the remaining resources available to her. Gamda could have chosen to do nothing as Sakia progressed to the online translation site, leaving Sakia to do the work. I say 'feigns a proactive response' because Gamda picks up the translator but does not use it. She picks it up, opens and closes the lid, but in only four seconds and quickly changes her focus of attention to the screen to watch Sakia's interaction with www.ling.pl. This feigned mirroring behaviour would suggest positive alignment. Gamda demonstrates a willingness to 'pull her weight' in the task. Though Gamda appears to do very little and contributes nothing in the first ten seconds, her shifting gaze and time-lagged mirroring would suggest she is fully engaged in the activity.

Figure 5-3 (00:01 to 00:06)

Sakia's talk is in red font and Gamda's is black. In the first nine seconds there are three verbal utterances, 01, 03 and 06 seconds as above. The primary catalyst for the interaction is the written word 'environmentalism' which Sakia speaks aloud at 00:01. It is said slowly and quietly as she faces the worksheets, which would suggest that the

speaking of the word is for Sakia's own purpose rather than being communicative. She wants to understand the meaning of the word. Gamda imitates Sakia at 00:02 seconds by turning her attention to the worksheets; presumably to find the word Sakia has spoken. At 00:03 seconds Sakia turns to face Gamda and says in Polish 'zaraz sprawdzimy,' meaning 'we're going to check.' Gamda's only verbal response is a confirmatory 'mhm' at 00:06 when she turns to face her. Talk in this interaction would not appear to be the best mode to solve the problem of what is meant by the word 'environmentalism.' Instead, they use different tools: the internet and an abandoned attempt with a translation device. Each learner could have talked about the word, or raised their hand to attract the teacher's attention, but the resources they have at hand give them independence. There is still a significant amount of interaction in the first nine seconds but talk has a lesser role in the meaning-making process because other modes and tools would appear to be more beneficial.

Extract 5-2 (00:09 – 00:18)

Ref:		Seconds:	9s	10s	11s	12s	13s	14s	15s	16s	17s	18s
13.	Sakia	talk		en:		vi:						
14.		actions										
15.		keyboard		e n v i r o n m e n t a l i s m								
16.		screen	⇨-----⇨	e n v i r o n m e n t a l i s m								
17.		mouse	🖱-----🖱									
18.		gaze	💻----------💻📖----------📖 💻------💻📖----------📖💻----------💻									
Ref:		Seconds:	9s	10s	11s	12s	13s	14s	15s	16s	17s	18s
19.	Gamda	talk							ron	>mental<		
20.		actions										✋
21.		keyboard										
22.		screen										
23.		mouse										
24.		gaze	📖-------📖❑----------❑📖----------📖❑----------❑									

The transcript continues from the ninth second of their interaction as they proceed to try and understand what is meant by the word ‘environmentalism.’ Of interest in Extract 5-2 is the coordination of talk, reading, writing on-screen, and the completion of each other’s talk as they linguistically align. The same word is repeated, ‘environmentalism,’ but its production is very different from Extract 5-1. At Ref: 16-18(9s) Sakia shifts her gaze to the computer screen and using the mouse clicks inside the search box at www.ling.pl. She begins to type the word ‘environmentalism’ at Ref: 15(10s) as she pronounces the first two syllables slowly at Ref: 13(10-12s). Each letter appears slightly later on the screen, 16(10s-18s). Sakia’s gaze at Ref: 18 alternates between the keyboard and the screen and back to the worksheets, on average every second, as she checks the spelling on the worksheets to inform her typing of the word on the screen. Gamda’s gaze at Ref: 24 switches between the worksheets on the table and the screen as she appears to be checking the spelling of the word ‘environmentalism’ as Sakia types it in the search-box at www.ling.pl.

At Ref: 19(15s-17s) Gamda takes over the verbalisation of the word (perhaps as a form of linguistic scaffolding to help with the spelling) and speaks aloud the penultimate three syllables, 'ron' and 'mental.' Their talk is slowed down to temporally map the sounds off-screen onto the on-screen typing of the word. As the final letters of the word appear in the www.ling.pl search box Gamda raises her hand at Ref: 19(18s). This is possibly a metaphoric gesture from Gamda, open palm, back of hand, which would appear to suggest a stop signal as Sakia reaches the end of the word.

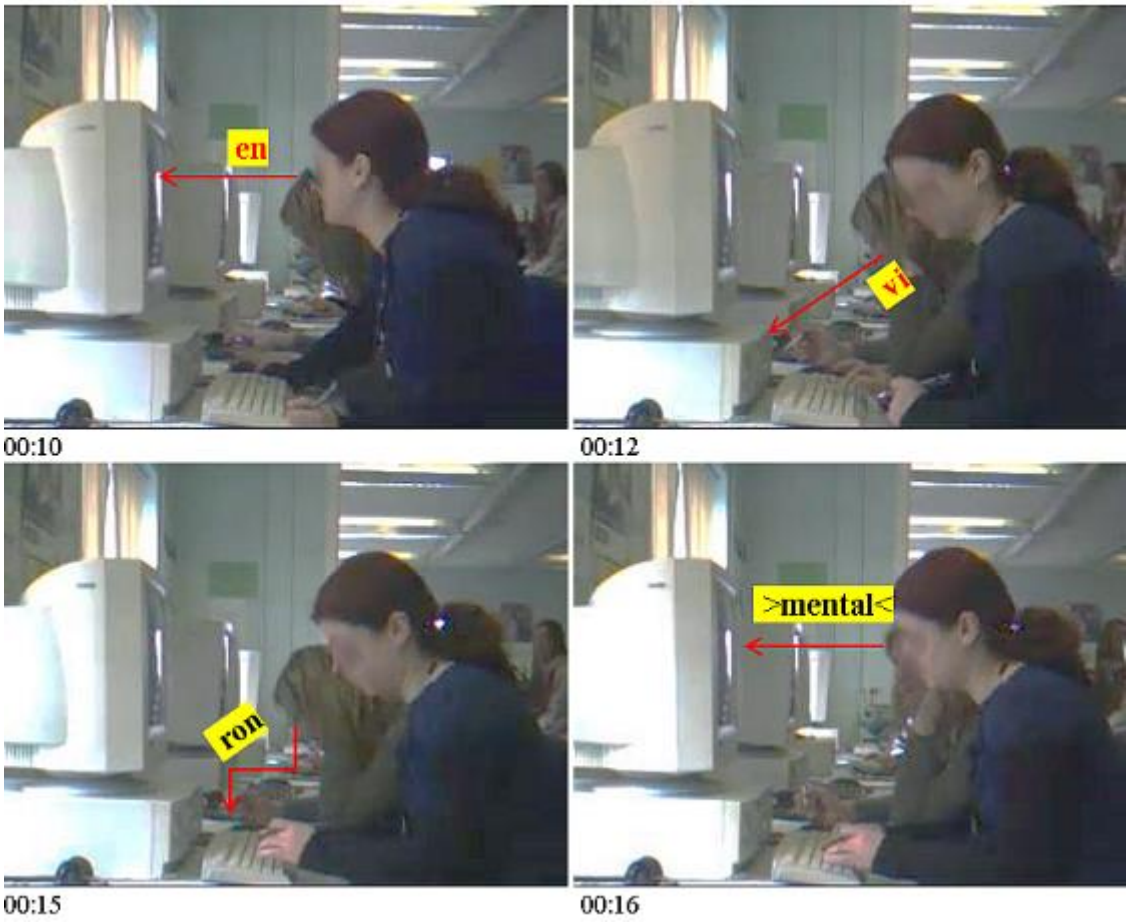


Figure 5-4 (00:10 - 00:16)

In Extract 5-2 there are four verbal utterances. Talk in these nine seconds (albeit a single word) is punctuated to syllabic utterances; punctuated by the slowness of (1.📖) reading the word 'environmentalism' on the worksheet (2.📻) looking at the keyboard to find the corresponding letters as Sakia simultaneously pronounces a syllable at a time (3.💻) looking up to the screen to check the correctness of the typing and (4.📖) back to the worksheet to look at the word again. In the typing of a single word there is significant cross-modal coordination by Sakia. However, typing and speaking of the word do not occur in isolation. The word 'environmentalism' takes eight seconds to type. It takes

Sakia a lengthy two seconds to say aloud the first two syllables and six seconds for the complete articulation of the word across two people. Sakia has control of the mouse and the keyboard and is not dependent on Gamda. In collaborative computing it would be easy to assume that the person in control of the mouse and keyboard is the one most active in the collaboration. A quick modal comparison would suggest that Gamda is very active in Extract 5-2. Sakia reads, speaks and types. Gamda read, listens, speaks and gestures. What evidence is there that she is listening? Gamda starts speaking, and perhaps overlaps Sakia on the third syllable of the word 'ron,' leading directly on from the second syllable spoken by Sakia 'vi.' Gamda speaks the syllable 'ron' as she reads from the worksheets (15 seconds) suggesting she is synchronised with Sakia's reading and verbalisation. She then looks up to the screen and speaks the next two syllables 'mental' very quickly. Noticeably her gaze is directed at the screen, possibly to check how much of the word had already been typed. At no point in Extract 5-2 does either learner look towards the other. Sakia's multimodal episode is isolated to herself. Gamda's leads on from Sakia and is fully engaged with the actions of Sakia.

In Extract 5-2 there is mode-switching from reading in print to verbal utterance to writing digitally. Mode-switching is a prevalent action in this sequence and the synchronisation of modes is revealing. Linguistically, Gamda aligns successfully with Sakia. Sakia's actions are isolated to herself as her attention is focussed on typing. To maintain the interaction, Gamda synchronises with the actions of Sakia and she does this successfully through linguistic alignment; albeit one word. Synchronisation, alignment and mimicry would appear to be features of an inclusionary peer-interaction framework.

Extract 5-3 (00:23 - 00:32)

Ref:		Seconds:	23s	24s	25s	26s	27s	28s	29s	30s	31s	32s
25.	Sakia	talk		ekolog						ekologia↑		
26.		actions										
27.		keyboard										
28.		screen								----------		
29.		mouse								----------		
30.		gaze	----------									
Ref:		Seconds:	23s	24s	25s	26s	27s	28s	29s	30s	31s	32s
31.	Gamda	talk										
32.		actions								----------		
33.		keyboard										
34.		screen										
35.		mouse										
36.		gaze	----------							----------		

From 18 to 23 seconds the two learners sit motionless as they wait for the website to provide a translation of ‘environmentalism.’ Extract 5-3 follows on from second 24. Sakia is first to react when the translation appears and she reads aloud from the screen ‘ekolog’ at Ref: 25(24s). Both learners pause for six seconds, perhaps reading the full translation. The sentence on-screen following on from ekolog is as follows: ‘ekolog; dzialacz na rzecz ochrony środowiska’ which translates word-for-word as ‘ecologist; activist for environmental protection’ but in general Polish terms would be understood to mean ‘environmental activist.’ After the six second gap, Sakia moves the mouse down from the search box to the on-screen word ekolog and says aloud ‘ekologia↑’ with rising intonation; Ref: 25(30s). It is uncertain if this word is meant for Gamda or just a verbal means for confirming to herself that she now understands the word ‘environmentalism.’ The fact that her gaze is on the screen would suggest the latter, though Gamda is likely to have heard her. At the same time as Sakia says the word ‘ekologia’ Gamda begins to write on her worksheets, Ref: 32(30s).

Sakia's movement of the cursor to the word ekolog acts like a visual gesture, a proxy for a finger that points to draw attention to something, but it is uncertain if this gesture is meant for Sakia or Gamda; though Gamda responds immediately to the gesture and the verbal ekologia by writing something on her worksheets. Exactly what she writes is uncertain but presumably it is a translation of the word environmentalism.

The word ekologia is not on the screen. Also, Sakia has a rising tone when she pronounces the word. The rising tone does not suggest uncertainty, in the form of a question, as the learner's ensuing actions would suggest that they now understand the word environmentalism. Rather, the rising tone might suggest an eureka moment, an element of surprise. Sakia draws on her first-language and is able to supplement the less than perfect online translation of ekolog into something that makes sense to her. Ekolog is a noun for a person who is active in preserving the environment. There is a direct translation here in Polish for environmentalist. It is however a less than perfect translation for environmentalism. Ekologia is a noun for the subject of ecology; i.e. the study of the relationship between organisms and their environment. Again, it is less than perfect but is a best-fit. While the ling.pl website has modal affordances in terms of quick access to translation and definitions of unknown words, there are also constraints. In this scenario the ling.pl website failed to give an accurate translation of environmentalism. 'Ekologizm' is a better translation and is provided by other English-Polish translation websites. Sakia, perhaps unaware, has mitigated this problem by substituting ekolog with her own word ekologia. The learners seem happy with this interpretation and the action is complete. They use the outcome of this interaction to inform what they do next. In light of their new understanding they begin to discuss the specific content of the sentences they have been asked to write about.

00:24 00:30

Figure 5-5 (00:24 - 00:30)

Looking back on Extract 5-1 to Extract 5-3 and the completion of these lower-level actions, how are the communicative modes configured to help them achieve their aim? Interaction between the two is achieved through a number of modes which can be analysed separately but in real-time they are collective, they mimic, they align, they synchronise. Single utterances contribute to a chain of utterances. A gaze is returned. A glance at a worksheet is mirrored. A spoken word is overlapped. Over a thirty second period the entire verbal output of both learners is as below:

Transcript 5-3 (00:01 – 00:24)

00:01 Sakia	<°environmentalism°> (2) zaraz sprawdzimy	{we're going to check}
00:07 Gamda	mhm↑	
00:13 Sakia	en: vi:	
00:15 Gamda	ron >mental<	
00:24 Sakia	ekolog (6) ekologia↑ (14)	{activist (6) ecology}

Modal deconstruction in the extracts so far has revealed how language(s), literacy and technology (LLT) combine in a productive but unexpected way: in this instance, independent translanguaging strategies evidenced through the technology of a personal translation device and the internet; noticeably not a bilingual book. The temporal ordering of the different modes structures the sequentiality of the interaction and helps them to achieve understanding. Both are involved and neither is marginalised. The coordination of LLT is explored further in the next section as the learners begin to formulate sentences verbally off-screen before writing on-screen.

5.3 composing sentence 1

The learners sit quietly for fourteen seconds as they think how best to continue with the task of writing about environmentalism. This silence is a valuable moment as it enables them to reflect on the recent translations of 'ekolog' and 'ekologia' and to think about the specific content of the forthcoming writing. They choose recycling and this suggestion comes from Gamda; it is then adopted by Sakia who begins with a few words and the sentences develop between them through off-screen dialogue and on-screen typing. A product approach to writing would focus on the finished text as a means for making learning visible. Here the focus is on the off-screen negotiation of writing, the fluid real-time discussion and modal alignment of the learners to each other and the objects around them, to understand the sequentiality of how talk becomes on-

screen text and if there is evidence of linguistic scaffolding between the learners; i.e. meta-talk about the writing where the learners help each another with language construction. Below is a transcription of the talk for this section of analysis, followed by multimodal transcription.

Transcript 5-4 (00:40 – 01:25)

00:40 Gamda	we can speak about recy	[cling
00:43 Sakia		[aha
00:44 Gamda	and on website	
00:46 Sakia	yea (11) like we should more care about recycling (2) if we (4) we want	
01.10 Gamda	want to keep the	
01.11 Sakia	want to keep our planet	
01.14 Gamda	as	[long
01.14 Sakia		[nice and clean
01.15 Gamda	mhm	
01.25 Sakia	haha	

Extract 5-4 (00:40 - 00:49)

Ref:		Seconds:	40s	41s	42s	43s	44s	45s	46s	47s	48s	49s
37.	Sakia	talk			[aha				yea {long pause of 11 seconds}			
38.		actions										
39.		keyboard										
40.		screen										
41.		mouse	{releases mouse and sits back in chair}									
42.		gaze	❑----------		❑🗣----------			🗣⌨----	⌨📖----------			📖
Ref:		Seconds:	40s	41s	42s	43s	44s	45s	46s	47s	48s	49s
43.	Gamda	talk	we can speak about recy [cling				and on website					
44.		actions										
45.		keyboard										
46.		screen										
47.		mouse										
48.		gaze	❑----------						❑😐---{looks over shoulder}---😐			

In Extract 5-4, Ref: 43(40s) Gamda takes the initiative and suggests recycling as a topic for their writing. At the same time Sakia releases the mouse and sits back in her chair and overlaps Gamda's talk at 42s with an 'aha' to show her agreement. Sakia also shifts her gaze from the screen to Gamda as she says 'aha.' At Ref: 43(44s) Gamda adds, thinking forwards to the next task, that they can also use the same topic for the design of their website: 'and on website.' Sakia quickly agrees at Ref: 37(46s) with a 'yea' and then both learners don't speak for eleven seconds.

Figure 5-6 (00:40 - 00:46)

There is evidence of linguistic alignment between the two as they verbally agree on what to write. This is evident in the turn-taking adjacency pair of suggestion (40s) confirmation (43s); suggestion (44s) confirmation (46s). But then the learners' proxemic alignment becomes disjointed. The reason for this temporary lack of alignment is possibly because of their uncertainty in how to proceed. They know they need to write something, and whilst they have agreed on a topic, the structuring of the sentences requires thinking and silence. In the eleven second phase of uncertainty which follows, Gamda looks over her shoulder at the tutor speaking to someone Ref: 48(46s). Sakia's gaze shifts from the keyboard to the worksheets and then out of the window.

Figure 5-7 (00:46 - 00:57)

For a total of 11 seconds (46s to 57s) Gamda is seemingly ‘detached’ from the task. Sakia sits very still in what appears to be deep thought. Gamda is preoccupied with what else is happening in the room and with the screen. There is an absence of proxemic alignment. Physically, Gamda is animated but Sakia sits passively with crossed arms (46s to 57s). The outward behaviour would suggest that one is deep in thought on what to write and the other is not. In Extract 5-5, Ref: 58-60(57s) Gamda holds the mouse and begins to ‘play’ with the objects on the screen. Sakia breaks the silence at Ref: 49(57s) and offers the beginning of a sentence they might write.

Extract 5-5 (00:57 – 01:06)

Ref:		Seconds:	57s 58s 59s 01.00 01.02 01.06
49.	Sakia	talk	like we should more care about recycling (2) if we (4) we want
50.		actions	----------{arms crossed}----------
51.		keyboard	
52.		screen	
53.		mouse	
54.		gaze	😐---------=----------{appears to be staring blankly out of the window as she thinks}--=----------😐
Ref:		Seconds:	57s 58s 59s 01.00 01.02 01.06
55.	Gamda	talk	
56.		actions	
57.		keyboard	
58.		screen	{Gamda repositions the text box and image on the screen as Sakia is talking}
59.		mouse	🖱----------🖱
60.		gaze	❑----------❑

'Like' can be used as a filler in communication but is perhaps used here as a preposition to suggest similarity, that the words she is about to say are not final, but that they may be similar to what is eventually written. The production of writing is rarely verbalised in such a way, spoken out loud before being written, so we can perhaps assume from this verbalisation that it is spoken for the benefit of Gamda. In this way, Sakia enables equal textual production (or linguistic equality) in the writing of the text; that she is not taking over but suggesting possibilities. This can be seen in Extract 5-5 and then Extract 5-6 when Sakia's verbalisation acts as a catalyst for Gamda and she offers a language chunk which logically and sequentially continues the sentence. Gamda picks up the tail end of Sakia's sentence 'want' at Ref: 49(01.09) and makes it a cognitive baton for the beginning of her sentence 'want to keep the' - at Extract 5-6, Ref: 67(01.10).

Extract 5-6 (01.10 – 01.17)

Ref:		Seconds:	01.10	01.11	01.12	01.13	01.14	01.15	01.16	01.17
61.	Sakia	talk		want to keep our planet			[nice and clean			
62.		actions	----------{arms crossed}----------							
63.		keyboard								
64.		screen								
65.		mouse								
66.		gaze	❑----------				----------❑	🗣---------{turns to Gamda}----------🗣		
Ref:		Seconds:	01.10	01.11	01.12	01.13	01.14	01.15	01.16	01.17
67.	Gamda	talk	want to keep the			as	[long	mhm (5)		
68.		actions					🗣---{rests her chin on her hand}------🗣			
69.		keyboard								
70.		screen	❑----------{moves the image}----------❑							
71.		mouse	🖱----------				----------🖱			
72.		gaze	❑----------				----------❑	🗣--------{turns to Sakia}----------🗣		

Repeating the final word(s) acts both, as recognition of the other speaker to show active participation, which promotes linguistic equality in the conversation, but also as a cognitive baton to build their first sentence. The colour coding is used to suggest where these cognitive batons are causally related. In the blue coding ‘want’ is repeated three times, from Sakia [Ref: 49(01.06)] to Gamda [Ref: 67(01.10)] and back to Sakia [Ref: 61(01.11s)], and in the yellow coding ‘to keep’ is copied by Sakia from Gamda. At Ref: 61(01.14) and Ref: 67(01.13) both learners follow their own linguistic sequencing and produce different language chunks leading on from Sakia’s ‘our planet’ at Ref: 61(01.12) and ‘nice and clean’ from Sakia compared to ‘as long’ from Gamda. Sequentiality momentarily breaks down but Sakia’s contribution makes more grammatical and contextual sense than Gamda’s. Perhaps it is for this reason that Gamda ‘defers’ to Sakia with the back-channelling ‘mhm’ at Ref: 67(01.15) to show she is listening and that she approves of this alternative. Sequentiality is restored and the grey colour coding at (01.14) shows through their gaze that they both proxemically re-align as they turn simultaneously to each other and make eye-contact.

Figure 5-8 (01.10 – 01.20)

At 01:01 Gamda is looking at the screen and continues the sentence Sakia started by suggesting ‘want to keep the.’ Sakia adapts Gamda’s suggestion at 01:11 with ‘want to keep our planet.’ At 01:14 they overlap each other in conversation and simultaneously turn to face each other as Gamda agrees with a ‘mhm.’ Gamda follows this back-channelling ‘mhm’ with another back-channelling technique 01.20 by turning to Sakia and smiling. There is mutual coherence on how to proceed; which is now to start typing on the screen.

5.4 writing sentence 1

In the previous section, the sequentiality of off-screen text collaboration was analysed. It was shown how the Polish learners verbally agreed that they would write the following sentence: ‘we should more care about recycling if we want to keep our planet nice and clean’ - which is highlighted in red in Transcript 5-5 showing who contributed which words.

Transcript 5-5 (00:46 – 01:14)

00:46 Sakia	yea (11) like we should more care about recycling (2) if we (4) we want
01.10 Gamda	want to keep the
01.11 Sakia	want to keep our planet
01.14 Gamda	as [long
01.14 Sakia	[nice and clean

The sentence as it is later finished on-screen is: '*We should more care about recycling~~,~~ ~~more~~ if we want to keep our planet cle~~e~~an and healthy.*' 'Clean and healthy' is changed from what they originally agreed, '*nice and clean.*' This section will micro-analyse the shift of talk to text and the sequential temporal mapping of off-screen talk to on-screen text; from one learner to another.

In Transcript 5-5, the composing of the text off-screen takes 28 seconds and there is chunking of the sentence, sequentially developed through a process of cognitive batons passed back and forth between them as a form of linguistic scaffolding. In Transcript 5-6, the on-screen typing of their sentence takes 01:15 seconds.

Transcript 5-6 (01:25 – 02:40)

01.25 Sakia	haha		
01.28 Gamda	best fit		
01.35 Sakia	maybe in the middle	lepiej	{that's better}
01.57 Gamda	a few sentences ( ) more than one		{they both laugh}
02.05 Gamda	if we want to keep our planet		{reads from the screen}
02.09 Sakia	clean and healthy↑		
02.10 Gamda	mhm		
02:12 Sakia	( )		{unclear}
02:14 Gamda	aha		
02.24 Sakia	or we should <u>more</u> care about recycling↑ (6) I never know if it should be comma or not		
02.36 Gamda	ah		
02.39 Sakia	no [I don't think		
02.40 Gamda	[mhm		

The following extracts provide a multimodal microanalysis of Transcript 5-6 to explore the talk and writing of the learners.

Extract 5-7 (01.19 – 01.27)

Ref:		Seconds:	01.19	01.20	01.21	0 1.22	01.23	01.24	01.25	01.26	01.28
73.	Sakia	talk							ha-ha-ha		
74.		actions									
75.		keyboard							W e		
76.		screen							W e		
77.		mouse			🖱----------						🖱
78.		gaze			💻----------		{turns to monitor and keyboard to type}				----------💻
Ref:		Seconds:	01.19	01.20	01.21	0 1.22	01.23	01.24	01.25	01.26	01.28
79.	Gamda	talk									best fit
80.		actions	----------{rests her chin on her hand}----------								
81.		keyboard									
82.		screen									
83.		mouse									
84.		gaze		☺-------{smiles at Sakia}----------			☺ ❑		----------{turns to monitor}----------		❑

The first word to be typed is ‘we’ at Extract 5-7, Ref: 75(01.25), and Transcript 5-6, 01:25. Sakia simultaneously laughs. It would have been unclear why she laughed at this point but with the screen-capture video it can be seen that the laugh coincides with the on-screen display of ‘*we*’ at Ref: 76(01.25). The writing of ‘*we*’ is so small it is barely readable and it is possibly this that causes the laugh. At Ref: 79(01:28) Gamda makes a suggestion to increase the font size when she says ‘best fit.’ Best Fit is a function for automatically resizing text in a text box so that it fills the box. From 01:35 to 01:57 (Transcript 5-6) Sakia writes very quickly: ‘*We should more care about recycling if we want to keep our planet.’* She makes a comment about the on-screen text at 01:35, changing it from left-aligned to centre-aligned: ‘maybe in the middle’ and adds a further comment in Polish: ‘lepiej’ {that’s better}. Meanwhile, Gamda looks at her worksheets at 01:57 and reads, speaking aloud: ‘a few sentences () more than one’. They both laugh at Gamda’s realisation that a ‘few sentences’ means ‘more than one’ so they have to think of more sentences than the one they have thought of so far. The laugh ends very quickly as Gamda brings them back on task a few seconds later. Extract 5-8 starts from 02:05 when Gamda reads on-screen what Sakia has finished writing at 01:57.

Extract 5-8 (02.05 – 02:18)

Ref:		Seconds:	02.05	02:07	02:08	02:09	02:10	02:11	02:18
85.	Sakia	talk				clean and healthy↑			
86.		actions				⌨------ pauses with hands over keyboard------⌨			
87.		keyboard						⌨-- cleean and healthy--⌨	
88.		screen							cleean and healthy
89.		mouse		🖱--------🖱					
90.		gaze	💻----------						----------💻
Ref:		Seconds:	02.05			02:09	02:10	02:11	02:18
91.	Gamda	talk	if we want to keep our planet				mhm		
92.		actions							
93.		keyboard							
94.		screen							
95.		mouse							
96.		gaze	❑----------			----------❑	🗣--------🗣	❑----------	----------❑

Gamda looks at the screen Ref: 96(02:05) and repeats what Sakia has written: ‘if we want to keep our planet’ – Ref: 91(02:05). At Ref: 89(02:06) Sakia holds the mouse and clicks on-screen at the point where she next wants to begin writing. Sakia then switches her hands to the keyboard [Ref: 86(02:08)] and holds her hands over the keyboard as she looks at the screen; she appears to be considering how to complete the sentence. There is a four second pause as they both sit in silence. In Transcript 5-5 (01:14) they agreed to continue the sentence ‘keep our planet’ with the phrase ‘nice and clean.’ This was verbally agreed but nothing was written down.

It would appear they have both forgotten this language chunk in the brief sixty seconds which have passed since they agreed. Instead, Sakia offers a logical alternative at Ref: 85(02:09) - 'clean and healthy↑' but says it with rising intonation as a suggestion. Gamda agrees with a confirmatory 'mhm' at 02:10 and turns to face Sakia at Ref: 96 as she speaks. Sakia types '*cleean and healthy*' at 02:11, misspelling and then deleting the double 'e'.

Between 02:18 and 02:24, when Sakia finishes the typing of '*healthy*', they are both silent and inactive. They both appear to be looking at the sentence on screen which reads as: '*We should care about recycling more if we want to keep our planet clean and healthy.*' It is a well-structured, grammatically correct sentence. As in any text composition, there is drafting, reading back and re-writing phases. Silence (or pausing) is a recurring mode in their collaboration and operates for different reasons: hesitation, self-correction, reading, thinking, typing, etc. Silence alongside the absence of any physical activity is very rare in their interaction; where there is no gesture, no head nor body movement. Sometimes their conversation is slowed down to help with the processing of ideas for sentence construction, as in the four second gap at 02:05 when Gamda says 'if we want to keep our planet' followed by Sakia at 02.09 saying 'clean and healthy↑'. Sometimes the silence would appear to be a pause to read for accuracy; as in the silence between 02:14 and 02:24. Both learners appear to be looking at the screen and they say nothing, until Sakia suggests an alternative to the sentence chunk: '*We should more care about recycling*' at Ref: 97(02:24).

Extract 5-9 (02.24 – 02:40)

Ref:		Seconds:	02.24	02:26	02:28	02:30	02:32	02:34	02:36	02:38	02:40
97.	Sakia	talk	or we should <u>more</u> care about recycling↑				I never know if it should be comma or not				no I don't think
98.		actions						☜ points at comma on screen		rubs hands	
99.		keyboard			⌨-- more--⌨			⌨----⌨ ,			⌨
100.		screen	⇖⇖⇖⇖⇖⇖		more		⇘⇘	~~m o r e~~ ,			;
101.		mouse	🖱---------🖱{clicks after *should*}					🖱--🖱{clicks after *more*}			
102.		gaze	💻----------					----------💻🗣--------🗣💻----------			----------💻
Ref:		Seconds:	02.24	02:26	02:28	02:30	02:32	02:34	02:36	02:38	02:40
103.	Gamda	talk							ah		mhm
104.		actions								rests chin on her hand	
105.		keyboard									
106.		screen									
107.		mouse									
108.		gaze	❑----------						----------❑🗣------🗣❑----------		----------❑

At the same time as she speaks ('or we should <u>more</u> care about recycling↑'), placing stress on '<u>more</u>' by saying it louder, Sakia moves the mouse cursor up the screen at Ref: 100(02:24) and clicks after the word '*should*' and types the word '*more*' at Ref: 99(02:28). It now reads as: '*We should more care about recycling more if we want to keep our planet clean and healthy.*' Gamda sits quietly watching. Sakia moves the mouse at Ref: 100(02:30) to the end of the second occurrence of '*more*' and deletes it. She then completes a number of actions simultaneously and with split-second time difference between the actions. At Ref: 97(02:32) Sakia says 'I never know if it should be a comma or not' as she presses the comma key on her keyboard and gestures at the screen to indicate for Gamda where she means. She turns to face Gamda at Ref: 102(02:34), as if to seek confirmation. Gamda mirrors Sakia's behaviour by turning to her at Ref: 108(02:35) and says 'ah'.

Sakia quickly dismisses the idea of a comma and deletes it at 02:40, saying 'no I don't think' as Gamda overlaps her with an agreeing 'mhm'. This action completes the first sentence. It went through the following composition and deletions: '*We should more care about recycling~~,~~ ~~more~~ if we want to keep our planet cleean and healthy.*' The learners contributed the following language chunks: Sakia, Gamda

We should more care about recycling if we want to keep our planet clean and healthy.

The composing of the sentence off-screen took twenty-eight seconds and there was chunking of the sentence, sequentially developed through a process of cognitive batons passed back and forth between Sakia and Gamda in a form of linguistic scaffolding. The on-screen typing of that sentence took 01:15 seconds and there was an absence of the linguistic equality and cognitive batons as seen in the previous sections, perhaps because they had already verbally agreed on what the sentence should be and Sakia only had to type it. Section 5.3, which examined the on-screen typing, showed there was little cross-modal activity by Gamda, compared to Sakia, but from a conversation analysis approach there was clear evidence that she was an active collaborator and engaged with the task. There is Gamda's comment on the design process 'best fit' at 01:24; then a critical comment on the requirements of the task, to write 'a few sentences' at 01:57; then reading what Sakia has already typed at 02:05 followed by four back-channelling signals (02:10 to 02:40) to show agreement.

01.28 Gamda	best fit	
01.57 Gamda	a few sentences ( ) more than one	
02.05 Gamda	if we want to keep our planet	{reads from the screen}
02.10 Gamda	mhm	
02:14 Gamda	aha	
02.36 Gamda	ah	
02.40 Gamda	mhm	

These back-channelling signals and snippets of language chunks evidence mutual coherence and function as indicators that Gamda is on task and sharing the collaboration; even though she is not contributing much to the on-screen typing. *Mutual coherence* is here used to suggest where there is agreed understanding, or a lack of, as seen later between the Kurdish learners. It is an important finding of the analysis which will be elaborated in the discussion chapter.

5.5 writing sentence 2

Whilst sentence one was quickly agreed verbally and then written without too much variation and difficulty, the second sentence in composition proves to be more difficult and they become dependent on alternative resources to complete the writing. The second sentence on completion reads as: '*It takes only few minutes to put the litters in order and prevent the Earth from pollution.*' The analysis of the second sentence will be divided into two sections. 5.4 will examine the first half of their sentence: '*It takes only few minutes to put the litters in order*'. 5.5 will examine the second half of their sentence: '*and prevent the Earth from pollution.*' Both sections will micro-analyse the real-time shift of talk to text by focussing on the interaction of the learners' talk and writing and the modes they utilise to help complete the process; it proves to be a polyvocal event combining language, literacy and technology. It is worth noting the activity and context which preceded their writing of these sentences for page four of their booklet. For page one of their booklet they had to design an image on environmentalism and they designed the image below which illustrates the impact of not recycling with a red 'NOW!' juxtaposed against the green 'RECYCLE' adjacent to some recycling bins. Their writing would appear to be a transmodal shift from the visual to text and will be further discussed in section 5.5.

Figure 5-9 (HLA 2.0 - create a protest image for page 1 of booklet; Appendix B)

At 02:40 Sakia completed the writing of sentence one and Gamda said ‘mhm’ to agree with the deletion of the comma in the sentence. Gamda then immediately drops her head into her hand and offers a potential beginning for the next sentence at 02:42.

Transcript 5-7 (02:42 – 03:18)

02:42 Gamda	but (1) we can write it as well (1) like
02.45 Sakia	mhm
02:50 Gamda	it like (1) only takes (1) a few seconds to sort them out separately
02.56 Sakia	aha
02:59 Gamda	and
03:14 Sakia	s e pa
03:15 Gamda	ra sepa <separately>
03:18 Sakia	separately nie wiem {I don’t know} ºit takes only few seconds to put the litters separatelyº {reads from the screen}

In the opening few seconds of sentence two there are multiple and complex cross-modal configurations in each learner’s communication with the other. To best describe this multimodal interaction, the above language exchange will be: (1.) deconstructed at modal level in the extract below and discussed; (2.) then visually through images of the learners to understand how gesture combines with talk.

Extract 5-10 (02.42 – 03:00)

Ref:		Seconds:	02:42	02:44	02:46	02:50	02:52	02:54	02:56	02:58	03:00
109.	Sakia	talk			mhm				aha		
110.		actions									
111.		keyboard			⌨{paragraph break}					I t	
112.		screen	⇩⇩⇩⇩⇩⇩							I t	
113.		mouse	🖱---------								
114.		gaze	💻----------			💻🗣----------			🗣💻----------		💻
Ref:		Seconds:	02:42	02:44	02:46	02:50					
115.	Gamda	talk	but we can write it as well like			it like only takes a few seconds to sort them out separately					and
116.		actions	head in hand		chin in hand	opens hand			waves hand side to side		
117.		keyboard									
118.		screen									
119.		mouse									
120.		gaze	😐----------		😐❑	----------					❑

At Ref: 115(02:42) Gamda starts the next sentence, offering a suggestion rather than dictating; and there is evidence for this being suggestive through her repetition of the word ‘like’ at 02:44 and at 02:50. Previously at 00:57, sentence one, Sakia started a sentence similarly when she said ‘like we should more care about recycling’. ‘Like’ can be used as a ‘filler’ in conversation but is perhaps here functioning in each instance as a preposition to suggest similarity. The words are an offer, a starting point. This lack of assertion promotes linguistic equality in their text production; it is evident throughout their language collaboration off-screen and editing when on-screen. In CA studies the language function is similar to ‘negative politeness strategies’ (Brown & Levinson, 1987) in which speakers avoid giving offence by showing deference to the other. Here it functions as respect for the potential language choices of the other.

At the exact same moment that Gamda offers a sentence at 02:42, Sakia is already preparing on-screen where to start writing. At Ref: 112 & 113(02:42), she moves the cursor down the screen and clicks after the full-stop at '*clean and healthy*'. At Ref: 111(02:45) she presses the enter key on the keyboard to create a paragraph break to start the next sentence on a new line; Figure 5-10.

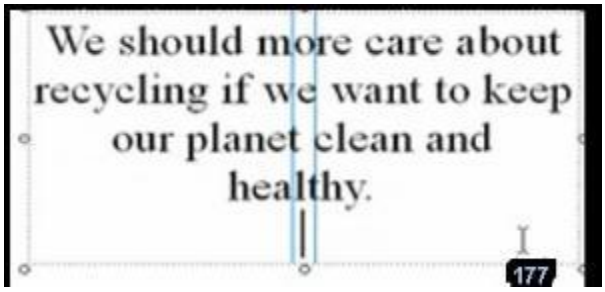


Figure 5-10 (start point of second sentence)

Simultaneously, as Sakia is preparing the text area, she acknowledges Gamda with a back-channelling 'mhm' at Ref: 109(02:45). Sakia then turns to Gamda at Ref: 114(02:50) as the latter gestures with an open hand and says: 'it like (1) only takes (1) a few seconds to sort them out separately' - Ref: 116 & 117(02:50). Sakia responds with a confirmatory 'aha' at 02:56 and immediately starts typing the first word 'it'. They have the start of their next sentence.

The images below will help to illustrate how gesture combines with talk in the same eighteen second sequence of interaction. Sakia's talk is in red and Gamda's talk is in black.

Figure 5-11 (gesture combined with talk)

As a brief reminder there are broadly four types of gesture:

- Deictic - pointing to objects or people often in synchrony with language: 'this' 'that' 'him' 'her'; from the Greek deiktos meaning 'able to show directly'
- Iconic – using the hands to visualise a concrete object such as writing a word on a table with a finger to try and spell it
- Metaphoric – visualising the abstract such as tapping on one's head to suggest 'craziness' in another
- Beat – (in/out) or (up/down) movements such as a tapping a foot to music.

At 02:44, Gamda drops her head into her hand. This might be considered a metaphoric gesture to externalise the process of thinking about what the next sentence might be. She speaks downwards at the table 'but we can write it as well like' before lifting her head up to look at the screen (02:50) and says 'it like (1) only takes (1) a few seconds'. As she speaks she uses another gesture which is an opening and closing hand, the fingers in an upwards claw shape. This too is possibly a metaphoric gesture for her mind trying to grab the words she is looking for. When she suggests the main clause in the second half of her sentence ('sort them out separately') at 02:56 her hand moves side to side in what appears to be a beat gesture. However, it is more likely functioning as an iconic gesture to visualise the sorting out of household waste into different recyclable categories. From 03:00 Sakia types quickly '*It takes only few seconds to put the litters sepa*' and then she pauses.

03:14	Sakia	s e pa
03:15	Gamda	ra sepa <separately>
03:18	Sakia	separately

Sakia is uncertain how to spell the word so she slows down her speech, at which point Gamda takes over to help and they 'work' the syllables together to try and figure out the spelling. The same strategy was seen previously when Sakia was typing 'environmentalism' into a search-box, sounding out the syllables at 00:13 which were then taken over by Gamda at 00:15.

00:01 Sakia	<°environmentalism°> (2) zaraz sprawdzimy	{we're going to check}
00:07 Gamda	mhm↑	
00:13 Sakia	en: vi:	
00:15 Gamda	ron >mental<	

Spoken language is slowed down to help with the on-screen spelling and this is a recurring feature of temporal mapping which will be explored further in the discussion chapter as a significant finding.

Upon completion of typing 'separately', Sakia then looks briefly at Gamda at 03:18 and says 'nie wiem' {I don't know} followed at 03:19 by her leaning towards the screen and reading the sentence back to herself.

Figure 5-12 (Sakia uncertain if their text is correct)

What is lost and gained so far in the shift from talk to the writing of the first sentence?

- Gamda's spoken words: it only takes a few seconds to sort them out separately
- Sakia's written text: *it takes only few seconds to put the litters separately*

‘It’ clauses (sometimes referred to as cleft sentences) are the most common type of cleft clause. Whilst a cleft sentence can appear in a simple sentence, they frequently appear in complex sentences where there is a main clause and a dependent clause. It, as a cleft clause to start a sentence, commonly contains information that is already known. Clefts can put emphasis on the main clause and make for persuasive language. In Gamda’s sentence, ‘it’ opens the cleft sentence as a dependent clause and is a cataphoric reference to the main clause in the second half of the sentence; ‘it’ can only be understood by reading on to understand that ‘it’ refers to the sorting of household rubbish for recycling: ‘sort them out’. However, Gamda has used the third person plural pronoun ‘them’ for the referent ‘rubbish’ which makes the sentence more difficult to understand because the reader/listener has to determine from the context of recycling what ‘it’ and ‘them’ (as indirect objects) actually refers to; in this case one has to go backwards anaphorically to the previously written sentence ‘*we should more care about recycling*’ to understand the new sentence.

Sakia is more precise in her language use, substituting ‘them’ for ‘*litters*’, so that the opening cleft ‘it’ can be understood cataphorically to the main clause of sorting litter for recycling. Though ‘*litters*’ is not ideal as she has incorrectly used an uncountable noun and pluralised it, we can still determine from the context what is meant. Accuracy of noun use is one of the constraints of writing in a second language. The English language can have multiple synonyms and sometimes there is not an obvious word-for-word translation. In English there is general acceptance that the most appropriate word is ‘rubbish’ when discussing the recycling of household waste. The most direct comparison for ‘rubbish’ in Polish is ‘śmieci’ which translates as garbage, trash, litter. ‘Litter’ is the word Sakia has chosen to use. ‘Litter’ in the UK is generally understood to refer to the problem of empty food packaging, tins, bottles, etc (or rubbish) in public areas. Again, it is not perfect but close enough to be understood. This problem of the most appropriate noun was seen earlier when the learners wanted to understand the word ‘environmentalism;’ finally settling on ekologia and ekolog.

Sakia drops the indefinite article ‘a’ before the phrase ‘only few seconds’. Articles are not present in the Polish language; though Gamda uses it in speech Sakia fails to keep it in the text. It is a recurring feature in Polish learners that they can self-correct and add articles to spoken English, even though articles are not used in Polish, but then when writing can revert back to learned forms in the first language. Gamda’s ‘only’ functions

as an adverb in ‘it only takes’ but Sakia changes the placement of ‘only’ so that it becomes an adjective in ‘only few seconds’. In both instances ‘only’ functions as an intensifier for a ‘few seconds’; evidence of the learners again using persuasive language features; here stressing the shortness of time required to recycle: ‘only a few seconds’. Finally, Sakia changes Gamda’s verb phrase ‘sort them out’ with the imperative verb ‘put’ which has more immediacy and impact.

In the process of writing seen so far in sentence two, do the learners realise the extent to which Gamda’s oral sentence has changed following Sakia’s transcription? Has Sakia deliberately edited the words of Gamda? If so, in grammatical terms, Gamda’s is more accurate than Sakia’s. Or is it simply that spoken words are easily forgotten and Sakia is filling the gaps in memory with her own words? The words in red show which learners contributed which words to the final sentence, as it later appears on-screen.

- Gamda’s spoken words: it only takes a few seconds to sort them out separately
- Sakia’s written text: *it takes only few seconds to put the litters separately*

5.6 completing sentence 2

In section 5.2 it was shown how the learners arrived verbally at the construction of their first sentence: ‘We should more care about recycling if we want to keep our planet nice and clean.’ In 5.3 they made the transmodal shift from talk to text and wrote: ‘*We should more care about recycling if we want to keep our planet clean and healthy.*’ The second sentence on completion reads as: ‘*It takes only few minutes to put the litters in order and prevent the Earth from pollution.*’ Section 5.4 examined the first half of their second sentence: ‘*It takes only few seconds to put the litters separately*’ and considered how fluid the shift from talk to writing is; in terms of being slowed down for comprehension and fragmented between speakers as it appears on-screen. Section 5.5 will examine the production of the second half of their sentence: ‘*and prevent the Earth from pollution.*’

At 03:29 Sakia types what is eventually an abandoned continuation of their second sentence – ‘*- it takes years*’ so that the sentence now reads as: ‘*It takes only few seconds to put the litters separately – it takes years.’* She is perhaps trying to juxtapose two timescales: the ‘few seconds’ it takes to recycle against the ‘years’ it takes for landfill to

erode. This juxtaposed concept makes sense in relation to the image they previously designed.

We should more care about recycling if we want to keep our planet clean and healthy.

It takes only few minutes to put the litters in order and prevent the Earth from pollution.

Figure 5-13 (the visual recycle compared to the writing of recycle)

The ‘narrative’ of the visual reads from top to bottom: the positive imagery and green colour of recycling down to the negative images and red colour of not recycling. It also reads from right to left; the waste on the right being bulldozed into the green earth on the left. The ‘narrative’ of the textual is a transmodal shift from the visual design to their sentences and reads from left to right. The first sentence at Figure 5-13 corresponds broadly to the top and left side of the image with the language of ‘recycling’ and ‘clean and healthy’ transposed from the green ‘RECYCLE’ with the bins and the green earth and blue water. The second sentence at Figure 5-13 corresponds broadly to the bottom and right side of the image with the language of ‘litter’ and ‘pollution’ and ‘Earth’ transposed from the images of the bulldozer filling waste into the earth and the warning red of ‘NOW!’ An affordance of the visual mode to the learners’ design is one of spatial logic. It can be read in any direction. The transmodal shift of concepts from the visual to text brings with it the constraint of linearity and second-language grammar. The text can only be read in one direction: left to right; and this constraint begins to be felt as they struggle to compose the second part of their sentence. They become stuck at ‘- *it takes years*’.

The next four minutes of their interaction will be summarised as nothing new is written. At 03:47 Sakia says ‘rozklad’ {decay} and opens the internet and types rozklad into the translation website ling.pl. Gamda opens her pocket translator. At 04:05 Sakia types ‘degradacja’ into ling.pl which translates as 'degradation' and 'deterioration.’ At 04:30

Gamda reads aloud from her translator: ‘peace, resolution, corruption’. Between these small episodes of talk and activity there are longer passages of silence and inactivity with frequent dropping of heads into hands as they think.

04:49 05:47

Figure 5-14 (thinking what to write)

At 06:13 Gamda offers a suggestion to continue the sentence when she says: ‘It takes years to neutralise the’ but Sakia replies with ‘nie wiem’ {I don’t know}. Sakia tries Google, searching the word ‘recycling’ and clicks on random websites but does not find anything. At 07:41, Sakia clicks on a Google link to Wikipedia and she appears to see something. From 03:29 to 07:42, over four minutes, they have not produced a new word to continue their sentence. At 07:43, Sakia reads aloud from the Wikipedia page: ‘to prevent … in order to prevent.’ Gamda responds at 07:46, echoing her: ‘to prevent … to prevent from what.’ This would appear to be the word they have been looking for to continue their next sentence as at 07:54 Sakia deletes the following words ‘*It takes only few ~~seconds to put the litters separately – it takes years~~*’ and begins to retype the sentence. The final piece of language for analysis is in the transcript below and the following extracts. The analysis continues from this point at 07:56 when Sakia starts typing again.

Transcript 5-8 (07:56 – 08:49)

07:56 Sakia	it takes only few minutes	
07:58 Gamda	to put the litters in order	
08:08 Sakia	and [prevent	
08:08 Gamda	[prevent	
08:10 Sakia	the earth from degradation (08) from what↑	
08:20 Gamda	from (1) zaśmiecenie	{littering}
08:27 Sakia	from pollution (08) jeden odcinek	{that’s one part}

08:42 Sakia	zaśmiecenie to ok a degradacja	{littering is ok but what about degradation}
08:43 Gamda	mhm	
08:44 Sakia	degradacja	{degradation}
08:49 Gamda	myśle że to (pollution)	{I think that's it}

They speak in English as they talk and write (07:56 – 08:10). However, as they begin to think about what else they might write, and they start to use the Polish website ling.pl in conjunction with Wikipedia and Google search, they become more reliant on Polish to communicate. The first-language of Polish becomes a meta-language for discussing the second-language of English. The learners also use the internet to help them find collocates for their sentence. It proves to be an effective language learning strategy; if a little bit slow. The following extracts will detail the above transcript to think about how talk and writing are sequenced and sustained through a series of cognitive batons, linguistic scaffolding, echo utterance, translanguaging and creative 'borrowing' from the internet as corpus.

Extract 5-11 (07:55 – 08:10)

Ref:		Seconds:	07:55	07:56	07:57	07:58	08:00	08:04	08:08	08:09	08:10
121.	Sakia	talk		it takes only few minutes					prevent		the earth
122.		actions									
123.		keyboard	minutes				to pre	put the litters in order and		prevent	
124.		screen	minutes				to ~~pre~~	put the litters in order and		prevent	
125.		mouse									
126.		gaze	💻----------								----------💻
Ref:		Seconds:	07:55	07:56	07:57	07:58	08:00	08:04	08:08	08:09	08:10
127.	Gamda	talk				to put the litters in order			prevent		
128.		actions									
129.		keyboard									
130.		screen									
131.		mouse									
132.		gaze	❑--------❑🗔----------			----------🗔❑			----------		----------❑

At Ref: 123(07:55) Sakia begins to retype the sentence, deleting 'seconds' and adding 'minutes' so that it now reads as: '*It takes only few ~~seconds~~ minutes'.* At Ref: 121(07:56) Sakia reads back the sentence so far as it has been rewritten. At the same time as Sakia reads 'it takes only few minutes', Gamda looks briefly at her electronic translator [Ref: 132(07:56)] then lifts her head to offer a continuation of the sentence at Ref: 127(07:58) when she says: 'to put the litters in order'. Sakia immediately types verbatim on-screen what Gamda has said at Ref: 123(07:59). She then continues the sentence by typing '*and*' though she does not say it. Unusually, both learners simultaneously say 'prevent' at 08:08. It is likely that they both understand the next logical piece of language because at 07:43 Sakia read from the Wikipedia page: 'to prevent … in order to prevent' and Gamda repeated her at 07:46. This is the likely catalyst for the start of rewriting the sentence only twelve seconds later at 07:55. The word appears in their sentence twenty-five seconds later at 08:08 and both learners say the word at precisely the same time as the sentence is written in real-time. Sakia continues the sentence at Ref: 121(08:10) when she says: 'the earth' and at 08:11 'from degradation;' though she does not write the word degradation, only '*the Earth from*.'

Extract 5-12 (08:11 – 08:28)

Ref:		Seconds:	08:11	08:12	08:13	08:18	08:20	08:24	08:26	08:28
133.	Sakia	talk	from degradation			from what↑				from pollution
134.		actions	rubs her neck					internet page opens		
135.		keyboard			the Earth from					
136.		screen			the Earth from		↖↖↖opens Wikipedia↖↖↖	↖↖↖↖↖↖↖↖↖↖↖↖↖		
137.		mouse								
138.		gaze	🗣----------		🗣💻		----------			💻
Ref:		Seconds:								
139.	Gamda	talk					from zaśmiecenie {littering}			
140.		actions	nods her head				looks over her right shoulder			
141.		keyboard								
142.		screen								
143.		mouse								
144.		gaze	🗣----------		🗣❑	----------	❑🗣----------		🗣🗔	---------- 🗔

In Extract 5-12, Ref: 133(08:11), as Sakia says ‘from degradation’ she turns to Gamda and rubs her neck. Gamda mirrors Sakia by turning to her and nods in agreement at Ref: 140(08:11). Three seconds later at Ref: 135(08:13) Sakia types the words ‘*the Earth from*’ but would appear to be uncertain as at 08:18 she asks ‘from what↑’. Gamda replies with ‘from (1) zaśmiecenie’ {littering} at Ref: 139(08:20) but then looks at her electronic translator at Ref: 144(08:21). As Gamda does this Sakia opens the Wikipedia page and the cursor hovers near a sentence on the page: ‘reduce air pollution … and water pollution’. At this point 08:27 Sakia says ‘from pollution’, suggesting ‘pollution’ from Wikipedia to continue their sentence ‘*the Earth from*’

A noticeable feature in their talk above is the echo utterance of the preposition ‘from’, which helps to sustain the sequencing of their interaction and sentence development.

08:10 Sakia	the earth from degradation (08) from what↑	
08:20 Gamda	from (1) zaśmiecenie	{littering}
08:27 Sakia	from pollution	

The analysis continues from 08:35 in Extract 5-13.

Extract 5-13 (08:35 – 08:49)

Ref:		Seconds:	08:35	08:42	08:43	08:44	08:47	08:48	08:49
145.	Sakia	talk	jeden odcinek	zaśmiecenie to ok a degradacja		degradacja			
146.		actions	opens www.ling.pl					Publisher opens	
147.		keyboard	zaśmiecenie						
148.		screen	zaśmiecenie	translation appears on-screen as ‘littering’					
149.		mouse	⇳ scrolls down the translations⇳	⇖⇖clicks on Publisher⇖⇖⇖⇖⇖			⇖⇖⇖clicks on Wikipedia⇖⇖⇖		
150.		gaze	💻----------						----------💻
Ref:		Seconds:							
151.	Gamda	talk			mhm				myśle że to (pollution)
152.		actions							☜ points at the screen
153.		keyboard							
154.		screen							
155.		mouse							
156.		gaze	❑----------						----------❑

As Gamda looks up from her translator at 08:35 Sakia says ‘jeden odcinek’ {that’s one part}; i.e. ‘pollution’ is the word they need. Perhaps ‘one part’ refers to the possibility that Sakia is looking for the sentence to have two words joined by a conjunction; for example, ‘pollution and littering’ or ‘pollution and degradation’. At 08:42 Sakia confirms Gamda’s suggestion of ‘zaśmiecenie’ {littering} but perhaps thinks her suggestion ‘degradacja’ {degradation} might be the better choice. At Ref: 145(08:42) Sakia says: ‘zaśmiecenie to ok a degradacja’ {littering is ok but what about degradation}.

She would appear to be wrestling with the choice of each. This could evidence linguistic equality between the two, as has been seen throughout their text construction, as both appear to recognise and value the language contributions of each other. As Publisher opens at 08:48 Gamda becomes animated and speaks quickly in Polish as she points at the screen in four places.

Figure 5-15 (choosing pollution)

At Ref: 151(08:49) Gamda says: ‘myśle że to (pollution)’ {I think that’s it}. She points at the space at the bottom of the screen near ‘*Earth from*’ and then at the top of the screen when Sakia switches to Wikipedia, approximately where the word ‘pollution’ is, then back to the bottom of the screen 08:51 and up again, 08:52. These deictic gestures signal for Sakia which word Gamda wants transposing from Wikipedia to their sentence. The deictic gesture of Gamda to Sakia suggests a spatial relocation of the word from one place on-screen to a different place on-screen. At 08:58 Sakia pauses the mouse over the word ‘pollution’ in Wikipedia; Figure 5-16. They say something quickly and unclear in Polish and at 09:02 Sakia returns to Publisher and completes their sentence by typing ‘*pollution*’.

Figure 5-16 (pollution in Wikipedia)

To summarise, the temporal mapping in sentence two is sequenced by an ongoing chunking of the sentence between them, where each language chunk becomes a cognitive baton for the listener to support linguistic scaffolding. There are examples of the listener taking the tail-end of the previous speaker and continuing with the next logical language chunk, and vice-versa as in Table 5-1.

Sakia	Gamda	Sakia	Both	Gamda
it takes only few minutes 07:56	to put the litters in order 07:58	and 08:08	prevent 08:08	the earth from degradation 08:10

Table 5-1 (sentence 2)

Talk is slowed down to allow for the cognitive processing of the sentence structure when spoken and is also slowed down to allow for the temporal mapping of off-screen talk to on-screen text. They both contribute and even use some words from the internet, substituting ‘degradation’ for ‘pollution.’ On-screen text then can be a polyvocal

construction. To explore the polyvocality of temporal mapping in real-time, the waveforms below show three timescales, from 07:56 to 08:02 at the top, which continues on the middle timeline from 08:03 to 08:11, and on the bottom from 08:11.05 to 08:16. The spoken words of Sakia are in red, Gamda's in black and typing on-screen is white font on black background. What the waveforms evidence is the asynchronous and synchronous language production in temporal talk: (1.) the temporal mapping of talk to text and the time-lag between the two; (2.) the gaps of 'thinking' silence between the two learners as they take turns to talk; (3.) where talk and text are synchronised; (4.) the amplification and duration of sounds at syllable level.

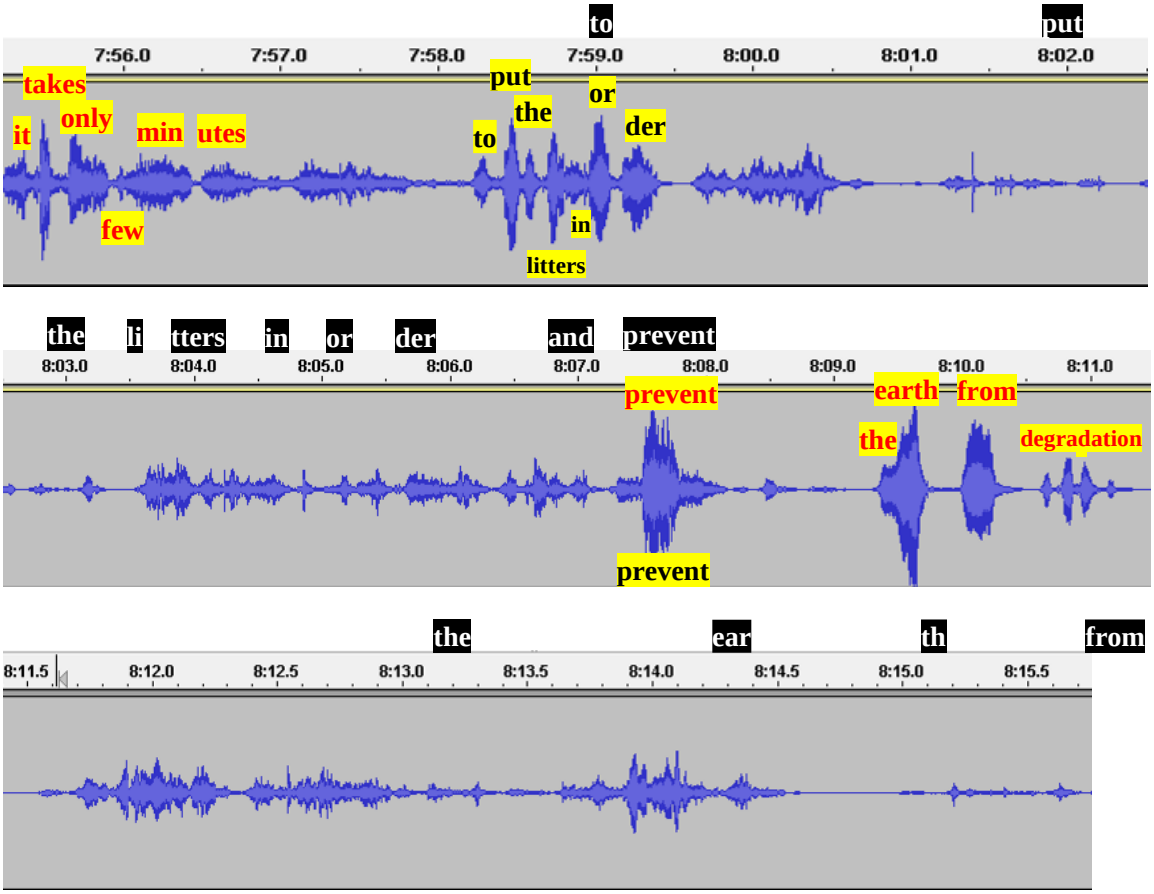


Figure 5-17 (Sakia, Gamda, typing)

At 07:55.5 Sakia reads aloud what she has previously typed 'it takes only few minutes.' The relatively low peaks and valleys of the amplification for each syllable (compared to the amplification of other words) suggest that Sakia is reading to herself rather than speaking to be heard. It takes approximately one second for her to read this language chunk aloud and there is a two second pause before Gamda continues the sentence with 'to put the litters in order' which again takes only a second to verbalise.

At 07:59, one second after Gamda begins her language chunk of 'to put the litters', Sakia begins to type this on the screen; synchronising her listening with typing. A typing error slows her down at 08:01 when she types '*pre*' and then deletes it. Perhaps she was beginning to type the word 'prevent' but corrects herself to type verbatim what Gamda has said. From 08:02 to 08:07 both learners are silent to allow the text to be typed and the low peaks on the timeline evidence only the sound of typing. At 08:07 Sakia types '*and*' but does not speak it and at 08:08 there is synchronisation between the two learners and typing as the word 'prevent' appears on-screen and can be heard spoken by both. The high peak of amplification on this word evidences how loudly it was spoken and perhaps the importance of the word to them as the missing piece in the language sequence they spent four minutes trying to solve, from 03:29 to 07:42. At 08:09.5 Sakia continues the sentence by saying: 'the earth from degradation'.

There is a noticeable gap of four seconds before the words appears on the screen from 08:13.5 to 08:16. The reason for this gap is because Sakia stops typing and turns to Gamda for confirmation. Gamda nods in agreement and Sakia continues to type, but stops before she writes degradation; seemingly uncertain about this word. The fact that Sakia pauses and turns to Gamda is further evidence of linguistic equality; how both learners contribute to, and sustain, an inclusionary participation framework. The language contributions and interactions of each are valued and equally valid. At 09:02 Sakia completes the sentence with the word 'pollution' taken from Wikipedia. Temporal mapping between both learners is sequenced and sustained through a series of cognitive batons, including linguistic scaffolding, echo utterance, polyvocal inter-languaging and creative 'borrowing' from the internet as corpus.

5.7 conclusion

A conversational analysis approach can tell us much about the sequentiality of the learners' temporal mapping and interaction. How they use:

- echo utterance to enable cognitive batons,
- standard turn-taking,
- adjacency pairs of suggestion: confirmation
- linguistic equality through the recognition of each other's contributions,
- back-channelling techniques,
- humour,
- linguistic scaffolding,
- silence.

However, as discussed throughout the sections, there are a number of cross-modal configurations in their interaction alongside language which positively structure their interaction and help them to achieve their aim. There is a temporal ordering of actions and utilisation of different modes to complete the task. All these features of language and interaction contribute to, and sustain, an inclusionary participation framework.

In the next chapter with the Kurdish learners we find *competitive interaction* rather than *collaborative interaction*. Temporal mapping is co-constructed between the learners but rather than *collaborative co-construction* of meaning, as with the Polish learners in chapter 5, there is *competitive co-construction*; evidenced by an ongoing 'battle' for tools and spatial control. The result is a contesting of participation with some features of interaction different to that of chapter 5.

6 Analysis 2

6.1 introduction

This chapter describes the interaction between Darras and Shourok and the resources and modes they utilise to help them complete writing for the fourth page of their booklet, as was seen with the Polish learners in chapter 5. Combining multimodal and conversation analysis, the microanalysis will focus on the cross-modal configurations in interaction and the relevance of the different modes to the collaborative design process of generating on-screen text from off-screen talk. Darras and Shourok collaborate to write three sentences for page four of their booklet on the environment. They are using Microsoft Publisher. There is an identifiable beginning (a blank page on the computer) and an identifiable end (when the writing is completed). Within the overall classroom sequence (Appendix C) this focal event is higher-level action 1.7 within the larger first action sequence of creating a booklet about Greenpeace and environmentalism. The video of their interaction is available on the accompanying CD called **Kurdish_1.7** and a close-up video of their writing is called **Kurdish_writing**. This focal point has been chosen because it provides a unique insight into the real-time design process where language, literacy and technology come together. The equivalent focal event was seen in chapter 5 with the collaboration of the Polish learners.

Darras and Shourok have been working on the first three pages of their booklet for one hour. At the beginning of the interaction for analysis the learners write a title for the fourth page called 'My thoughts' and insert a recycling image they found on the internet. They then add a textbox under the image in which to write their text. The completed text is three sentences long and when finished reads as: *'I think we can always help our Earth at any time when we wake up, knowing how to use electricity, water, gas, petrol and many other things around us. We should be responsible for recycling our bins. Be responsible for our streets, our cities., then we can say we took a place in Greenpeace;'* see Figure 6-1 overleaf. All on-screen text typed by the learners will be presented in the analysis using *italics*.

Figure 6-1 (page four of recycling booklet and their previous image design)

The totality of language for analysis between sections 6.2 and 6.5 is as follows. The transcription conventions are at Appendix A.

00:01	Darras	page 4
00:02	Shourok	°add a text box [write a few sentences about Greenpeace or environmentalism what do you think°
00:06	Darras	[page 4
01:09	Darras	make a space
01:13	Shourok	°centre°
01:21	Darras	you can make it best fit (5) but make make space
01:29	Shourok	space for what↑
01:30	Darras	my thoughts space
01:38	Shourok	بگیر عینک ! نمیبینی؟. بهتره ان با . {It is better with the other one. Can't you see? Get glasses!}
01:40	Darras	haha (2) haha (2) nice
03:12	Darras	what do you mean↑ getting show
03:14	Shourok	shower shower
03:16	Darras	shower haha
03:17	Shourok	shower it's when the ()
03:25	Darras	take (1) take shower
03:27	Shourok	بگیری! {take a shower}
03:55	Shourok	()
03:56	Darras	elec
03:57	Shourok	lec(h)tri
03:58	Darras	e:lec: tri (3) r
04:01	Shourok	() (14) I can't do it
04:26	Darras	haha
04:41	Darras	last one you should put and here
04:44	Shourok	no there's new gas (10) <petrol>
05:05	Darras	many thinKS <things> thing things

05:10	Darras	and many other things around us (1) <many other> (3) things (2) thing
05:22	Shourok	ha
05:30	Darras	things around us
06:33	Darras	err finished you need to make bigger
06:36	Shourok	دیگه؟ {finished what else} ننویسیم؟ بنویسیم؟ بیشتر still {shouldn't we still continue writing}
06:42	Darras	I think that (8) we we should be responsible for recycle our bins
07:04	Shourok	for
07:06	Darras	recycling recycling our bins
07:08	Shourok	((*grunt*))
07:15	Darras	it's for recycling (5) our <bins> <recycling> <recycling> <bins>
07:40	Darras	if you change this to recycle
07:44	Shourok	((*grunt*))
07:45	Darras	recycle our bins

Transcript 6-1 (full transcript)

The following sections will take extracts from the total language exchange above and offer multimodal transcription to analyse what else is happening in addition to talk. Section 6.2 (00:01 – 00:16) will look at how the learners orientate themselves to the task of writing about environmentalism; there is no negotiation of who will do what and there is no planning or drafting of the sentences. Shourok dominates the computer space and Darras, though peripheral, tries to engage in the design process. A central question is who is empowered and disempowered by the tools of collaboration: keyboard, mouse and worksheets? Section 6.3 (01:09 – 01:40) examines their collaboration when writing just a two word heading called '*My thoughts*.' Close multimodal analysis reveals in this instance what is happening in addition to language as learners 'compete' for space and tools in collaboration. There is very little modal, proxemic and linguistic alignment between the learners.

Section 6.4 (03:12 – 05:30) is a lengthy section as the learners construct their first sentence. From a conversation analysis approach, this section reveals that though there is conversational grounding and conventional turn-taking synchronised with gesture and gaze, there is predominantly a lack of mutual coherence about what to write and for Darras this destabilises the peer-interaction framework he is a part of with Shourok. Section 6.5 (06:33 – 07:45) examines the construction of the second sentence and considers in particular the transmodal shift of talk to text and the sequential temporal mapping of off-screen talk to on-screen text; from one learner to another. How language and text is mediated by this process is the focus of section 6.5. Section 6.6 examines Darras's contesting of the design process and Shourok's reluctance to collaborate when

Darras has control of the computer, resulting in a breakdown of the participation framework. Particular attention will be paid to the cross-modal configurations in their interaction and the relevance of these modes to the participation framework.

6.2 understanding the task

In Extract 6-1 the analysis focuses on the learners establishing what they need to do for the next activity, which is to write some sentences about environmentalism for the fourth page of their booklet. Their combined verbal output during this extract is as follows:

Transcript 6-2 (00:01 – 00:06)

00:01 Darras	page 4	
00:02 Shourok	°add a text box	[write a few sentences about Greenpeace or environmentalism what do you think°
00:06 Darras		[page 4

Upon first listening and watching there does not appear to be much of interest in these opening sixteen seconds. Close multimodal analysis however reveals much about the proxemics of the learners and how one learner is constrained by the spatial arrangement of the objects around him, as well as by the learner he is collaborating with. Talk does not appear to contribute much in this interaction. This raises the question of what other communicative modes are present.

Extract 6-1 (00:00 – 00:16)

Ref:		Seconds:	0s	2s	4s	6s	8s	10s	12s	14s	16s
1.	Darras	talk	page 4		page 4						
2.		actions	taps on 📖		slaps his knees					leans to 📖	bangs the table
3.		keyboard									
4.		screen				⇖----------	----------	----------	----------⇖		
5.		mouse				🖱----------	----------	----------	----------🖱		
6.		gaze	📖----------	----------	📖❑----------	----------	----------	----------	----------❑	📖--------	📖❑----------❑
Ref:		Seconds:	0s	2s	4s	6s	8s	10s	12s	14s	16s
7.	Shourok	talk		ºadd a text box write a few sentences about Greenpeace or environmentalism what do you thinkº							
8.		actions		he rocks backwards and forwards as he quietly reads							
9.		keyboard									
10.		screen									
11.		mouse									🖱----------🖱
12.		gaze		📖----------	----------	----------	----------	----------	----------	----------📖	❑----------❑

At Ref: 1(0s) Darras is positioned to the right of the computer and stretches across the keyboard and across Shourok. His left hand taps on the page of the worksheets Ref: 2(1s) to indicate to Shourok what the next step is as he simultaneously says ‘page 4.’ Shourok responds to the gesture and verbalisation by looking at the worksheets at Ref: 12(1s). At Ref: 7(2s) he reads aloud the instructions quietly to himself. At the same time Shourok rocks backwards and forwards on his chair as he reads. Darras repeats ‘page 4’ at Ref: 1(4s) to reiterate the next stage of the task and at the same time slaps his knees, Ref: 2(4s). At Ref: 6(4s) Darras’s gaze shifts to the screen and at Refs: 4-5(6s) he holds the mouse and navigates to the fourth page of their booklet to prepare them for the on-screen element of the next task. He clicks on page four of their booklet. Shourok continues to read the instructions. At Ref: 2 and 6(13s) Darras again leans across the space between him and the worksheets to read the instructions.

However, at Ref: 11(16s) Shourok reaches for the mouse, breaking into the space Darras had leant into to read the instructions, forcing Darras back into a sitting position, and at the same time, Ref: 2(16s), Darras bangs his fist lightly on the table.

Figure 6-2 (00:01 – 00:16)

In Figure 6-2 (00:01) we see Darras lean forwards to tap the student worksheets in front of Shourok and say 'page 4.' Shourok's talk is in red font and Darras's is black. Darras uses this deictic gesture of tapping on the page, in synchrony with language to instruct Shourok on what they need to do next. This multimodal episode of communication, from Darras to Shourok, evidences his cooperation in the collaborative task. At 00:06 seconds, Darras shows his ongoing collaboration in the task by holding the mouse and moving to the next stage of the task; page 4 to do their writing. His gaze is fixed on the screen. The gaze of Shourok is fixed on the page as he reads aloud the instructions. Darras is ahead of Shourok in that he knows what they need to do next and has arranged the on-screen mouse position and page in the Publisher program to enable them to continue. In terms of meaning making, Darras's off-screen and on-screen coordination would seem to suggest he is confident in understanding the task and 'leads' them by telling Shourok what to do next and restructuring the immediate tools around them to

begin. In the meantime Shourok continues to read the instructions. At 00:14, however, the dynamic changes. As Darras leans across Shourok to read the same page, perhaps mirroring Shourok, the hand of Shourok pushes across the space in front of Darras and he holds the mouse. This movement pushes Darras backwards away from the keyboard and the mouse and at 00:16 Shourok looks at the screen and begins the task. Simultaneously, Darras lightly bangs his fist on the table.

In Extract 6-1 and Figure 6-2 there are two communicative utterances ('page 4') and from only one person: Darras. Shourok's verbalisation is for him alone when he reads aloud. Darras also uses gesture to communicate to Shourok. Darras also uses the mouse to help them in the task and prepare the ground. Conversation analysis provides a framework for exploring 'the rule that a current action (…) requires the production of a reciprocal action' (Goodwin & Heritage 1990: 287). Darras speaks twice to Shourok and gestures for Shourok but none of these communicative instances are verbally reciprocated or gesturally acknowledged. Shourok continues to read quietly to himself. Given there is so little language use, is it possible to interpret meaning from other modes such as gesture, proxemics and spatial arrangement? As a brief reminder there are broadly four types of gesture: deictic, iconic, metaphoric and beat.

The deictic gesture of pointing as Darras says 'page 4' is functional and part of the communicative landscape. The second time he says 'page 4' he simultaneously slaps his thighs as a beat gesture; perhaps to show his readiness to start the task. The final gesture Darras uses in Extract 6-1 is another beat gesture as his fist pumps up and down. The gesture is perhaps an expression of frustration for Darras. What evidence is there for this?

Figure 6-3 (00:38 – 00:42)

Figure 6-3 is taken from 23:21 seconds into the beginning of their collaboration at the start of the lesson. At 00:38 Shourok looks to his left to find his worksheets. When he sees they are not there he reaches across at 00:40 and 'takes' the worksheets from Darras. He does not speak, does not ask, just takes them and for the rest of their interaction (over 2 hours) keeps the worksheets in his hand or to the left of him when he is using the keyboard. Four seconds later (at 00:42) there is the same light bang of the fist by Darras on the table, which would seem to be causally linked to Shourok's taking of a 'tool' – the worksheets. We see the same action repeated later under similar circumstances when Shourok takes a different tool from Darras – the mouse.

Proxemics is the distance people adopt in relation to other people and objects. The discussion chapter will detail the findings of proxemic behaviour in the two analysis chapters. For now, at a broad level, there are four basic types of distance adopted between people: intimate, personal, social and public distance (Hall, 1966); though these can vary dependent on context and between genders. Proxemic behaviour in relation to objects, such as a computer, keyboard, mouse, need to be considered as the proxemic distances people adopt can be controlled by the spatial arrangement of such objects. Pairing individuals at a computer is commonly understood to have a triangular placement as below.

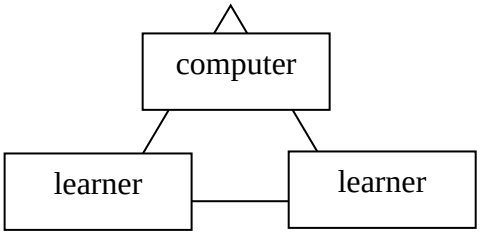


Figure 6-4 (triangular layout)

The distance between the learners and the tools (keyboard, screen, monitor, worksheets) requires a physical closeness which the two male adults might otherwise avoid. The layout suggests spatial equality in proximity to the screen and tools. The proxemics of their interaction (with each other and with the computer) is mediated by a triangular layout; they both face the computer, shoulder-to-shoulder rather than face-to-face. However, the video of their interaction shows spatial inequality. A truer representation would look like this:

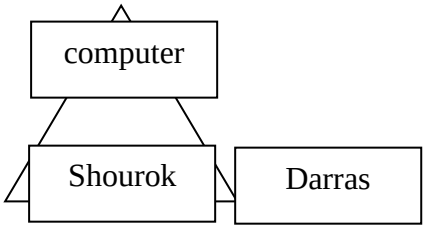


Figure 6-5 (triangular layout 2)

In terms of meaning making, Darras's off-screen and on-screen coordination would seem to suggest he is confident in understanding the task and 'leads' them by telling Shourok what to do next and he restructures the immediate tools around them so they can begin. Shourok however might be seen to be 'disempowering' Darras by taking control, and not sharing, the tools and spatial arrangement required to interact.

6.3 typing the heading

From 00:17 seconds to 01:09, after the light fist-bump on the table, Darras sits without movement or sound as he watches Shourok insert a picture onto page 4 of their booklet and insert a text box above the picture to write a heading. This small span of time will not be analysed as very little happens. Shourok enters the title for the fourth page and writes *'My thoughts*.' Darras leans into the screen to see what Shourok is writing and makes a suggestion at 01:09: 'make a space.'

Transcript 6-3 (01:09 – 01:40)

01:09	Darras	make a space
01:13	Shourok	ºcentreº
01:21	Darras	you can make it best fit (5) but make make space
01:29	Shourok	space for what↑
01:30	Darras	my thoughts space
01:38	Shourok	بگیر عینک ! نمیبینی؟. بهتره ان با .
		{It is better with the other one. Can't you see? Get glasses!}
01:40	Darras	haha (2) haha (2) nice

At 01:09 Darras is making a suggestion to Shourok that he needs to put a space between the words '*My'* and '*thoughts*.' Shourok appears to ignore the comment by Darras, or not hear him, as he gives no indication of hearing him. Instead Shourok speaks quietly to himself 'ºcentreº' as a vocalisation of what he is doing as he doing it: he makes the heading centre-aligned in the text box. The writing on the screen is very small so Darras offers another suggestion at 01:21 'you can make it best fit.' Best Fit is a function in Microsoft Publisher for automatically resizing text in a text box so that it fills the box. This is a very good suggestion but again Shourok appears to ignore this. Darras returns to his previously unacknowledged suggestion to put a space between the two words and says 'you can make space.' Shourok now replies with rising intonation 'space for what.' Darras replies 'my thoughts space'; i.e. put a space between the two words. Shourok

switches to Azerbaijani (Azeri Turkish); in translation: ‘It is better with the other one. Can’t you see? Get glasses!’ . بگیر عینک ! نمیبینی؟. بهتره ان با Darras laughs.

A speech act typically requires a reciprocal action. Even if that is not reciprocated with speech, one would expect a gesture or some acknowledgement. In this transcript there are two instances of ‘failed’ communication, where Shourok does not acknowledge Darras has said anything: ‘make a space’ and ‘make it best fit.’ Extract 6-2 looks at the same language episode but takes into consideration the on-screen and off-screen multimodal interaction.

Extract 6-2 (01:09 – 01:30)

Ref:		Seconds:	01:09	01:13	01:21	01:26	01:29	01:30
13.	Darras	talk	make a space		you can make it best fit	but make make space		my thoughts space
14.		actions				☞points at screen ❑		☞❑ again
15.		keyboard						
16.		screen						
17.		mouse						
18.		gaze	❑----------					----------❑
Ref:		Seconds:	01:09	01:13	01:21	01:26	01:29	01:30
19.	Shourok	talk		ºcentreº			space for what↑	
20.		actions						
21.		keyboard						
22.		screen	⇧⇧clicks centre icon⇧⇧⇨makes text box bigger⇨⇦⇦⇦⇦clicks font size icon 14⇦⇦⇦⇦⇦					
23.		mouse	----------|				|----------|
24.		gaze	----------|				|----------|

In Extract 6-2, Darras attempts to be an active participant in the design process, explaining to Shourok the next step in the task of creating a booklet. He then sits quietly for fifty-two seconds as Shourok begins the task. In Extract 6-2, at Ref: 13(01:09) Darras shows his eagerness to be active in the design process as he thinks he has seen a mistake so tells Shourok: 'make a space.' It is an imperative verb to change something, functional language which lacks the social pleasantries of more interactional discourse. It is transactional language, brief and to the point to move the task quickly forwards. Shourok does not respond to the suggestion, possibly because he does not hear him, which is unlikely given their close proximity, or perhaps because he does not understand. Make a space where? The writing on the screen is very small, so small it looks like the two-word heading *Mythoughts* is one word. Darras is telling Shourok to put a space between the two words. We know this from Ref: 13 and 14(01:26) where Darras later uses a deictic gesture by pointing at the screen where the heading is and repeats 'you can make space.'

His non-response to Darras's 'make a space' is a quiet verbal utterance to himself of what he is doing, Ref: 19(01:13) 'centre.' At Ref: 22(01:09) Shourok moves the cursor up to the menu bar and selects centre to move the text to the middle of the text box.

At Ref: 31(01:21) Darras says 'you can make it best fit.' This verbal output is synchronised with the mouse movements of Shourok at Ref: 22(01:21). Shourok is making the text box bigger by stretching it left and right. The text however remains the same size: too small. Darras has remembered from previous learning that the function 'Best Fit' in Microsoft Publisher is the simplest method for automatically resizing text to fill a text box. It would seem Shourok has forgotten this function so on seeing Shourok stretching the text box, but the text remaining the same size, Darras makes a highly relevant suggestion. Again it would appear that Shourok ignores him as he gives no signal (such as a word, a noise, a gesture) to acknowledge that Darras has said anything. Shourok does the opposite of what Darras suggests and at Ref: 22(01:26) makes the font bigger by clicking '14' on the font size drop-down box. This is the worst choice compared to the 'Best Fit' suggestion of Darras. Again, at the same time as we see an on-screen action by Shourok, there is another simultaneous re-action by Darras. As the font size is increased at Ref: 22(01:26) Darras repeats 'but make make space' at Ref: 13(01:26). The font has marginally increased in size but Darras wants a space between the two words of the heading. At this point we see the first adjacency-pair as Shourok responds by asking for clarification at Ref: 19(01:29) 'space for what↑.' There have been five verbal utterances by Darras to Shourok. Four have had no verbal response over a ninety second period:

- Ref: 1(0s) 'page 4'
- Ref: 1(4s) 'page 4'
- Ref: 13(01:09) 'make a space'
- Ref: 13(01:21s) 'you can make it best fit'

Darras immediately responds with a clarification, Ref: 13(01:30s) 'my thoughts space.' Four seconds later in Extract 6-3, Ref: 32(01:34) Shourok leans back for a second and then forwards again.

Extract 6-3 (01:34 – 01:46)

Ref:		Seconds:	01:34	01:38	01:40	01:41	01:42	01:46
25.	Darras	talk			haha		haha	nice
26.		actions	reaches for the mouse		swings left and right on chair			
27.		keyboard						
28.		screen						
29.		mouse						
30.		gaze						
Ref:		Seconds:	01:34	01:38	01:40	01:41	01:42	01:46
31.	Shourok	talk		{It is better with the other one. Can't you see?	Get glasses!}			
32.		actions	leans back (1s)	then forwards		turns to face Darras		
33.		keyboard						
34.		screen					clicks font size icon 28	
35.		mouse						
36.		gaze						

Perhaps Darras misreads the leaning back as he responds immediately at Ref: 26(01:34) by leaning forwards and reaching for the mouse. A leaning back gesture might be interpreted proxemically as an invitation to occupy a space; in the same way a person might step aside to allow a person to enter their space such as when opening a door. It would appear that Darras wants to use the mouse to correct the problem of the spacing in the heading and briefly leans into the space (Figure 6-6, 01:34) and places his hand over the mouse. Shourok counters this movement towards the mouse by immediately leaning forwards to the monitor and not releasing the mouse. The space is contested and Darras yields, leaning backwards, as Shourok takes back the space in front of the screen and maintains possession of the mouse (Figure 6-6, 01:35). The interaction lasts only a second but close multimodal analysis reveals in this instance what is happening in addition to language as learners 'compete' for space and tools in collaboration. Compared to the Polish students, there is very little modal, proxemic and linguistic alignment between the two Kurdish learners.

Figure 6-6 (01:33 - 01:35)

In Extract 6-3, Ref: 31 and 34, Shourok moves the mouse back to the font drop-down box to select size 28 and at the same time speaks in Azerbaijani: نمیبینی؟. بهتره ان با . بگیر عینک ! 'It is better with the other one. Can't you see? Get glasses!' Shourok appears to be making a comparison between the heading at font size 14 and font size 28; suggesting he preferred the former. It is also a strange response to Darras's 'make a space' suggestion; perhaps he misinterpreted what Darras was referring to, as at no point does he actually confirm exactly where he thinks a space needs to go.

- Ref: 13(01:09) 'make a space'
- Ref: 13(01:26) 'you can make space'
- Ref: 13(01:30) 'my thoughts space'

These are all imprecise uses of language (perhaps because of English as a second language) where the signifier is vague and hence what is being signified is not understood. In addition, the heading '*My thoughts*' does not require a space in the middle. It only appears as such on-screen when the font was very small. Once the font is enlarged it is clear there is already a space in the middle; all the more confusing for Shourok when someone keeps repeating the word 'space' but he cannot see where a space is needed. Perhaps it is for this reason of Darras's confusion that Shourok switches to Azerbaijani and sarcasm and says: 'get some glasses.' Darras laughs at his response at Ref: 25(01:40) and Shourok turns to face Darras at Ref: 34(01:41) and smiles. Once Darras sees the increased font size of the heading he accepts the change with a complimentary 'nice' at Ref: 25(01:46).

In summary, rather than mirroring behaviour and alignment, such as seen with the Polish learners, here we see oppositional positions and language. Oppositional behaviour evidences misalignment and sometimes the contesting of interaction which

can undermine an inclusionary participation framework; not all learners in the collaboration will have equal input and voice. Darras lets his voice be heard as an active contributor, even though Shourok maintains possession of all the tools – worksheets, keyboard and mouse. Certain elements of the design process are therefore contested by the learners and there is a lack of mutual coherence. The common ground is there (write a heading) but Shourok makes little effort to acknowledge the comments of Darras, nor give any visual cues, to enable a conversation to take place about the writing of the heading. This does not mean their interaction is unsuccessful, compared to the Polish learners, but 'looks' different even though the same outcome is achieved.

6.4 writing sentence 1

At 02:06 Shourok begins to write the first sentence under the image he inserted on page 4. He speaks aloud the words he is typing, deleting '*can'* almost as soon he writes it. '*I ~~can~~ think we can always help our Earth at any time when we wake up, getting shower* …' Darras is a spectator for the fifty seconds it takes Shourok to write this. At no point does Shourok ask Darras for any input or what he thinks about the construction of the sentence. The analysis starts again when Shourok pauses in thought, his hands resting on the keyboard. At which point Darras questions what Shourok means by the phrase 'getting show,' misreading 'shower' as 'show.'

Transcript 6-4 (03:12 – 03:27)

03:12	Darras	what do you mean↑ getting show	
03:14	Shourok	shower shower	
03:16	Darras	shower haha	
03:17	Shourok	shower it's when the ()	
03:25	Darras	take (1) take shower	
03:27	Shourok	بگیری!	{take a shower}

The following extract details the same conversation but adds the multimodal and description.

Extract 6-4 (03:12 – 03:27)

Ref:		Seconds:	03:12	03:16	03:25
37.	Darras	talk	what do you mean↑ getting show	shower haha	take take shower
38.		actions	rocks left to right on his chair as he watches the screen		
39.		keyboard			
40.		screen			
41.		mouse			
42.		gaze	❑----------		----------❑

Ref:		Seconds:	03:12	03:14	03:17	03:18	03:21	03:27
43.	Shourok	talk		shower shower	shower it's when the ()			بگیری! {taking a shower}
44.		actions		☺	-----------			
45.		keyboard						----------
46.		screen					----------	⇦ getting ~~shower~~
47.		mouse		----------			----------	
48.		gaze	----------	----------		----------		----------

Darras appears to be reading what Shourok is writing, even though he not an active participant in its construction. His gaze never leaves the screen; Ref: 42(03:12). As Shourok pauses, Darras asks a question at Ref: 37(03:12): 'what do you mean getting show.' Darras rocks slightly in his chair. He mistakenly reads on the screen 'show' instead of shower. Shourok replies at Ref: 43(03:14) with repetition for emphasis 'shower shower' and smiles to himself; possibly because of Darras's mistake. Darras repeats the word 'shower,' laughing; perhaps laughing at his misreading of the word. At Ref: 43(03:17) Shourok tries to explain what he means by 'getting shower' but the sentence is very brief and is difficult to understand: 'it's when the ().' There are a couple of extra words he speaks which might be a language-switch but it is not possible to transcribe in English or Azerbaijani. As Shourok explains he turns to face Darras at Ref: 48(03:17) and uses a gesture towards his body which might be a visualisation of washing himself.

Given the context of environmentalism, it is probable that what Shourok wants to communicate is the water-saving principle of using a shower instead of a bath. What he perhaps lacks though is the language to communicate that. Shourok moves the cursor on-screen to the end of the sentence Ref: 46-47(03:21) and then presses backspace on the keyboard and deletes 'getting shower' – one letter at a time; Ref: 46(03:27). At the same time the Learning Support Assistant (LSA) sits on the table behind the two learners. She yawns and says nothing; perhaps because neither learner asks for her help. She sits quietly and does not interact with them.

Figure 6-7 (03:17 and 03:27)

There is conversational grounding and conventional turn-taking synchronised with gesture and gaze but again a lack of mutual coherence about the sentences. It has taken Shourok fifty seconds (from 02:06 – 02:56) to write '*I ~~can~~ think we can always help our Earth at any time when we wake up, getting shower …*' but then pauses after he types 'shower.' Interesting to note that Shourok uses the first-person '*I think*' at the start of the sentence as opposed to 'we think.' This is a collaborative task but the language choice of Shourok, and his domination of the tools around them (worksheets, keyboard, mouse) mean that Darras is at risk of being peripheral and marginalised in the task. In addition, Shourok uses the first person '*my thoughts*' rather than 'our thoughts' when writing the heading. To what extent does this language choice of the singular marginalise Darras in the task as a subordinate collaborator? How does he feel about Shourok's language choice of first person pronouns? He cannot be an equal collaborator when the referents of the written language on-screen are singular instead of plural and he has limited access to the tools so that he might be more active in the

design process. We will come back to this later in section 6.6. For now he does not complain but shortly he will contest the design process.

Shourok pauses as he thinks about how to continue the sentence and quietly speaks aloud the words to himself as he types them: ‘knowing how to use elcti’ to replace the confusing ‘getting shower.’ ‘*I ~~can~~ think we can always help our Earth at any time when we wake up, ~~getting shower~~ knowing how to use elcti.’* Shourok then pauses writing and mumbles something unclear at 03:55.

Transcript 6-5 (03:55 – 04:26)

03:55	Shourok	()
03:56	Darras	elec
03:57	Shourok	lec(h)tri
03:58	Darras	e:lec: tri (3) r
04:01	Shourok	() (14) I can’t do it
04:26	Darras	haha

Extract 6-5 and detail the same conversation but add the multimodal interaction to detail how the learners try to spell the word ‘electricity.’

Extract 6-5 (03:55 – 04:06)

Ref:		Seconds:	03:55	03:56	03:57	03:58	04:01	04:02	04:03	04:04	04:05	04:06
49.	Darras	talk		elec		e:	lec: r					
50.		actions					(points at the R on the keyboard)					
51.		keyboard										
52.		screen										
53.		mouse										
54.		gaze										
Ref:		Seconds:	03:55	03:56	03:57	03:58	04:01	04:02	04:03	04:04	04:05	04:06
55.	Shourok	talk	()		lec(h)tri							
56.		actions										
57.		keyboard									r i	
58.		screen								~~r~~ ~~i~~	r i	
59.		mouse										
60.		gaze										

At Ref: 55(03:55) Shourok mumbles something, possibly a complaint or request for help to spell the word. He has typed '*elcti*.' Darras replies by pronouncing the first two syllables at Ref: 49(03:56); 'elec' – adding the 'e' vowel. Darras turns his gaze to the keyboard [Ref: 54(04:01)] as well as leaning forward to point to the 'R' on the keyboard. Shourok responds by deleting the 'i' and typing '*ri'* [Ref: 57, 58(04:05)] so the on-screen spelling of the word now is '*elctri*.'

Extract 6-6 (04:18 – 04:27)

Ref:		Seconds:	04:18	04:19	04:20	04:21	04:22	04:23	04:24	04:25	04:26	04:27
61.	Darras	talk									haha	
62.		actions	(tries to spell the word with his finger on the table then uses a pen on the mousemat)									
63.		keyboard										
64.		screen										
65.		mouse										
66.		gaze										
Ref:		Seconds:	04:18	04:19	04:20	04:21	04:22	04:23	04:24	04:25	04:26	04:27
67.	Shourok	talk								I can't do it		
68.		actions										
69.		keyboard		l	e	c	t	r	i	c	i	t e
70.		screen	e ~~l e t r i~~	l	e	c	t	r	i	c	i	t e
71.		mouse										
72.		gaze										

At Ref: 70(04:18) Shourok deletes all the letters in the word leaving only 'e' and then starts to retype the word slowly a letter at a time. As he presses a letter on the keyboard (Ref: 69) we see the letter appear on the screen (Ref: 71) and at the same time Shourok looks at the screen to check the developing spelling of the word; Ref: 72. Darras at Ref: 62(04:18) first tries to spell the word using his finger on the desk then picks up a pen and appears to be trying to spell the word on the mousemat. Both of Darras's iconic gestures last less than two seconds so we might surmise that it was something of a half-attempt to spell the word. Nevertheless, Darras shows his willingness to collaborate in the task by trying to mime a word on the desk, verbalise a likely spelling at syllable level and even point to the keyboard to suggest the next letter in the sequence. Although Shourok remains in 'control' of the primary tools of the design process, Darras continues to engage with the process. At Ref: 67(04:25) Shourok says 'I can't do it,' and Darras laughs loudly at Ref: 61(04:26).

Shourok quickly types the word 'water' at the end then leans backwards and stretches with his hands above his head. The sentence now reads as: *I ~~can~~ think we can always help our Earth at any time when we wake up, ~~getting shower~~ knowing how to use electricite, water.'* Shourok must know the word is still wrong as the on-screen display shows a red underline on '*electricite*.' Perhaps the red visualisation of the auto-spellcheck feature is why he declares 'I can't do it.' He seems to accept it is wrong and that he cannot change it. Why does he not ask Darras for his thoughts? The LSA is sat behind them and she could help. He could use the spellcheck function of the Publisher Program to suggest a spelling. He could call for the teacher. He could go online. There are many alternative tools he could use but he appears to be content with leaving it as wrong for now.

Figure 6-8 (03:57 – 04:18)

The first image shows Shourok turning to face Darras for help with spelling the word electricity. This is a rare instance of Shourok asking for help from Darras. The second image (04:01) shows Darras stretching forwards to identify that 'R' is the next letter in the word. The third image (04:18) is of Darras using an iconic gesture to try and spell the word; first with his finger and then miming the word with a pen without actually writing any letters on the mousemat.

In Extract 6-6 there is temporal mapping of speech where sounds are slowed down to try to help phonically with spelling the word. The oral syllabification of the sounds is also passed between each learner as a form of echo utterance. In Extract 6-7 there is a similar temporal mapping where the typing of a single word is slowed down to one letter at a time with synchronised checking on-screen of each letter. The temporal ordering of these different modes structures the sequentiality of their interaction and helps them to achieve understanding: i.e. how to spell the word, even though it remains

incorrect at the moment. Both learners are involved in this short interaction and neither is marginalised.

The sentence so far in its ongoing construction reads as follows: *I ~~can~~ think we can always help our Earth at any time when we wake up, ~~getting shower~~ knowing how to use ~~elcti~~ electricite, water'*. The sentence ends without punctuation, suggesting there is more to write, and Shourok stops at this point and leans backwards and stretches. Darras says 'and,' wondering what the next part of the sentence might be. At 04:41 Darras holds the mouse and clicks on the screen to place the cursor at the end of the sentence. In doing so he draws Shourok back into the design process.

Transcript 6-6 (04:41 – 05:30)

04:41	Darras	last one you should put and here
04:44	Shourok	no there's new gas (10) <petrol>
05:05	Darras	many thinKS <things> thing things
05:10	Darras	and many other things around us (1) <many other> (3) things (2) thing
05:22	Shourok	ha
05:30	Darras	things around us

Transcript 6-6 is an unusual language interchange and is a good example of the extent to which talk is mediated by other modes in the immediate environment. 'The way in which artefacts and other contextual phenomena are incorporated into the talk is emergent, where the talk moves forward alongside the ever-changing multimodal text that is being created on the computer' (Gardner & Levy, 2010: 3). Extract 6-7 to Extract 6-9 itemise the differrent modes and these help to better understand the interaction and the mediation of talk in Transcript 6-6. At 04:41 Darras tells Shourok 'last one you should put and here.' Darras is drawing attention to the punctuation point that if you are writing a list of objects and using commas to separate them, then on the last item you don't use a comma but add 'and' before the final object in the list. Shourok though is not finished with the list. He is the one composing the sentence and Darras has peripheral agency in the process. Shourok disagrees at 04:44 with 'no there's new,' suggesting there is more he wants to write. Extract 6-7 continues from this point when Shourok starts typing the rest of the sentence.

Extract 6-7 (04:41 – 04:59)

Ref:		Seconds:	04:44	04:51	04:52	04:53	04:54	04:55	04:56	04:57	04:58	04:59
73.	Darras	talk	gas				<petrol>					
74.		actions										
75.		keyboard										
76.		screen										
77.		mouse										
78.		gaze	❑----------							❑🗣--------	🗣❑----	❑
Ref:		Seconds:	04:44	04:51	04:52	04:53	04:54	04:55	04:56	04:57	04:58	04:59
79.	Shourok	talk										
80.		actions										
81.		keyboard		, gas		,	p	etr	ol		an	d
82.		screen		, gas		,	p	etr	ol		an	d
83.		mouse										
84.		gaze	💻----------									----------💻

Darras remains static with the exception of turning his gaze towards Shourok at Ref: 78(04:58). He otherwise looks at the screen as Shourok types. Speech is slowed down as Shourok temporally maps what he is thinking and saying onto the screen through the keyboard. He says the word ‘gas’ at Ref: 80(04:44) but the typing and the appearance on the screen do not appear for a few seconds later at Ref: 80(04:51). His utterance of ‘petrol’ has a ten second delay and he says the word slowly as he types at Ref: 81(04:54) which appears on the screen at Ref: 81 in synchrony with him saying the word. Extract 6-8 continues with the same interaction and talk from Transcript 6-6.

Extract 6-8 (05:00 – 05:10)

Ref:		Seconds:	05:00	05:01	05:02	05:03	05:04	05:05	05:06	05:07	05:08	05:10
85.	Darras	talk						many thinKS	<things>	thing	things	
86.		actions						(rocks left and right on his chair as he talks)				
87.		keyboard										
88.		screen										
89.		mouse										
90.		gaze	❑----------									----------❑
Ref:		Seconds:	05:00	05:01	05:02	05:03	05:04	05:05	05:06	05:07	05:08	05:10
91.	Shourok	talk	and many things									
92.		actions										
93.		keyboard	man		y		thin	ks	ar	ound	us	
94.		screen	man		y		thin	ks	ar	ound	us	
95.		mouse										🖱---🖱
96.		gaze	💻----------									----------💻

At Ref: 91(05:00) Shourok says to himself quietly ‘and many things’ and the final word ‘things’ appears on the screen five seconds later at 05:05. There is a typographical error though written as ‘*thinks*’ and Darras immediately spots the error and draws Shourok’s attention to it at Ref: 85(05:05) by saying ‘many thinKS.’ He adds stress on the ‘K’ (saying it more loudly) possibly to phonetically highlight the on-screen error for Shourok. At 05:06 Darras stresses the word ‘things’ again by saying the word slowly but pronouncing correctly what the word should be. At 05:07 he again repeats ‘thing,’ dropping the ‘s’ and at 05:08 repeats ‘things’ putting the ‘s’ back on. Darras has used four tactics to draw Shourok’s attention to the error: loudness, slowness, singular and back to the plural. Over the four seconds Darras repeats the word Shourok does not acknowledge having heard him. Perhaps this is why Darras keeps repeating. However, at Ref: 95(05:10) Shourok reaches for the mouse and moves the cursor to the end of the word ‘*thinks’* in preparation to delete it. He has heard him but gives Darras no verbal confirmation. The effect of which leaves Darras stuck in a message loop. It is customary in communication that if someone does not acknowledge something you have said then they might not have heard you so the same words are repeated. It is a standard convention of turn-taking in dialogue (Sacks et al, 1974). Extract 6-9 continues with Transcript 6-6 and we find the same problem of Darras repeating the word for Shourok to change the error.

Extract 6-9 (05:12 – 05:30)

Ref:		Seconds:	05:12	05:14	05:16	05:18	05:20	05:22	05:24	05:26	05:28	05:30
97.	Darras	talk	many other things around us <many other>				things	thing				things around us
98.		actions										
99.		keyboard								things		
100.		screen								things		
101.		mouse										
102.		gaze	🖵 ------------	------------	------------	------------	------------	------------	🖵 ⌨ ------	------------	------ ⌨ 🖵 ------	🖵
Ref:		**Seconds:**	**05:12**	**05:14**	**05:16**	**05:18**	**05:20**	**05:22**	**05:24**	**05:26**	**05:28**	**05:30**
103.	Shourok	talk						ha				
104.		actions		✋ (rubs his forehead and face)				☺	✋ (stop, then grabs the recorder)			
105.		keyboard				~~thinks~~	o	ther				
106.		screen	⇦⇦⇦⇦⇦							⇨⇨⇨⇨⇨	⇨⇨⇨⇨	⇨⇨
107.		mouse	🖱---------	🖱						🖱------------	------------	------🖱
108.		gaze		⌨------------	------------	------------	------⌨	🗣--------	🗣 💻------	------------	------------	------💻

At Ref: 97(05:12) Darras continues to repeat ‘things’ but now adds ‘other’ to the noun phrase: ‘many other things around us.’ At the same time, Shourok uses the mouse to move the cursor to the end of ‘*thinks*’ and begins to delete it using the backspace with his right hand. At Ref: 104(05:12) Shourok rubs his forehead and face with this left hand. Darras perseveres with his word choice at Ref: 97(05:18) and repeats but more slowly: ‘<many other>.’ He pauses for two seconds and repeats ‘things’ and then a second later says again ‘thing.’ At this point Shourok reacts with what might be a laugh of incredulity or frustration. He turns to face Darras at Ref: 108(05:22), raising his right hand in what looks like an iconic stop gesture.

Shourok then reaches for the microphone, in what can only be interpreted as an attempt to block out whatever he is going to say to Darras. At the exact moment that Shourok releases the keyboard, Darras leans into the space and types ‘things’ at Ref: 97(324s). It is unclear if Shourok says anything. When Darras leans away from the keyboard Shourok holds the mouse at Ref: 106-107(05:25) and moves the cursor to the end of the sentence to add a full-stop. At Ref: 97(05:30) Darras once again repeats ‘things around us.’

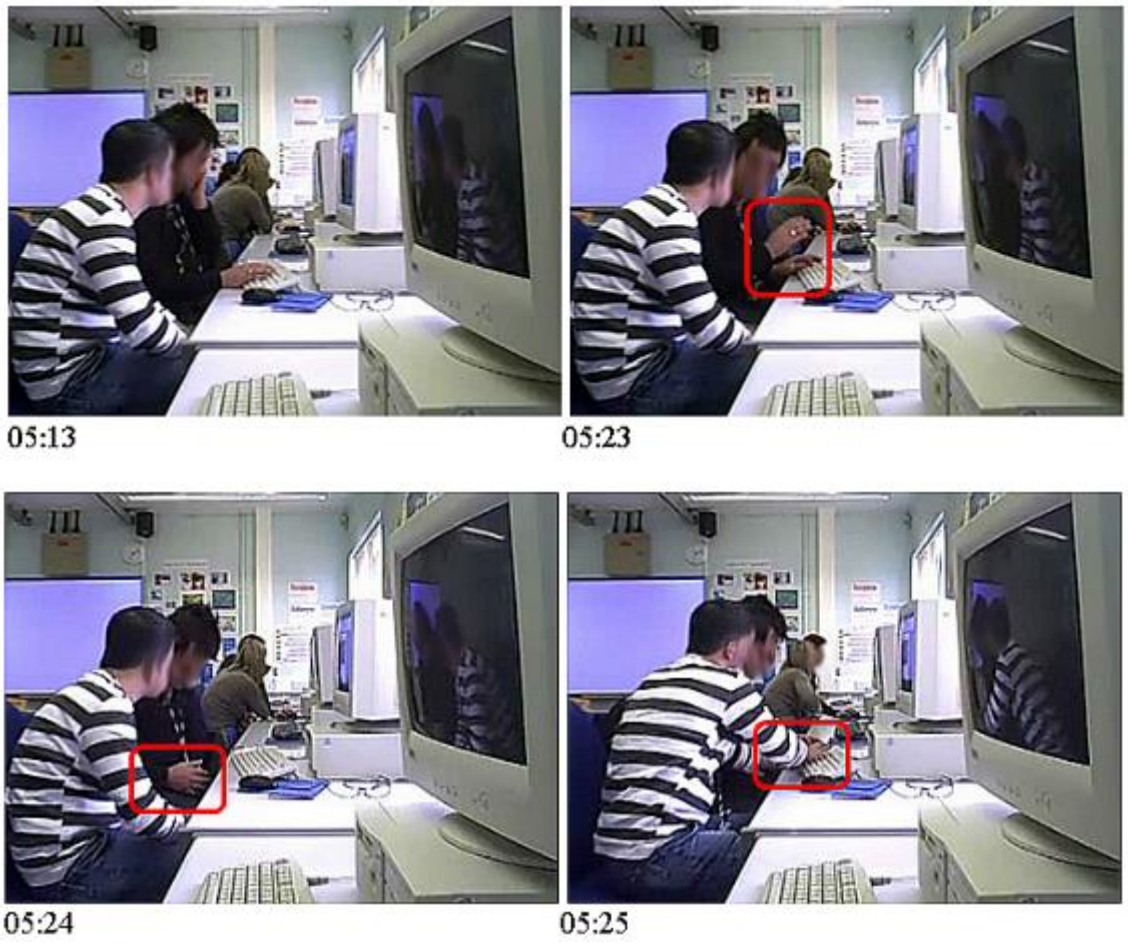


Figure 6-9 (05:13 – 05:25)

Figure 6-9 provides some images from Extract 6-9 to help with the visualisation of the learner interaction. At 05:13 Shourok looks tired (or frustrated) and rubs his forehead and face with his left hand as he simultaneously deletes ‘*thinks*.’ The image at 05:23 shows Shourok raising his hand in an iconic stop gesture. He laughs at the same time but the tone does not suggest that this is a laugh of humour, perhaps annoyance at Darras’s repeated utterance of the word ‘things;’ a total of eight times over a fifteen second period. At 05:24 Shourok grabs the voice-recorder in what must be an attempt to conceal something he is about to say to Darras. If he does say something it cannot be

heard in the recording. Interestingly though, because Shourok is distracted by the voice-recorder, this has left the keyboard free to use and Darras instantly leans forward (05:25) to write the correct the spelling of '*things*.'

At 05:44 Shourok adds a full-stop to complete the sentence. He leans back and releases the mouse. Immediately Darras holds the mouse, right-clicks on '*electricte*' and uses the spell-check feature to change to the correct spelling. He turns to face Shourok. Darras has changed what Shourok failed to do but Shourok does not acknowledge through gesture or speech what Darras has done. The first sentence is now complete and reads: '*I ~~can~~ think we can always help our Earth at any time when we wake up, ~~getting shower~~ knowing how to use ~~elcti~~ electricit~~e~~y, water, gas, petrol and many other thing~~k~~s around us.'* For Darras, the only words he helped with in the construction of the sentence are 'getting shower,' 'electricity' and 'other things.' Noticeably these were all grammatical or typographical errors by Shourok. These inputs from Darras predominantly arise when Shourok voluntarily leans away from the computer, or is distracted, and at that moment, Darras leans into the vacated space to use the mouse or/and keyboard. In the construction of the first sentence Darras 'steals' the tools, in Shourok's attentional absence, at the following points:

- 04:41 – uses the mouse to move to the end of the sentence and says 'and' when Shourok leans away from the computer
- 05:24 – to type 'things' when Shourok grabs his voice-recorder
- 05:44 – to correct 'electricite' using the mouse when Shourok leans away from the computer

From an educational perspective, if this is a collaborative task, then we need to consider the management of tools and design process to enable equality of learning opportunities. A finding here is *competitive interaction* rather than *collaborative interaction*. Temporal mapping is still co-constructed between the learners but rather than *collaborative co-construction* of meaning there is *competitive co-construction*.

6.5 writing sentence 2

Section 6.5 is a relatively short section compared to 6.4. The first sentence took 03:38 to write, between: 02:06 – 05.44. The second sentence by comparison takes only 01:12 to write, between 06:33 – 07:45. The second sentence is much shorter to construct but interestingly is very different in construction to the first sentence, where Shourok largely wrote from his own head and Darras managed to have some input where there

were mistakes. Sentence two is mostly constructed by Darras where he speaks aloud and Shourok writes down verbatim, correcting in text some of Darras's spoken grammatical mistakes. This process provides an interesting perspective on the temporal mapping of the learners. How talk and text are mediated by this process is the focus of section 6.5. The total language exchange for this section of the chapter is in Transcript 6-7:

Transcript 6-7 (06:33 – 07:45)

06:33	Darras	err finished you need to make bigger
06:36	Shourok	دیگه؟ {finished what else} ننویسیم؟ بنویسیم؟ بیشتر still {shouldn't we still continue writing}
06:42	Darras	I think that (8) we we should be responsible for recycle our bins
07:04	Shourok	for
07:06	Darras	recycling recycling our bins
07:08	Shourok	((*grunt*))
07:15	Darras	it's for recycling (5) our <bins> <recycling> <recycling> <bins>
07:40	Darras	if you change this to recycle
07:44	Shourok	((*grunt*))
07:45	Darras	recycle our bins

It is difficult to make any sense of this language exchange, even though it utilises standard turn-taking. Darras is verbally composing a sentence for Shourok to write and Shourok is reduced to a series of back-channelling 'grunts' to confirm he understands what Darras is saying. Extract 6-10 looks at the same language episode but takes into consideration the multimodal.

Extract 6-10 (06:33 – 06:42)

Ref:		Seconds:	06:33	06:34	06:35	06:36	06:37	06:38	06:39	06:40	06:41	06:42
109.	Darras	talk	err finished you need to make bigger									
110.		actions	holds the mouse and highlights the text						rests chin on right hand			
111.		keyboard										
112.		screen	⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖⇖									
113.		mouse	🖱--🖱									
114.		gaze	🖵 -- 🖵 🗣 -------------------- 🗣									
Ref:		Seconds:	06:33	06:34	06:35	06:36	06:37	06:38	06:39	06:40	06:41	06:42
115.	Shourok	talk				دیگه؟ {finished what else}			نتویسیم؟ بنویسیم؟ بیشتر still {shouldn’t we still continue writing}			
116.		actions							cancels the highlighted text			
117.		keyboard										
118.		screen							⇨⇨⇨			
119.		mouse						🖱------🖱				
120.		gaze	🖵 -- 🖵 🖱------🖱 🖵 --------- 🖵 🗣 ------------ 🗣									

At 06.33 Darras holds the mouse, highlights the text on screen and says at Ref: 109 ‘err finished you need to make it bigger.’ He thinks they have finished writing so he highlights all the text to enlarge the font size. Shourok contests Darras’s belief that they have finished by language-switching to Azerbaijani at Ref: 115(06:36) and says ‘finished what else.’ In other words, no, we have not finished; what else can we write? This interpretation is confirmed at Ref: 120(06:38) as Shourok’s gaze shifts from the screen to the mouse and he takes the mouse away from Darras and at Ref: 118-119(06:39) clicks on the screen to cancel the highlighted text. He says at Ref: 115(06:39) ‘shouldn’t we still continue writing’ but mixes English with Azerbaijani نتویسیم؟ بنویسیم؟ بیشتر still.

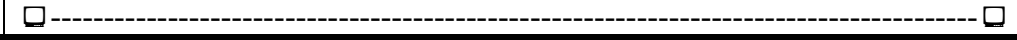
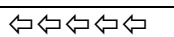
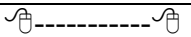
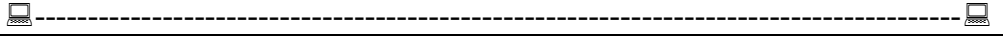
At Ref: 114(06:40) Darras turns to face Shourok as he speaks to him and Shourok mirrors his gaze at Ref: 120(06:41). It is worth noting that Shourok did not pause and wait for Darras to release the mouse, nor did he ask for it.

Figure 6-10 (06:38 – 06:39)

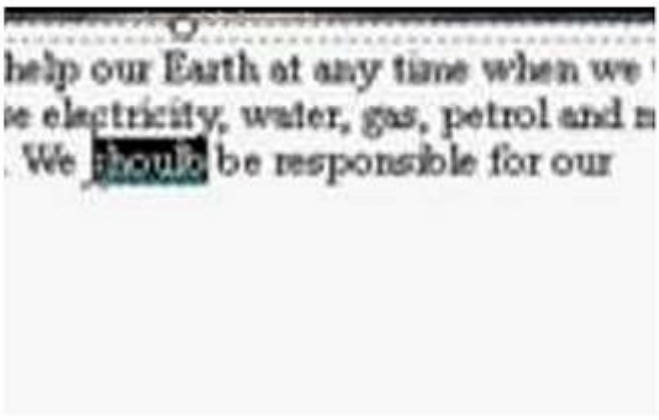
In Figure 6-10 it took one second for Shourok to take the mouse from Darras. Shourok stretches his hand out across the space of Darras and even though Darras's gaze never seems to leave the screen he knows the hand is there as he immediately releases the mouse and Shourok continues. Darras could have chosen not to release the mouse but he yields as he has done throughout their collaboration.

At 06:42, after Shourok cancels the highlighting of Darras and explains they need to write more, Darras gestures at the screen and says: 'I think that (8) we we should be responsible for recycle our bins.' There is a noticeable eight second pause before Darras says what he thinks the sentence should be. Shourok writes quickly '*We shoulb be responsible …*'

Extract 6-11 (07:04 – 07:14)

Ref:		Seconds:	07:04	07:06	07:08	07:09	07:10	07:11	07:12	07:13	07:14
121.	Darras	talk		recycling recycling our bins							
122.		actions	rests chin on right hand			leans into the keyboard to type		/	rests chin on right hand		
123.		keyboard									
124.		screen									
125.		mouse									
126.		gaze									
Ref:		Seconds:	07:04	07:06	07:08	07:09	07:10	07:11	07:12	07:13	07:14
127.	Shourok	talk	for		((*grunt*)) suggesting a pause						
128.		actions									
129.		keyboard	for our						should		
130.		screen	for our					~~shoulb~~	should		
131.		mouse									
132.		gaze									

At Ref: 127, 129, 130(07:04) Shourok says ‘for’ as he types the word and it appears on the screen at the same time. At 07:05 he types ‘*our*’ but has dropped (intentionally or not) the word in the middle which Darras offered: ‘for recycle our.’ At Ref: 121(07:06) Darras hears Shourok say ‘for’ and continues the sentence verbally ‘recycling recycling our bins.’ At Ref: 127(07:08) Shourok grunts in response to Darras’s ‘recycling’ and the grunt is suggestive of a wait, asking Darras to wait a moment. At the same time as he grunts, Shourok highlights the word ‘*shouldb’* (Ref: 130,131 (07:08). It would seem that the grunt to pause is because Shourok has seen the misspelt ‘*should’* and wants to correct it but Darras is pushing him on with the sentence so that at Ref: 122(07:09) Darras actually leans into the keyboard, in the gap where Shourok stops typing but holds the mouse at 07:08, presumably to try and type ‘recycling.’



07:09 (Shourok deleting) 07:09 (Darras trying to type)

Figure 6-11 (07:09)

Here is another example of two learners working collaboratively but seemingly at odds in the process. We find recurring instances where proxemic and linguistic alignment with each other is lost; possibly because of the continuing *competitive interaction* rather than *collaborative interaction*. Darras provided a straightforward sentence. Shourok is typing it but dropped the word 'recycling' so Darras repeats it and leans into the computer to type it but fails to see that Shourok has moved backwards to a previous error. The sequential process of off-screen talk to on-screen text is momentarily broken. Why? Is the fault with Shourok who rarely explains what he is doing and why and would appear to accept Darras in the process as a witness rather than a participant? For example, he grunts at Darras to wait rather than explaining that he wants to backtrack to change something. Is it then a communication failure? In that small window of uncertainty when Shourok stops typing and Darras repeats 'recycling' it would appear that Darras is being proactive by typing the word himself. Perhaps he misreads Shourok's pause as one of confusion so tries to help. Darras's hand pauses over the keyboard for two seconds until Shourok then starts to use the keyboard again at Ref: 130(07:11). He deletes '*shoulb'* and types '*should*.' Darras retracts his hovering hand away from the keyboard.

Extract 6-12 (07:15 – 07:33)

Ref:		Seconds:	07:15	07:17	07:19	07:21	07:23	07:25	07:27	07:29	07:31	07:33
133.	Darras	talk	it's for recycling			our <bins>	<recycling>		<recycling>	<bins>		
134.		actions	rests chin on right hand									
135.		keyboard										
136.		screen										
137.		mouse										
138.		gaze	❑----------									----------❑
Ref:		Seconds:										
139.	Shourok	talk										
140.		actions										
141.		keyboard		⇦	r e c y c l b n g	our	beens		b ⇦	bins		
142.		screen		~~our~~	r e c y c l b n g	our	beens		b~~eens~~	bins		
143.		mouse										
144.		gaze	💻----------									----------💻

At Ref: 133(07:15) Darras continues the second half of the sentence, identifying on-screen what has already been written ('*we should be responsible for*') so speaking only the second half the sentence: 'it's for recycling.' 'It's' is not part of the sentence construction but belongs to the conversation; an instruction. It might function as an anaphoric reference to the sentence spoken earlier by Darras. The identity of 'it's' is only understood by anaphoric reference back to the sentence he spoke earlier: 'be responsible for recycle our bins.' More likely, 'it's' is functioning as a cataphoric, future referent, i.e. 'The next words in the sequence you need to write are *for recycling*.' Either way, Shourok deletes 'our' at Ref: 142(07:17) so that he can write the first word of the language chunk Darras has just spoken. He correctly identifies the 'it's' referent as redundant to the sentence, and that '*for*' has already been written, so he ignores those and continues to type at Ref: 142(07:19) '*recyclbng*;' with a spelling error. Darras can see the word as it is being typed on the screen so speaks the next word in the sequence as soon as '*recyclbng*' is finished being typed on the screen: 'our <bins>' he says at Ref: 133(07:21).

Darras now begins to slow down his talk to allow his words to closer synchronise with the typing of Shourok. It takes Shourok two seconds to type the word '*reyclbng'* so it perhaps for this reason that Darras slows down his talk to allow Shourok to catch up; slurring the word '<bins>' at 07:22.

At Ref: 133(07:23 and 07:27) Darras again slows down his speaking of the word 'recycling' as Shourok types what Darras has just said, each typed word temporally lagging about a second behind its spoken equivalent. At Ref: 141(07:22 and 07:25) '*our beens*' appears on-screen a second after Darras pronounced each word. In the meantime, Darras slowly repeats the word '<recycling>' twice; perhaps because he can see the red underline on the on-screen display, meaning the word is spelled incorrectly. He might be trying to sound the word out to hear its spelling. At Ref: 141(07:27) Shourok deletes '*beens'* and retypes correctly as '*bins*.' At 07:29 both learners align linguistically as Darras says '<bins>' slowly and Shourok types the word; completing the sentence.

The original sentence spoken by Darras at 06:45 was 'we should be responsible for recycle our bins.' It took two seconds to say it. The completed sentence is virtually verbatim but for Shourok changing Darras's incorrect 'recycle' to 'recycling' by adding the 'ing' present continuous: *'We should be responsible for recycling our bins.'* With all the corrections the history of the typed sentence is as follows: ~~*w*~~*We shoul*~~*b*~~*d be responsible for* ~~*our*~~ *recycli*~~*b*~~*ng our b*~~*een*~~*ins.* It took Shourok fifty-eight seconds (06:47 – 07:45) to write the sentence. It can be seen from the analysis so far that there is a sequential mapping of off-screen talk to on-screen text which follows a time-lagged arrangement. The learners temporally map the spoken onto the textual by (1.) Darras uses pauses when speaking the sentence to allow Shourok to type, and (2.) drags out the sound of the syllables as Shourok is typing them, and (3.) repeats words.

The writing of the third and final sentence in their collaboration will not be analysed. It is the shortest of the three in construction as Shourok 'borrows' from the second sentence Darras composed and adds to it but with no input from Darras, other than his suggestion to pluralise 'city' to 'cities'. In completion it reads as: '*Be responsible for our streets, our cities., then we can say we took a place in Greenpeace*.' In writing it went through the following editing stages: '~~*b*~~*Be responsible for our streets, our cit*~~*y*~~*ies., then we can say we took a place in* ~~*g*~~*Greenpeace*.' The sentence takes seventy seconds

to write and there is no interaction as Darras sits passively with his head in his right hand.

6.6 contesting the design process

Darras and Shourok collaborate to create a webpage called 'information' for their website about the environment. They are using Macromedia Dreamweaver. They have completed their booklet and are now creating a website which is an online version of their booklet. Within the overall classroom sequence (Appendix C) this focal event is higher-level action 3.4 within the larger third action sequence of creating a website about Greenpeace and environmentalism. The video of this interaction is available on the accompanying CD called **Kurdish_3.4**.This focal point has been chosen because it evidences Darras's contesting of the design process and Shourok's subsequent reluctance to collaborate when Darras has control of the computer, resulting in a breakdown of their participation framework. It is an interesting dynamic which evidences all the modes seen in the previous two analysis chapters: proxemics, gesture, speech, text, spatial layout, posture, tone, gaze, etc. This section will bring all the modes together in a final detailed analysis to understand the cross-modal configurations in a contested communicative event.

The Learning Support Assistant (LSA) sits behind the two learners in a triadic configuration and draws attention to the fact that Darras is sitting quietly; Figure 6-12, by saying 00:09 'Darras you need to say something don't just sit quiet;' Transcript 6-8.

Figure 6-12 (LSA sitting behind the learners)

Transcript 6-8 (00:02 – 02:47)

00:02	LSA	you missed one one picture
00:05	Darras	you found that one ah
00:09	LSA	Darras you need to say something don't just sit quiet haha
00:16	Shourok	mhm

00:21	Darras	he doesn't listen to me haha
00:23	Shourok	what did you say
00:24	Darras	he's only doing himself
00:27	LSA	haha
00:29	Shourok	cheers uuh↑
00:31	Darras	nothing
00:32	LSA	without sharing ideas
00:34	Darras	yes
00:35	Shourok	yes yes we share ideas after work
00:39	Darras	yes after [work
00:40	LSA	[really haha
00:42	Shourok	no I'm joking come on say something
00:47	LSA	I know you're very quiet (3) yes
00:52	Darras	yes
01:01	Shourok	done all yours
01:07	Darras	she doesn't mean this
01:09	LSA	no(h) no(h) (h)no really (h)no (h)no I did(h)n't me(h)an it haha
01:13	Darras	this page yes (2) change that one↑
01:17	Shourok	save you can delete that one
01:19	Darras	eh↑
01:20	Shourok	you can delete
01:22	Darras	save (2) and here open↑ (6) this↑ (5) what is this↑
01:37	Shourok	(greenpeace)
01:40	Darras	greenpeace (06) and↑ (1) it should be what↑
01:49	Shourok	information
02:12	Darras	and↑ (1) what do we have for information
02:15	Shourok	nothing
02:16	Darras	th(h)at's very dreadful what can we write find something on the internet yes↑
02:27	Shourok	there is information ()
02:31	Darras	()
02:40	Shourok	open information
02:47	Darras	hand me

There are several intonation units in the form of adjacency pairs. The Q&A is a structuring feature in the conversation but it does not appear to be particularly effective in helping Darras. Turn-taking as a feature of transactional language has broken-down as the intonation units spoken by Darras receive little reciprocal response. For example, 01:22 has three intonation units in the form of questions directed at Shourok: 'and here open↑ (6) this↑ (5) what is this↑.' There are significant gaps of silence between Darras's unanswered questions for Shourok (as at 01:22) and in the response times between them both for the rest of the conversation. Language alone makes it difficult to explain these silences, so for analysis to understand the totality of the interaction, we need to identify the other modes alongside talk. The extracts and images below explore some of the other modes which belong to this language interchange in the above transcript.

Extract 6-13 (01:13 – 01:09)

Ref:		Seconds:	01:00	01:01	01:02	01:03	01:04	01:05	01:06	01:07	01:08	01:09
145.	Darras	talk								she doesn't mean this		
146.		actions										
147.		keyboard										
148.		screen										
149.		mouse								🖱----------		----------🖱
150.		gaze	🖵----------		🖵🗣----------		Shourok		----------🗣	🗣-------LSA-------		🗣
Ref:		Seconds:	01:00	01:01	01:02	01:03	01:04	01:05	01:06	01:07	01:08	01:09
151.	Shourok	talk		done all yours								
152.		actions		pushes keyboard to Darras and pushes himself away								
153.		keyboard										
154.		screen										
155.		mouse										
156.		gaze	🖵----------		🖵🗣----------							----------🗣

At Ref: 150, 156(01:00) both learners are looking at the screen. When Shourok says 'all yours' at 01:01, both learners look at each other at Ref: 150, 156(01:02). There are six seconds of silence when no one says anything and the two learners sit in 'awkward' silence. Darras eventually breaks the silence with what would appear to be an apologetic 'she doesn't mean this' at Ref: 145(01:07); as if he is trying to remove any blame from her. At the exact moment that Darras says 'she' he turns to look at the LSA at Ref: 150(01:07) and reaches for the mouse at Ref: 149. The LSA laughs excessively (maybe nervously) at 01:09: 'no(h) no(h) (h)no really (h)no (h)no I did(h)n't me(h)an it haha' and then walks away at 01:26. The LSA appears to be defending herself.

01:00 pushes the keyboard left

01:01 a 'giving' gesture

01:07 reaches for the mouse

01:26 LSA leaves the learners

Figure 6-13 (01:00 – 01:26)

At Figure 6-13, Shourok's talk is in red font and Darras's is black. At 01:00 and 01:01 Shourok nudges the keyboard towards Darras and pushes himself away from the table. He turns to face Darras, and with both hands together and gesturing towards Darras, says, 'all yours.' The modes of gaze, gesture, proxemics and speech combine in a pivotal moment. The proxemic act of pushing himself away from the table (and Darras) is a significant communicative moment. It opens the spatial potential for Darras to sit in front of the computer and take control of all the tools in the immediate vicinity: worksheets, mouse and keyboard. It is also perhaps a proxemic protest by Shourok because of Darras's accusations at 00:21 'he doesn't listen to me' and at 00:24 'he's only doing himself.' Shourok is in effect distancing himself from the spatial work-area around the computer and saying 'okay you do it then.' His behaviour and language disjoints from Darras and the task so that the proxemic arrangement now looks like Figure 6-14.

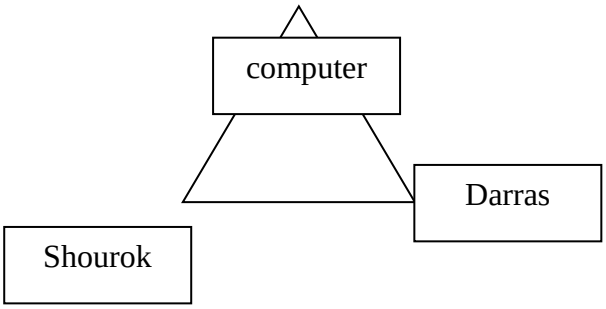


Figure 6-14 (triangular layout 3)

Shourok has distanced himself from the computer so that he is peripheral to the interaction space around the computer. Darras now has control of the computer but he retains a peripheral position to the right as if reluctant to position himself in a dominant central position. In the interaction that follows there is a sequence of proxemic actions, performed by both students, which evidence no proxemic alignment; they are in discord rather than harmony. The students have lost the personal distance required to work supportively and collaboratively. Shourok has chosen to proxemically position himself in the extreme boundaries of the layout space available to them. The microanalysis continues from 01:22. Darras is trying to understand what needs to happen next. There are a number of utterances with rising intonation, suggesting questions are being directed towards Shourok.

Transcript 6-9 (01:22 – 01:49)

01:22	Darras	save (2) and here open↑ (6) this↑ (5) what is this↑
01:37	Shourok	(greenpeace)
01:40	Darras	greenpeace (06) and↑ (1) it should be what↑
01:49	Shourok	information

Extract 6-14 (01:22 – 01:49)

Ref:		Seconds:	01:22	01:24	01:26	01:30	01:35	01:40	01:46	01:47
157.	Darras	talk	save	and here open↑		this↑	what is this↑	greenpeace	and↑	it should be what↑
158.		actions					☺ smiles briefly		moves himself central	
159.		keyboard								
160.		screen	🗁 file	⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩ opens the file greenpeace						
161.		mouse	🖱----------							----------🖱
162.		gaze	❑----------				❑🗣------🗣❑----------		❑🕮----------	🕮
Ref:		Seconds:			01:28		01:37			01:49
163.	Shourok	talk					(greenpeace)			information
164.		actions					☺ smiles briefly at Darras		looks to his left	
165.		keyboard								
166.		screen								
167.		mouse								
168.		gaze	❑----------		❑😐---😐❑----------		❑🗣------🗣❑----------		❑😐-yawns-😐🕮-🕮	

At Ref: 157,160 (01:22) Darras speaks aloud what he is doing, saying ‘save’ and then clicking File and Save on the computer. He has saved the previous page Shourok was working on and now needs to open a new ‘master’ webpage for the next page of their website which should be called ‘information.’ He spends the next twenty-five seconds trying to understand this by using his cursor to look for an item on the screen [Ref: 160(123-140)] and by talking to Shourok. At Ref: 157(01:24) Darras says ‘and here open↑’ with rising intonation but there is no response; just a six second silence. Silence is a mode which in this instance perhaps communicates an unwillingness to talk. In Figure 6-15, 01:28, whilst Darras is waiting for a reply, Shourok sits arms-crossed and peripheral to the activity, looking around the room.

Figure 6-15 (01:28 – 01:35)

Gaze is a problematic mode in terms of analysis because it is difficult to pinpoint precisely where people are looking. Though there is often structured gaze when people interact (Kendon, 1967; Goodwin, 1981) there is also sometimes randomness in where people look. We can see in the first image at 01:28 that Shourok is looking elsewhere, around the classroom. Gaze can tell us where individual attention is directed and in this instance it would seem to suggest that Shourok is off-task and that Darras's attention is on-task because he is leaning into and looking at the screen. After a six second silence, and no response, Darras tries again at Ref: 157(01:30) and says 'this↑' with rising intonation. After a four second silence he tries again at 01:35 'what is this↑' and turns to face Shourok at Ref: 162(01:35); Figure 6-15.

At 01:35, Figure 6-15, after Darras's interrogative tone on 'this↑' at 01:30, Shourok's gaze returns to the screen but his speech, posture, gesture and proxemic behaviour remains unchanged. He is still closed to any interaction and physically removed from the situation though his gaze would suggest he is showing some level of interest. Darras leans closer to the screen, his changed posture indicating greater attention towards the task, and his cursor moves backwards and forwards on-screen to find the file he needs; Ref: 160(123-140). Darras's 'this' is a demonstrative pronoun, substituted for the noun 'greenpeace' because the referent can be understood from the context, but the rising intonation in his voice conveys uncertainty and communicates to Shourok that he is asking a question. It is on the word 'this' that Shourok's attention is brought back to the interaction.

The semiotics of the spoken ‘this’ is interesting as part of the multimodal communicative act. There is Darras’s verbal ‘this,’ the visual cursor in the shape of an arrow on-screen and the blue highlighting of the word, which in itself is a visual sign to refer to the document ‘greenpeace’, each combining to communicate to Shourok that Darras is asking if this file is the correct one to open; Figure 6-16. Three signifiers working towards the one signified to achieve understanding: a webpage file name. Darras’s frequent use of ‘this’ in the two hours leading up to the computer-switch has nearly always appeared with a gesture, such as a pointing finger, but here the white cursor arrow is an embodied extension of Darras’s finger operated by proxy through his hand on the mouse. Rather than a tool within the windows environment to open and close things, Darras uses the arrow as a visual gesture and seeks confirmation from Shourok that he has the correct file.

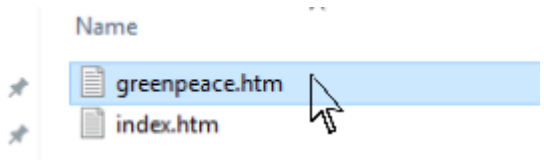


Figure 6-16 (greenpeace.htm)

This knowledge is assumed and understood. Both learners know how to interpret this brief, but complex, multimodal exchange. But how do they know? In semiotics ‘this’ is an indexical signifier (like ‘I,’ ‘you,’ ‘that,’ ‘there’) but the word also indexes both students in the interaction by positioning them contextually in relation to the task and the computer. ‘Indexicality is the property of the context-dependency of signs, especially language’ (Scollon & Scollon, 2003: 3). The meaning of language is located in the material, social world and it is the physical placement and experience of signs in this particular context which means the students can share understanding.

Figure 6-17 (01:35 – 01:36)

On-screen, the cursor hovers over the 'greenpeace' master page. Darras's gaze turns to Shourok at 01:35, Figure 6-17, as if he is checking for confirmation, and he smiles. They both smile. It is perhaps Darras's smile, more than the other modes, which communicates reconciliation but it does not yet draw Shourok back into the interaction. From Darras we see the modes of gaze, a smile and a visual cursor on the screen. Shourok's speech, posture, gesture and proxemic behaviour remain unchanged. The most prevalent mode is perhaps his silence. The picture at 01:35 shows Shourok leant backwards, arms crossed, closed to the interaction while Darras's posture is open.

Extract 6-15 repeated (01:22 – 01:49)

Ref:		Seconds:	01:22	01:24	01:26	01:30	01:35	01:40	01:46	01:47
169.	Darras	talk	save	and here open↑		this↑	what is this↑	greenpeace	and↑	it should be what↑
170.		actions					☺ smiles briefly		moves himself central	
171.		keyboard								
172.		screen	🗁 file	⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩⇩⇧⇧⇩ opens the file greenpeace						
173.		mouse	🖱----------							----------🖱
174.		gaze	❑----------				❑🗣------🗣❑----------		❑🕮----------	🕮
Ref:		Seconds:				01:28		01:37		01:49
175.	Shourok	talk						(greenpeace)		information
176.		actions						☺ smiles briefly at Darras	looks to his left	
177.		keyboard								
178.		screen								
179.		mouse								
180.		gaze	❑----------			❑😐---😐❑----------		❑🗣------🗣❑----------		❑😐-yawns-😐🕮-🕮

At Ref: 175(180), on the third attempt by Darras to get confirmation of the correct master file to use, Shourok turns to him and smiles briefly and says ‘(greenpeace).’ This is a best guess at transcription because it is unclear. At Ref: 169(01:40) Darras says ‘greenpeace’ and double-clicks the page to open it; Ref: 172(01:40).

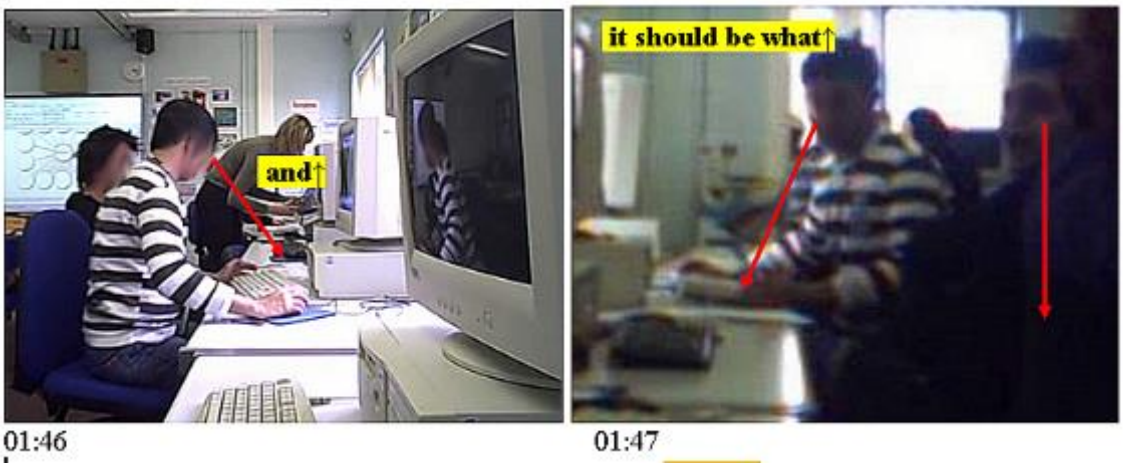


Figure 6-18 (01:46 – 01:47)

At 01:46 Darras pulls his chair to the left until it is central with the screen and he now has a dominant position in the work-space, and subsequently establishing for Shourok the peripheral position he took when he 'handed over' the computer to Darras. Darras says 'and↑' with rising intonation as he looks at the worksheets to find what he needs to do with the page he has opened. At 01:47 Shourok continues to sit with arms crossed and yawns as he looks at the Polish learners to his left. Darras tries again at Ref: 169(01:47) and says 'it should be what. ↑' Shourok finally confirms the new page he has opened should be called 'information.'

Posture is essentially about body positioning which may be open or closed and also includes postural direction (Dittman, 1987). The closed position of Shourok with his arms crossed and body leaning back into his chair throughout the extract so far, would suggest that he has cut himself off from further interaction. This is confirmed by other modes: his distanced proxemic position and verbal silence in response to Darras. With hands closed within the confines of his body, Shourok is unable to gesture either. These individual modes are powerful indicators that Shourok is not happy with the new configuration and is less than a willing participant. By contrast Darras's postural direction is focussed on the task.

This is further confirmed in the next transcript of their conversation; Transcript 6-10. Darras asks at 02:12 'what do we have for information' and Shourok's response is 'nothing' at 02:15, his tone flat and serious. Darras laughs at 02:16 and says 'th(h)at's very dreadful.'

Transcript 6-10 (02:12 – 02:47)

02:12	Darras	and↑ (1) what do we have for information
02:15	Shourok	nothing
02:16	Darras	th(h)at's very dreadful what can we write find something on the internet yes↑
02:27	Shourok	there is information ()
02:31	Darras	()
02:40	Shourok	open information
02:47	Darras	hand me

Figure 6-19 (02:16-02:17)

Darras laughs, lifts his left hand and gestures inwards with his finger, encouraging Shourok to join in. Darras says 'th(h)at's very dreadful what can we write find something on the internet yes↑' Shourok replies with a similar deictic gesture at 02:17 with 'there is information;' i.e. what you need to know is in the worksheets. Darras looks at the worksheets and appears unsure. He reads the instructions aloud to himself at 02:31; it is unclear what he says.

Figure 6-20 (02:40 – 02:47)

At 02:40 Shourok leans forward and says ‘open information.’ Darras releases the mouse and Shourok pulls the keyboard towards him, reaches across for the mouse with his right-hand, pushes his chair back into a central position opposite the screen and Darras is pushed back into his previous peripheral position to the right of the computer. Darras spent less than two minutes in control of the computer before Shourok ‘ousted’ him. At 02:47 Darras tries to reach across for the worksheets saying ‘hand me’ but Shourok does not move and Darras returns to the periphery, devoid of any tools. As previously discussed, a key finding here is the *competitive interaction* of the learners rather than *collaborative interaction*. At times their interaction is akin to a ‘battle’. Temporal mapping is co-constructed between the learners but rather than *collaborative co-construction* of meaning there is *competitive co-construction*; evidenced by the ongoing ‘battle’ for tools and spatial control.

6.7 conclusion

In peer-interaction, as teachers and educators, we would hope to see cooperative semiosis where the meaning-making process is an equal collaboration within the wider environment of the classroom. How communication technologies affect social interactions in the classroom, with differing modal alignments, are pedagogical strands for consideration in the next chapter. It was demonstrated how Shourok dominated the computer space and Darras, though peripheral, consistently tried to engage in the design process. From a conversation analysis approach, this chapter revealed that though there was conversational grounding and conventional turn-taking synchronised with gesture and gaze, there was a predominant lack of mutual coherence and alignment. Like the Polish learners, the sequential temporal mapping of off-screen talk to on-screen text was evident across syllable, word and chunk level utterances. There were some examples of equal collaboration but more examples of task-inequality with less evidence of cooperation between the two males, compared to the Polish learners at chapter 5.

7 Cross-case analysis

7.1 introduction

The previous two analysis chapters provided a microanalysis of interaction between two pairs of learners. Both pairings had the same task and both completed their writing tasks, but in one pairing there was *collaborative co-construction* and in the second pairing *competitive co-construction*, where in the former the Polish learners worked cooperatively together to complete all tasks but for the Kurdish learners there was less cooperation and more conflict. This chapter will integrate the analysis of the two pairs of learners into a single comparative discussion, collating the findings from the analysis chapters with reference to the literature. There is an accepted orthodoxy of pairing learners in language learning contexts. This chapter will demonstrate what the benefits are to this pedagogy, in relation to second-language collaborative writing, and how such pairings are structured multimodally.

7.2 introduces the term *transmodal talk* as an outcome of the data analysis and identifies the etymological variations in the literature of how others have discussed the movement of meaning from one mode to another. This coining of the term will help to frame the discussion on the coordination of off-screen talk and on-screen text which is understood to have five stages. There is a sequential pattern in how learners transpose talk to text and this section identifies the terms of reference to frame the discussion in sections 7.3 through to 7.6. 7.7 is the conclusion of this discussion chapter.

7.2 transmodal talk

A significant action in both pairs of learners is the sequential temporal mapping of off-screen talk to on-screen text. This section will discuss in further detail, with examples from the Polish and Kurdish learners, the features of that process. A more precise term is required to name the synchronous action of writing down on-screen the spoken words off-screen. *Transcribing* is close but this typically describes the typing of pre-recorded conversation where one can pause and playback; a verbatim transcription is customary. *Amanuensis* names the scribing in real-time of someone else's words (usually because that person is unable to write or type) but again this is typically a verbatim transcription. What the learners in this study are doing, making a transmodal shift of talk to text, is fluid and dynamic and not as static as transcribing and amanuensis.

Kress (2003) uses the term ‘transduction’ to name the process of moving meaning from one mode to another; such as expressing meaning through a picture rather than words. Bezemer & Mavers (2011) talk in similar terms of ‘transmodal redesign’ as does Newfield (2013) when discussing the ‘transmodal moment’. None of these terms however are accurate enough for this study as they refer more precisely to how meaning ‘travels’ from one mode to a completely different mode of semiotic signs. Ulhirová (1994) analysed the linguistic encounter of two engineers carrying out a task with a single computer. She called their language performance at the computer ‘PC talk’. Gardner & Levy (2010) also adopt the term ‘PC talk’ when analysing the conversation of two teenage pupils at a computer. Talk and text, though different semiotic systems, are presumed to approximate each other. The learners in this study are collaborating at a computer. Spoken and written language goes through a textual cycle of transformation. For simplicity and accuracy, the term ‘transmodal talk’ will be used when referring to the fluid process of transposing off-screen dialogue to on-screen text; they both shape and mediate each other through temporal mapping. *PC talk* refers to general discussion at the computer. *Transmodal talk* refers specifically to the verbal drafting of language off-screen before, and as, it appears on-screen as text in a cycle of speaking, typing, editing and re-typing; it is a writing process shared by two or more people in real-time using a shared computer. There is a general sequence to transmodal talk which is here classified through the following five stages:

- cognitive-orientation,
- off-screen drafting,
- on-screen writing,
- off-screen noticing,
- on-screen correcting.

Each of these stages will be discussed in the following sections to understand the coordination of talk and text and the framework of multimodal collaboration.

7.3 cognitive-orientation

This section will discuss how the learners identified what they needed to do. Before off-screen drafting and on-screen writing, both pairs of learners spent a short amount of time understanding what was required of them in the task. This is a typical precursor to any learning activity: understanding the task. In many contexts this would be negotiated verbally prior to starting the task, which could be tutor input to explain or/and learners reading instructions. In this context, all learners were given worksheets with a brief

instruction to the task: ‘For page 4 of your booklet … Write a few sentences about Greenpeace or environmentalism. What do you think?’ There was no tutor input to further clarify so all pairs of learners needed to decide between themselves what to do. *Cognitive-orientation* to the task is an essential part of the learning process. In this research the learners ‘orientate’ themselves differently across the pairs, resulting in inclusionary and exclusionary peer-interaction frameworks. The differences are interesting as the patterning of modal alignment at this precursor stage is indicative of the ensuing interaction across both pairs of learners’ transmodal talk.

Section 5.2 from chapter 5 provided a comprehensive analysis of the Polish learners identifying what they needed to do in the task. They successfully arrived at a shared understanding of what to do. What was evident was a process of *mutual cognitive-orientation* which lasted thirty-two seconds. This yields important findings into cross-modal coordination, including proxemic mirroring, modal alignment and symmetrical configurations. Section 6.2 from chapter 6 provided a full analysis of the Kurdish learners identifying what they needed to do before they started the task. There is little evidence that they arrived at a shared understanding of what to do. There was a process of *divided cognitive-orientation* which lasted sixteen seconds; the evidence for which can be seen in an absence of proxemic mirroring and failed modal alignment. With the Kurdish learners we see *misalignment* made evident by unreciprocated actions, specifically from Shourok to Darras.

At 5.2 Sakia reads the word ‘environmentalism’ on the worksheets, speaks the word aloud, then turns to Gamda and says in Polish ‘zaraz sprawdzimy’ {we’re going to check}. Gamda mirrors her by looking at the worksheets (presumably to find the word) and then turns to return Sakia’s gaze. *Mutual cognitive-orientation* is evidenced through how they modally align themselves: (1.) the interrelation of reciprocal actions and (2.) mirrored actions. (1.) An action does not occur in monomodal isolation but belongs to a larger modal density of cross-modal coordination. (2.) Actions are also copied, i.e. mirrored, by others in the interaction. The Kurdish learners’ interaction with the worksheets and each other looks significantly different. In Section 6.2 Darras leans forwards to tap the student worksheets in front of Shourok and he says ‘page 4.’ The gaze of Shourok is on the worksheets as he quietly reads the instructions to himself. As Darras leans across the table to read the same page, perhaps mirroring Shourok, the hand of Shourok pushes across the space in front of Darras and he holds the mouse.

This movement pushes Darras backwards away from the keyboard and the mouse and Shourok looks at the screen and begins the task. *Divided cognitive-orientation* is evidenced by how they do not align modally.

Sakia			
Polish		gaze	
		worksheets	
English		internet	
	objects	layout	proxemics

Gamda			
English		gaze	
worksheets		translation device	
	objects	layout	proxemics

Table 7-1 (modal density 3)

The modal colouring is used to suggest where some of these actions are causally related and mirrored for the Polish learners. In the yellow colouring, Sakia looks at her worksheets and says in English 'environmentalism'. Gamda mirrors her by looking at her own worksheets. Gamda's interaction with the worksheets is causally related to Sakia's interaction with her own worksheets. In the green colouring, Sakia turns to face Gamda and says in Polish 'We're going to check'. Gamda responds by turning to face Sakia and says 'mhm', evidencing postural and language alignment. In the grey colouring, Sakia looks at the computer and opens a Polish translation website. Gamda mirrors her by opening her electronic translator. The modal colouring visualises the likely interrelation of these modes in cross-modal coordination and mirroring of actions. In these three instances of mirroring we can see the embodiment of cognition. Sakia's uncertainty is expressed externally and Gamda mirrors that. They both then use *cognitive technologies* (Clark, 2001) with language-switching to arrive at shared cognition in the task. This mirroring gives evidence to theories of social mimicry, where individuals imitate others. We might say that *mimicry* is a feature of an inclusionary peer-interaction framework because it allows for modal alignment. Though they both say very little, Gamda's proxemic alignment with Sakia helps to maintain their interaction. All these features of modal alignment give evidence to mutual cognitive-orientation in the task.

Perhaps what we see in the opening seconds for the Kurdish learners, as they try to understand the task, is a lack of alignment because there is no reciprocation from

Shourok. Their participatory framework is structured by differing levels of modal density and asymmetrical configurations compared to the Polish learners.

Shourok			
mouse		worksheets	
	objects	layout	proxemics

Darras			
mouse		worksheets	
language		gesture	
	objects	layout	proxemics

Table 7-2 (modal density 4)

We can see that the modal density for Darras is more complex than Shourok's. He speaks. He gestures. He uses the mouse. The interrelation of modes is highlighted in yellow: for Darras there is a larger modal density of cross-modal coordination. For Shourok his actions occur in isolation. In the yellow modal colouring, Shourok's action with the worksheets is causally related to Darras's action of tapping the worksheets and saying 'page 4.' In this instance he draws his attention to where Darras tapped the page. However, he doesn't reciprocate or mirror any further actions. Two verbal utterances from Darras are ignored by Shourok and he offers no discussion to enable mutual orientation to the task. Darras tries to align posturally with Shourok to read the worksheets but Shourok's actions prevent this. The two tables of modal density reveal that for the Polish learners there is greater *mutual coherence*, or shared cognitive understanding, compared to the Kurdish learners and this is predominantly achieved, or not, through how well learners align multimodally with each other.

'Coherence is the consequence of interaction between the addresser and addressee, which can be achieved by mutual efforts of both communicator and addressee' (Wang & Guo, 2014: 465). *Mutual coherence* is used here to name a situation in which two or more people have a shared understanding. It is achieved through 'mutual efforts' which work collaboratively towards achieving a goal. If learners are paired or grouped together then mutual coherence towards the task must be a primary concern for the teacher. There is evidence for this in the Polish learners' cognitive-orientation to the task. The term *divided coherence* is used to name the opposite of this, where it is unclear if there is any shared understanding. A learner may understand the task at hand and could be said to have cognitive coherence but if she or he is paired with another learner and does not interact then it could hardly be considered mutual learning. The adjective *divided* is used to describe a situation in which there is no evidence that two or

more learners are sharing; be that through language, understanding, tasks or actions. There is evidence of *divided coherence* in the Kurdish learners' cognitive-orientation to the task.

In this opening sequence of cognitive-orientation, the Polish learners' framework of participation looks symmetrical because of the following features:

→ mutual cognitive-orientation is achieved
→ through mirroring behaviour
→ evidencing modal alignment
→ leading to mutual coherence
→ and an inclusionary peer-interaction framework.

When a participant makes meaning of an utterance or text it achieves coherence. The Polish learners both arrive at mutual coherence about the task and we see evidence for this through their ongoing discussion of the translated 'environmentalism' and their discussion of the 'recycling' topic in the first stage of off-screen drafting; discussed in the next section. For the Kurdish learners, their framework of participation looks asymmetrical because of the following features:

→ an absence of mutual cognitive-orientation to the task
→ because of no mirroring behaviour
→ and no reciprocation of language
→ evidencing misalignment
→ leading to divided coherence
→ and an exclusionary interaction framework.

The evidence for divided coherence towards the task is seen through misalignment and absence of discussion which continues into the first stage of off-screen drafting; discussed in the next section. In terms of cognition, Darras's off-screen and on-screen coordination would seem to suggest he is confident in understanding the task and 'leads' them by telling Shourok what to do next and he restructures the immediate tools around them so they can begin. Shourok however might be seen to be 'disempowering' Darras by not aligning with the reciprocal actions of Darras and taking control, rather than sharing, the cognitive tools around them. This includes the mouse and the worksheets which Shourok previously took from Darras and placed on his side of the table. In terms of 'objects,' the Polish learners both use the cognitive technologies available to them to complete the action. They each hold a pen. They each have the

worksheets in front of them. They are both proxemically equidistant from the keyboard and mouse, though Sakia is slightly more central to the screen and keyboard. They both use digital translation tools.

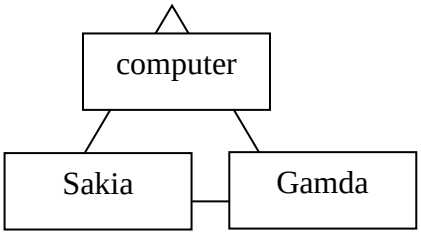


Figure 7-1 (equal triangular layout)

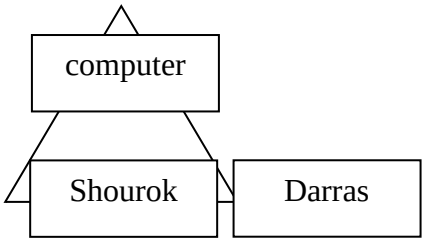


Figure 7-2 (inequal triangular layout)

By comparison, there is evidence of spatial inequality at the triangular layout of the computer, as Shourok pushes across the space between Darras and the mouse, which proxemically marginalises Darras on the periphery of the triangular layout. *Proxemic distance*, close and marginal, between each other and the tools around them and the space to work effectively, is another indicator of inclusionary and exclusionary frameworks for learning.

Conversation analysis provides a framework for exploring reciprocity. From a multimodal perspective, we have seen in this section that *reciprocal actions* can be achieved through the alignment of language, postural direction, tool use, gaze and mimicry/mirroring. Evidence of these in the Polish interaction suggests a *symmetrical configuration* between the learners which would indicate an inclusionary peer-interaction framework. In the Kurdish interaction we see *singular actions* (i.e. a lack of reciprocation) giving evidence of an *asymmetrical configuration*. If learning is essentially about the embodiment and extension of knowledge, we can see evidence of learning and knowledge through a range of actions in the classroom environment. ‘People cognize/learn not just mentally, but in environments comprised of bodies, cognitive tools, social practices, and environmental features’ (Atkinson, 2010: 609). Cognition is embodied in the environment through tools and modes. Here is where we can see evidence of learning and understanding. The absence of these features does not

therefore mean there is a lack of learning and understanding. Rather, there is an absence of collaboration and co-cognition which is not an ideal peer-interaction framework we would want for learners.

The reciprocal actions are identifiable in this section as (1.) language alignment where Gamda responds to Sakia's 'environmentalism with a 'mhm'; but not from Shourok back to Darras when Darras says 'page 4' twice. Reciprocal actions are evident in (2.) mirroring behaviour of postural direction and gaze between the Polish learners but not between the Kurdish learners as their postural direction and gaze are *dissimilar*. Reciprocal actions are evident in (3.) use of cognitive technologies where the Polish learners mirror each other with worksheets and translation tools; less so with the Kurdish learners. *Cognitive-orientation* to the task is an important precursory step in the learning process. In this section we have seen how the learners 'orientate' themselves differently across the pairs, resulting in *mutual* orientation and *divided* orientation to the task.

7.4 off-screen drafting

This section will focus on an extract from the analysis chapters to identify how off-screen drafting is sequenced and structured. The Polish learners draft their first sentence verbally off-screen, agreeing on the general content. The Kurdish learners do not draft their first sentence off-screen because Shourok writes the sentence without any discussion. Focussing on the opening stages of sentence one will allow for the contrasting of multimodal co-construction, co-cognition and the resulting peer-interaction frameworks of both pairs of learners.

Transcript 7-1 (00:46 – 01:15)

00:46 Sakia	like we should more care about recycling (2) if we (4) we want	
01.10 Gamda	want to keep the	
01.11 Sakia	want to keep our planet	
01.14 Gamda	as	[long
01.14 Sakia		[nice and clean
01.15 Gamda	mhm	

Section 5.3 provided detailed analysis of the off-screen drafting stage. Sakia and Gamda speak the above and in doing so facilitate *linguistic equality* in the planning of their text. Linguistic equality is an important characteristic of an inclusionary participation framework; one which recognises the voice and contributions of others. It is

recognisable in their conversation as a form of *echo utterance* or *latching*. Latching in conversation analysis recognises the retrospective turn continuation of two or more speakers with a slight difference from traditional turn-taking in that there is usually no gap in silence between the end of one speaker and the beginning of the next speaker; (Schegloff, 2007). In conversation terms it is anticipatory and important to show reciprocal interest. Echo utterance is the preferred term to describe what is happening here as in the turn-taking of Sakia and Gamda they are echoing each other verbatim. This begs the question why. Echo utterance in this context might be considered a *cognitive baton* to logically and sequentially build sentences off-screen between two or more people before transposing to on-screen. Their completed first sentence on-screen is: '*We should more care about recycling if we want to keep our planet clean and healthy*' and is strikingly close to what they agreed off-screen at Transcript 7-1.

In echo utterance (Tannen, 2005) speakers recognise what the other is saying by mirroring the tail of each other's verbal output. Repeating the final word(s) acts both, as recognition of the other speaker to show active participation, which promotes linguistic equality in the conversation, but also as a cognitive baton to build their first sentence. This would suggest a *linguistic scaffolding* of their text. The modal colouring is used to suggest where these cognitive batons are causally related and support the linguistic scaffolding of the sentence. In the blue colouring there is the verbal baton of 'want' repeated three times, from Sakia (01.06) to Gamda (01.10) and back to Sakia (01.11). In the yellow colouring 'to keep' is copied by Sakia from Gamda. The evidence then of echo utterance suggests that cognition is embodied through language and effectively distributed between the two learners. There is co-construction and co-scaffolding of language off-screen. For both learners there is subsequently mutual coherence and mutual collaboration which would suggest an inclusionary framework for learning.

These linguistic features are accompanied by eye contact and smiling. At 01:14 they overlap each other in conversation and simultaneously turn to face each other as Gamda agrees with Sakia saying 'mhm' at 01:15. Gamda follows this back-channelling 'mhm' with another back-channelling feature at 01.20 by turning to Sakia and smiling. Sakia mirrors her by returning the gaze; it is uncertain if she smiles as her head is turned away from the camera. A smile might be considered a 'back-channelling' technique because as a mode it can communicate agreement with what has previously been said. Humour and smiling could be considered as characteristics of an inclusionary participation

framework because smiles help to strengthen social relationships. In this short interaction of off-screen drafting we find the multimodal framework for learning is structured by the following modal density and both learners figure relatively equal in its co-construction:

Sakia

language		gaze	
gesture	objects	layout	proxemics

Gamda

language		gaze	
		smiling	
gesture	objects	layout	proxemics

Table 7-3 (modal density 5)

The larger the square the greater value, or intensity, the mode has in the interaction. In the framework of participation above, language is given greater relevance in terms of weighting because it structures the other modes of gaze and smiling. In terms of hierarchy, gaze and smiling are subordinate to language. Other tools and modes such as gesture, layout, proxemics and objects are present but relatively dormant in the interaction. These have least value in supporting the outcome of the interaction; which is to verbally draft their first sentence. Language has the greatest value to complete the action and has the following structure in their off-screen drafting:

→ linguistic equality is achieved
→ through echo utterance
→ which builds linguistic scaffolding
→ evidencing shared cognition
→ leading to mutual coherence
→ and an inclusionary interaction framework.

The other modes of gaze and smiling help to further structure a positive framework. It could be argued that *modal alignment* is central to the effectiveness of this peer-interaction framework. There is *linguistic alignment* where they are able to structure their talk through echo utterance where language chunks act as *cognitive batons*. There is *cognitive alignment* where they are able to agree on what to write. There is *proxemic alignment* where their bodies mirror each other through the modes of spatial positioning, turning to each other and smiling. Evidence of alignment is a recurring feature in this framework, as will be further discussed.

By contrast, there is a complete absence of off-screen drafting between the Kurdish learners across all their sentences. In the absence of any off-screen drafting in the Kurdish learners' transmodal talk, the following modal densities can be seen in the beginning of their first sentence:

Shourok			
text		keyboard	
		screen	
	mouse	layout	proxemics

Darras			
proxemics		gaze	
		layout	screen

Table 7-4 (modal density 6)

The coordination above belongs almost exclusively to Shourok as he types: '*I think we can always help our Earth at any time when we wake, getting shower …*' Text is the predominant mode as he writes on-screen. For Shourok cognition is embodied through the keyboard so we can see evidence of his understanding by what he is typing. For him there is greater modal coordination. Darras is almost totally excluded because Shourok starts typing immediately on-screen and for fifty seconds Darras sits watching, rocking side-to-side on his chair. He is noticeable only by his proxemic distance from the 'cognitive tools' which Shourok has control of. There may be some coherence for Darras as he is able to read Shourok's writing but it is a *divided coherence* as neither of them has mutually agreed on the content; he is predominantly a static participant. The absence of spoken language and negotiation results in the following framework:

→ linguistic inequality
→ resulting in no linguistic scaffolding
→ evidencing no shared cognition
→ leading to divided coherence
→ and an exclusionary peer-interaction framework.

This is not to suggest that Darras is wholly excluded from the totality of the learning encounter; participation frameworks, and the modal coordination therein, are in constant flux. Embodied and distributed cognition is fluid between the learners too as will be seen in the later sections.

Transmodal talk is not evident off-screen prior to writing for the Kurdish learners, as it is with the Polish learners, but there is evidence of transmodal talk in the later stages of

their writing. In contrast to the Polish learners, it is possible to say in the off-screen planning stage that a lack of *alignment* is central to the ineffectiveness of the Kurdish learners' peer-interaction framework. Lack of alignment is a recurring feature in the ongoing Kurdish frameworks, as will be further discussed. Their framework is ineffective in the sense that it lacks the features of collaboration seen with the Polish learners. It is still productive as Shourok is writing within the remit of the task; it is however one-sided so far, which is the antithesis of peer-interaction.

Off-screen drafting is an important step in the writing process if it is to enable linguistic equality and polyvocality in the task. In this section we have seen how the learners approach the planning stage differently. For the Polish learners there was evidence of linguistic equality achieved through echo utterance and linguistic scaffolding. This in turn evidences distributed cognition because the linguistic echoing performed as cognitive batons to build their first sentence. There is further evidence for cognitive batons as a function for collaborative writing in the Polish learners' construction of the second sentence; section 5.6 (07:56-08:27). It could be said with confidence that for the Polish learners there is mutual coherence and proxemic alignment. By comparison, for the Kurdish learners in the same off-screen drafting stage of transmodal talk, there was linguistic inequality and cognitive misalignment leading to divided coherence and proxemic misalignment. Shourok was still able to continue with the task but his learning partner was peripheral to the point of exclusion; there was no collaboration in their off-screen drafting stage. With the Polish learners it is clear how knowledge is embodied and distributed between them both through 'interpersonal interactions;' less so with the Kurdish learners. However, analysing the dialogues of pairs in pen-and-paper collaborative writing, Storch observes 'that although all the pairs spent some time on the planning phase, that phase in most cases was very brief (about 4–20 turns, taking up approximately 1 min on average). Most of the time was spent on the writing phase' (2005: 163). This would suggest some correlation with my findings as the drafting/planning stage also had the shortest duration.

7.5 on-screen writing

The next stage in the sequential process of transmodal talk is *on-screen writing*. As discussed, Shourok moved straight to this stage without any off-screen negotiation of what the wording might be. This section will provide extracts from the analysis chapters to identify the patterning of talk and typing, including cross-modal coordination, co-

cognition and lexical mutation as the students write their sentences on-screen. A central finding here is the *temporal mapping* of language from off-screen verbal to on-screen text. This finding is similar to Gardner & Levy: 'We find that speaking is slowed down, in order to map the talk onto the typing; participants start to speak most words slightly before typing them' (2010: 15). It has common features across both pairs of learners which will be discussed in this section. Evidence for the temporal sequencing of transmodal talk was seen in the analysis chapters: sections 5.4 and 5.6 for the Polish learners and sections 6.4 and 6.5 for the Kurdish learners. It is the findings from these sections which will be discussed next.

In the on-screen writing stage, the synchrony and coordination of transmodal talk is frequently slowed down as identified by Gardner & Levy, 2010. However, there are a number of significant causes to the temporal slowing down of language not identified by Gardner & Levy. Speech is mediated by technological tools when language is rendered into literacy; that is, spoken language is comparatively quicker in performance as it is not constrained by spelling accuracy, punctuation marks, letter case and the neat, linear, left-to-right grammatical logic of type. When typing on-screen, the writing is constrained by these features of written discourse so the customary flow of spoken language is slowed down and fragmented, syllabified and sounds repeated for a number of reasons. Off-screen language is slowed down to aid on-screen typing, including pockets of silence to allow for problem-solving, and is performed phonologically at *phoneme* level, *syllable* level, *lexeme* level and *lexical chunk* level between one or two learners.

- phoneme scaffolding: to build a word on-screen (graphemes) from individual letters and letter blends spoken off-screen (phonemes);
- syllable scaffolding: to build a word on-screen from utterance units off-screen;
- lexeme scaffolding: to build sentences on-screen from individual words off-screen;
- lexical scaffolding: to build sentences on-screen from chunks or collocations off-screen; that is, lexical units (or lexemes) were stringed into lexical bundles, which were then typed.

Scaffolding, as a metaphor for learning between two learners is used here in the same sense that Donato (1994) refers to 'collective scaffolding' in second-language peer-

interaction. Regardless of symmetrical and asymmetrical language proficiency, learners can still help each other at a number of linguistic levels.

At *phoneme* level, the typed words were constructed from the speaking of fragmented sounds to map the off-screen phoneme to the on-screen *grapheme*, including digraphs and trigraphs. When talk was broken down to the speaking of individual phonemes, this was usually to help with the spelling of difficult words, one letter or blend at a time. The accuracy of linguistic form on-screen was of priority to the learners. Whilst the Latin alphabet has 26 graphemes in script, the English language has 44 phonemes which can cause difficulty when mapping sounds to script, especially when a single sound might be represented by a blend of individual phonemes such as digraphs and trigraphs. The problem of phoneme-grapheme correspondence is magnified in second-language writing contexts. In the following discussion, phonemes will be notated with a forward slash /a/, graphemes with chevrons, <a> and phonetic equivalents of graphemes (such as hard /a/) with brackets, [a].

At *syllable* level, the learners' typed words were constructed from off-screen utterance units, including morphemes. Syllables were used as phonological building blocks for difficult words to spell. *Syllable scaffolding* is used in this context to describe how transmodal talk is slowed down to utterance units to help with the typing of problematic words. Off-screen, a word was syllabified to sound out possible spellings which were then tested on-screen for visual accuracy.

At *lexeme* level, words were sometimes spoken slowly and punctuated with silence between each word. At other times the word was problematic and needed resolving so again talk was punctuated. *Lexeme scaffolding* is used in this context to describe how transmodal talk is slowed down to word level to allow the typist time to write on-screen or/and cognitively resolve difficulties with meaning. Off-screen, words are repeated, dropped and even morph into variations of the same word: i.e. part of the same lexeme family or synonyms. It is a very fluid process which belies the frozen end-product of completed on-screen sentences. At *lexical chunk* level, spoken blocks are constructed from lexical units (lexemes); i.e. words. *Lexical chunk scaffolding* is used in this context to describe how transmodal talk is slowed down to help the typist type the words as they are spoken as small bundles of words with pausing after each verbalisation. At both levels there is a mutation of words and chunks from off-screen composition to on-screen

typing. This mutation is evidence of 'collective scaffolding' and a positive indication that learners are working to improve their language performance.

The four features of temporal linguistic scaffolding, which identify the organic movement of spoken language to the written, are performed at two levels:

- *monadic temporal mapping*: when a typist speaks as she or he is typing;
- *dyadic temporal mapping*: when the typist and non-typist are speaking and the typist functions as scribe and editor.

There was no evidence from the analysis of triadic-temporal mapping; i.e. a third party such as a teacher, LSA, or another learner involved. In these two instances of mapping, spoken language is slowed down to facilitate the slower typing or is slowed down because of uncertainty with form or/and meaning. Monadic temporal mapping is of less interest as dyadic as it was not a prevalent feature of transmodal talk. Typically, the typist would speak at word or chunk level, and then pause speaking, as he or she typed the words which would appear on screen a second or two later; as in the following examples from section 6.4 when Shourok is typing.

Extract 7-1 (monadic temporal mapping of lexemes)

Ref:		Seconds:	04:44	04:51	04:52	04:53	04:54	04:55	04:56	04:57	04:58	04:59
1.	Shourok	talk	gas				<petrol>					
2.		actions										
3.		keyboard		,	gas	,	p	etr	ol		an	d
4.		screen		,	gas	,	p	etr	ol		an	d

Extract 7-2 (monadic temporal mapping of lexical chunks)

Ref:		Seconds:	05:00	05:01	05:02	05:03	05:04	05:05	05:06	05:07	05:08	05:10
5.	Shourok	talk	and many things									
6.		actions										
7.		keyboard	man		y		thin	ks	ar	ound	us	
8.		screen	man		y		thin	ks	ar	ound	us	

At Ref: 1(04:44) Shourok says the word ‘gas’ which appears on-screen a few seconds later at Ref: 4(04:51). There is then a pause of seven seconds before the next word ‘petrol’ is spoken at Ref: 1(04:54). The lexical chunk ‘and many things’ at Ref: 5(05.00) takes less than a second to say but five seconds to type and appears on screen at Ref: 7-8(05.00-05:05). Pausing like this between words and chunks possibly serves two purposes: (1.) to allow time to type each word; and (2.) is a consequence of the cognitive process of thinking what the next word in the sequence might be. It is an obvious feature of script that when a person is writing there might be spoken, drafting composition before anything is written. The inevitable time-lag of the two does not merit extensive discussion. Dyadic temporal mapping, by comparison, is of significant interest as it is here where polyvocality, co-cognition and the potential for language development are found. To understand the fusion of language, literacy and technology in collaborative writing, the following discussion will offer some examples from the analysis chapters on phoneme, syllable, lexeme and lexical chunk scaffolding as a feature of dyadic temporal mapping in transmodal talk.

7.5.1 phoneme and syllable scaffolding in dyadic temporal mapping

In the analysis chapters it was shown how both pairs of learners map sounds to graphemes as a language strategy for spelling words. Two examples are included below.

Transcript 7-2 (from section 5.5)

03:14	Sakia	s e pa
03:15	Gamda	ra sepa <separately>
03:18	Sakia	separately nie wiem

Transcript 7-3 (from section 6.4)

03:56	Darras	elec		
03:57	Shourok	lec(h)tri		
03:58	Darras	e:lec: tri	(3)	r

The waveforms below show the learner's vocalisations on a timeline from Audacity with each utterance mapped to the soundwave peaks and valleys, showing amplification, frequency and duration of each utterance. The soundwaves can help to visualise how spoken language is temporally slowed down to aid with spelling, understanding and the mapping of off-screen talk (phonemes) to on-screen text (graphemes). At 02:50 Gamda suggested the start of their sentence could be: 'a few seconds to sort them out separately'. The speaking of 'separately' by Gamda, but as a soundwave at Figure 7-3, shows that it is vocalised in under a second. The visualisation of soundwaves can help to illustrate the duration of each spoken phoneme and syllable and the duration of the complete word spoken at normal speed.

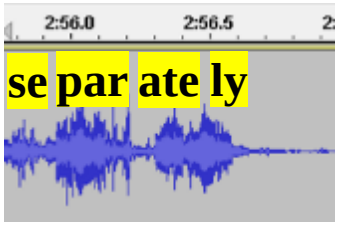


Figure 7-3 (waveform of 'separately' spoken by Gamda)

However, the timescale of the same word vocalised from 03:14 to 03:18 looks very different. It was relatively easy to say the word and this is revealed through how quickly it was spoken. Mapping the sounds to the correct on-screen graphemes is less easy and this is seen below when talk becomes fragmented and stretched. Sakia's talk is in red

and Gamda's is in black at Figure 7-4. The typed letters are in black beneath the soundwaves.

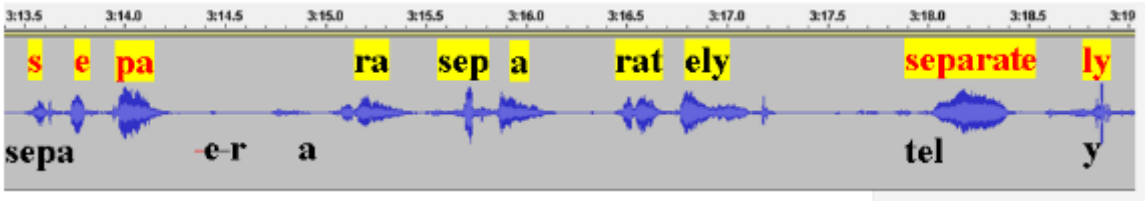
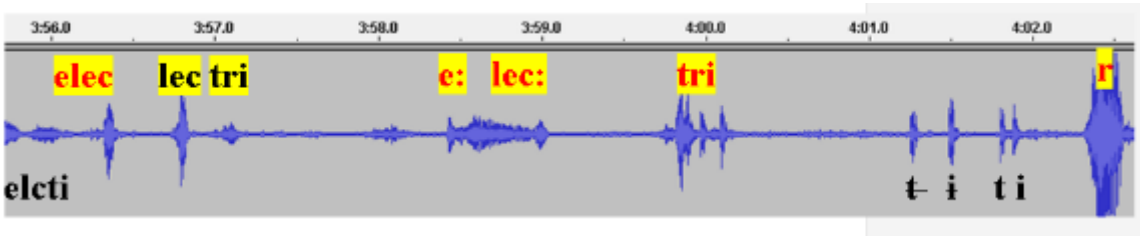


Figure 7-4 (waveform of 'separately' spoken by Sakia and Gamda)

Sakia's on-screen typing of the word pauses at '*sepa*' and she then pronounces the phonemes and syllables slowly from 03:13.5: /s/ /e/ /pa/. There is a one second pause at 03:14 on /pa/ and then Gamda takes over the pronunciation of the next syllable at 03:15 to help with the syllabic and phonemic scaffolding of spelling the word on-screen. Gamda goes back to the first syllable 'sep' at 03:15.5 and breaks the entire word apart, dragging the vocalisation of the word across two seconds; compared to her normal one second vocalisation of the word at Figure 7-3. This temporal slowing down of language occurs for the purpose of trying to work out how to correctly map the phonemes to graphemes. At 03:18 Sakia says the word much more quickly, 'separate', as can be seen from the waveform duration which now has only a single peak of amplification as opposed to the peaks and valleys of slow syllabification seen at 03:15.5. This increased speed of articulation would suggest cognitive acceptance of grapheme form. There is the slightest of pauses before the final syllable 'ly' and this corresponds to the simultaneous typing of the syllable as she says it; the vocalisation is synchronised with the final off-screen phoneme to the on-screen grapheme.

The Kurdish learners use a very similar language strategy of temporally slowing down language to a phonemic level to assist with the on-screen spelling of graphemes. Shourok's talk is in black and Darras's is in red. The typed letters are in black beneath the soundwaves at `Figure 7-5.



`Figure 7-5 (waveform of 'electricity' spoken by Shourok and Darras)

Shourok types '*elcti*' and then pauses, as Sakia did with '*sepa*'. Darras responds by pronouncing the first two syllables 'elec' (03:56) - adding the /e/ vowel which is missing from the on-screen text. Shourok echoes the final syllable of Darras's 'elec' and adds the third syllable of the word: 'lec(h)tri.' At the same time he laughs as he speaks the syllables and turns to face Darras. At 03:58.5, Darras stresses the phoneme /e/ followed by the syllable 'lec' and the phoneme /tr/, pronouncing them slowly over two seconds. Shourok deletes some letters and starts again. Darras then identifies the absent grapheme <r> on-screen which Shourok is missing and says loudly, as seen by the high peak at 04:02.5, the phoneme /r/. The challenge for Shourok here is possible uncertainty with the phoneme /tr/ and its corresponding digraph <tr>. They can articulate the sound but visually the on-screen grapheme is wrong. After sounding out the graphemes and syllables for six seconds it is Darras who finally realises the missing <r>. The difficulty of mapping sounds to letters can be seen in other examples. Shourok misspells 'electricit/e/', replacing the correct grapheme <y> with the phoneme /e/. It is phonologically correct but orthographically wrong. Sakia perhaps does the same when she writes 'cleen' instead of 'clean'. Phonologically, one can hear the long vowel phoneme /ē/ as in 'seen' and 'been' rather than the learned grapheme form of /ea/ as in 'seat' and 'meat'; same phoneme, different graphemes. In each instance it is the red underline feature of the technology which tells them they are wrong; something which is clearly not available in pen-and-paper format.

From this discussion, it can be seen there is evidence of co-cognition towards accuracy of language form between both pairs of learners. 'Writing is more "attended to" than speech, i.e. we are more conscious of what we are doing and tend to attach more importance to correctness' (Brookes & Grundy, 2002: 1-2). Writing, in one sense, is the visual permanence of the spoken and in this research there was collaboration by the learners towards the 'correctness' of form. The evidence from the analysis would suggest, as Brookes & Grundy state, that writing encourages learners to focus on accuracy in a way that spoken output might not. *Noticing* errors ('a hypothesis that input does not become intake for language learning unless it is noticed, that is, consciously registered' [Schmidt, 2010: 2]) is an important function demonstrated by all learners and is discussed in section 7.7. For educators, putting learners together to compose a joint piece of text can have positive outcomes for language development. What was consistent in the analysis chapters is that learners in general pause when writing a

difficult word; they notice a problem. This pause is an embodied indication of cognitive uncertainty. Working in pairs, the second learner is able to respond to that pause and between them share the cognitive load to resolve the problem. To do this they use echo utterance at the smallest phonemic level to scaffold sounds and work collaboratively towards spelling the word. The dyadic temporal mapping of phonemes and syllables to graphemes and ultimately words is a successful language strategy for accurate spelling. The Kurdish learners are not wholly successful, compared to the Polish learners, as their word remains incorrect until Darras uses the spellcheck feature of the software.

Learner focus on linguistic form in writing, as a collaborative venture, can offer a linguistic pool of lexical forms to work towards accuracy in a second language. This is the finding of Philp et al reviewing the research on second-language writing: 'Researchers conclude that discussing form in the context of creating a written text may help learners to use new and more accurate language forms' (2014: 160). As educators then, to support improvements in spelling form, teachers could offer more explicit instruction on the range of phonemes in the English language and how each maps to a corresponding grapheme. The learners in this research were level 2 language users so errors in form were minimal but even so they had strategies to deal with difficulties which lower level learners might not have. Teachers could show learners how to tease out graphemes, syllables and morphemes, as the base forms of lexemes, so that students are better able to scaffold sounds into written words. The phoneme /k/ for example appears differently as a grapheme in a range of words: **c**oo**k**, **k**ill, s**ch**ool. Enabling learners to work collaboratively, where the focus is on phoneme-grapheme agreement with pre-selected words, can allow for greater autonomy as learners are able to share a wider pool of linguistic resources. In the context of computers, the auto-spell function for writing tasks can provide added benefits.

7.5.2 lexeme and lexical chunk scaffolding in dyadic temporal mapping

In this section the focus is on word and chunking activity within transmodal talk, but examining (A.) the mutation of talk to text and (B.) the polyvocality of that process as a co-coordinated and co-cognitive, translanguaging interaction. This interaction can be identified at the level of lexemes and lexical bundles where words and chunks are scaffolded, but also go through a process of lexical mutation, to make sentences. In the first sentence of the Polish learners at Transcript 5-5, linguistic scaffolding of a sentence

is performed by two learners who sequentially contribute words and chunks as cognitive batons between each other: e.g. ‘want’ and ‘to keep’ and ‘nice and clean.’

- K: we should more care about recycling
- K: if we
- K: we want
- M: want to keep ~~the~~
- K: want to keep our planet
- M: ~~as long~~
- K: ~~nice and clean~~
- K: clean and healthy

The strikethrough words are dropped or mutate into other words on-screen. The off-screen lexical bundle ‘nice and clean’ is changed on-screen to ‘clean and healthy’ in Extract 5-8, but Sakia checks off-screen with Gamda to see if she agrees with that change before writing it. The change is arguably an improvement hence the suggestion. Rejected words include ‘the’ and ‘as long’ spoken by Gamda, and ‘nice and clean.’ In the findings of Storch (2005: 165) she also identifies the process of co-construction through cognitive batons; though she might not use that term: ‘Most of the pairs collaborated in the creation of the text by completing each other’s ideas, offering alternative suggestions, and feedback. These are illustrated in the excerpts that follow.

246 O: so do we have to write a concluding sentence?
247 S: yes concluding sentence
248 O: as a result . . .
249 S: as a result they improved their English
250 O: or their English is . . . their English level is increased after they
251 S: came . . .
252 O: yeah after they came to . . . after they arrived in Australia.’

Each student repeats a word or a chunk of the previous speaker and so builds, selects and deselects language to co-construct a sentence. There are similar instances of lexical change in sentence one of the Kurdish learners but no evidence of scaffolding and linguistic equality. In Transcript 6-3 Shourok pauses after the chunk ‘getting shower’ and this pause allows Darras to question what he means. Darras tries to offer an alternative verb: ‘take (1) take shower.’ This is a more appropriate verb form as ‘to have’ or ‘to take’ a shower is more commonplace. Shourok language-switches, perhaps to echo Darras’s verb form of ‘take,’ and says: بگیری! - in translation ‘take a shower.’ He then deletes ‘getting shower’ and replaces it with ‘knowing how to use electricity,’

ignoring Darras's chunk, which in comparison to the Polish learners, raises questions about polyvocal and monovocal writing in peer-interaction.

In the second sentence of the Kurdish learners, Transcript 6-7, the non-typist provides the full sentence 'we should be responsible for recycle our bins' but then has to fragment the sentence, repeat words and slow down his talking to synchronise with the on-screen typing of Shourok. If we break the language down to lexeme and lexical chunk level, where there are natural pauses between each utterance, it is easier to visually see how the non-typist scaffolds the language to support the typist. Off-screen, Darras says:

- I think that we
- we should be responsible for recycle our bins
- recycling
- recycling our bins
- it's for recycling
- our <bins>
- <recycling>
- <recycling>
- <bins>
- if you change this to recycle
- recycle our bins.

In real-time it is difficult to see, and appreciate the complexity, of how talk, text and action are interwoven in cross-modal and co-cognitive coordination. From a researcher perspective, videoing on-screen actions and off-screen interactions, with learner audio, and being able to slow down the playback of their behaviour for analysis, has enabled an insight into multimodal peer-interaction which would be missed by educators in the everyday teaching and learning of their classrooms. Darras works to synchronise with the actions of Shourok and we see the evidence for that in the mediation of his language, fragmenting speech down to word and chunk level to help Shourok temporally map sounds onto graphemes. There is cognitive contiguity between the learners as evidenced by their successful correlation of phonological and graphological interaction. This is not always the case as Shourok frequently ignores the suggestions of Darras, or chooses not to involve him. Distributed cognition is evidenced in this instance through the words of one and the corresponding actions of the other, suggesting a positive peer-interaction framework. The on-screen text has only one

change from what was originally spoken. Shourok adds the present continuous 'ing' to Darras's imprecise use of 'recycle:' *'We should be responsible for recycling our bins.'*

Sentence two of the Polish learners (sections 5.5 and 5.6) goes through the most significant lexical mutation. The language chunks off-screen undergo a number of on-screen changes from different sources; colour-coded as Wikipedia, Sakia and Gamda in the finished on-screen writing:

It takes only few minutes to put the litters in order and prevent the Earth from pollution.

Each learner retains something of the original word sequence but adds something new, scaffolding lexemes and lexical chunks between them as cognitive batons. They also draw on synonyms in their first-language 'zaśmiecenie {littering} degradacja {degradation}' before settling on 'pollution' which they find on the internet.

- M: it ~~like~~
- M: only takes
- M: ~~a~~ few ~~seconds~~ to ~~sort them out separately~~
- M: and
- K: it takes only few minutes
- M: to put the litters in order
- K: and
- Both: prevent
- K: the earth from degradation
- M: from
- K: from pollution

In the discussion so far it can be seen that the sequential mapping of talk to text follows a time-lagged and iterative arrangement using phonemes, syllables, lexemes and lexical chunk utterances. Talk and writing are fluid and non-linear as language goes forwards, is repaired and repeated, rephrased, written only to be deleted and go backwards before going forwards again. The chronemic configuration of language mediated by actions with tools is startling. Language is organic and unidirectional. The on-screen text is also a composite of different voices. Dyadic temporal mapping is therefore achieved between the typist and the non-typist through the following features:

- pausing between utterances to give the typist time to write;
- pausing to allow for the cognitive processing of what the next sequence of words might be;

- dragging out the sounds of syllables and words as the typist types;
- repetition of words and lexical chunks;
- using cataphoric references to temporally jump forwards to what the next words or chunks might be;
- using anaphoric references to temporally jump backwards to previously spoken words and chunks;
- stressing phonemic sounds to raise attention towards grapheme inaccuracies;
- lexical mutation and language experimentation to improve the accuracy of on-screen text;
- polyvocal expression in on-screen text collated from different sources;
- translanguaging performance, typically resorting back to a first-language to cognitively resolve an issue with the second language.

The mutability of language in the off-screen and on-screen dynamic is an outcome of polyvocal practice and experimentation as the learners use a range of 'vocabulary strategies' (McDonough, 1995) to find the most accurate lexemes and lexical chunks for their second-language writing. This plurality of voices gives evidence to the polyvocality of transmodal talk. Learners may draw on a range of linguistic resources to find and make meaning, including each other and the internet as corpus. Peer-interaction for writing, with access to digital, bilingual literacies on the internet, might thus be considered a language strategy to 'stretch and challenge' the learners' own vocabulary; using 'another's speech in another's language' (Bakhtin, 1981).

7.6 noticing and correcting

Editing is about changing the on-screen appearance of the learners' designs, be that language, layout, size, appearance, colours, etc. The editing process might be discussed between the learners or initiated individually. Instances of editing in the analysis chapters suggest there are two elements. (1.) *Noticing* is the students reading (silently or aloud) what has been written so far to identify any on-screen errors. (2.) *Correcting* is about making changes to try and be more accurate. Noticing and correcting are part of the same editing cycle. Noticing is essentially about accuracy; checking for mistakes and looking for improvements, whilst correcting implements change. In second-language learning contexts, the terms 'attention' and 'noticing' (Schmidt, 2001, 2010) are frequently used to describe a learner identifying errors, comparing target input against their own output, as a catalyst for making changes in their own language. These terms sit within the wider debate on comprehensible *input* (Krashen, 1985) and *output* (Swain, 1985) and *negotiation* (Long, 1996) which will be discussed in this section. Correcting is well recognised within research pedagogy as a relationship between 'expert' and 'novice' in such models as Bruner's *scaffolding* (1978) and Vygotsky's

zone of proximal development (1978). Section 2.4 problematised this pedagogy as it is usually applied to adult learners who are actually peers and symmetrically proficient in language and skills, not asymmetrical. When discussing scaffolding in this research I mean peer-scaffolding between learners of similar ability. As Dobao argues: 'Same-level learners can provide scaffolded assistance to each other and, pooling their individual knowledge and resources, achieve a level of performance that is beyond their individual level of competence' (2014b: 498). Correcting through peer-negotiation is the focus of this section, how the learners as peers *notice* problems and go about correcting them through *languaging* (Swain, 2006) and *translanguaging* (García, 2007) negotiations.

Identifying language-related episodes (LREs) is a process for measuring and analysing the collaborative dialogue of learners as they metacognitively think about the language they are using (Swain & Lapkin, 1998). Identifying when learners notice a difference between their actual output and potential output, that there may be errors in lexis, form, mechanics, is one way for any analysis to focus on significant moments of interaction (e.g. McDonough & Sunitham, 2009; Dobao, 2012; Amirkhiz et al, 2013 and Ajmi & Ali, 2014.) Noticing possible and actual errors leads to languaging and translanguaging negotiations, which from an analytical perspective, are identified as language-related episodes.

The context of F2FCCW is slightly different to the usual language environments in which *noticing* is explored in classrooms, however, as will be discussed, the process is very similar: identify errors in one's own language output and change them. There is evidence of synchronous and asynchronous *noticing*. Sometimes correcting happens *synchronously* with the writing so that errors are identified and rectified in real-time with the typing. Here the typist might see his or her errors on-screen, or the non-typist sees and explains, and the writing is corrected before progressing with the rest of the text; as seen in the previous section discussing phoneme and syllable scaffolding. Sometimes noticing happens *asynchronously* when a sentence(s) is complete, read back and the typist or non-typist discusses accuracy and any potential changes to the text. These two time differences of the same noticing process of transmodal talk can be further sub-divided.

- Checking for accuracy can occur synchronously through the monadic temporal mapping process as the typist is typing.

- It can occur synchronously through dyadic temporal mapping between the typist and the non-typist as the typist is typing.

- Checking for accuracy can occur asynchronously through the monadic temporal mapping process when the typist has completed the sentence or a text and he or she individually checks it by reading silently or aloud.

- It can occur asynchronously through the dyadic temporal mapping process between typist and non-typist when a sentence or a text is completed and the typist and non-typist might read the sentence back, either silently or aloud to check for errors.

There are then four dimensions to correcting in F2FCCW. In addition, evidence from the analysis chapters for noticing errors in form, and the subsequent negotiation and co-cognition of writing, could be categorised as language-orientated (or language-related episodes) and typography-orientated; that is, off-screen talk about on-screen text, be it synchronous or asynchronous, tends to focus on (A.) the accuracy of the language and (B.) improvements in the visual layout of the language. The examples below are drawn only from the analysis chapters. The full transcript of the learners' interaction across the entirety of the lesson provides a plethora of language and typography-orientated exchanges; where one *notices* an error and they work together to correct or improve the issue. The timings below are from the transcripts in the analysis introductions at chapters 5 and 6.

(A.) language-related episodes:

- focus on mechanics: uncertainty about the spelling or accuracy of a word
    - Kurdish: 03:56, 05:05
    - Polish: 00:01, 03:15
- focus on mechanics: punctuation uncertainty with commas and full-stops
    - Kurdish: 04:41
    - Polish: 02:24
- focus on form and lexis: uncertainty about word order and accuracy of meaning
    - Kurdish: 03:12, 03:25, 07:40
    - Polish: 02:39, 03:18, 08:42

(B.) typography-orientated:

- focus on font size, style, alignment
    - Kurdish: 01:09, 01:21, 06:33
    - Polish: 01:28, 01:35

Given the focus of this research is on language and peer-interaction, I shall not discuss typography-orientated editing nor monadic temporal mapping. These are the features of an individual working alone to improve the visual layout of language. Instead, I want to discuss the findings of language-orientated negotiations (i.e. languaging and translanguaging) through the process of dyadic temporal mapping and relate this to the long-standing debate on *input* and *output* hypotheses in language education; but from a peer-interaction perspective in line with the *Interaction Approach* (Long 1996). In brief, Krashen (1985) theorised that successful language acquisition was an outcome of comprehensible input +1. By providing an input of language slightly above current level, an individual acquired language. However, Swain (1985) noted that even after years of language input, an individual's output could still have significant errors in grammar. More focus on comprehensible output was needed. Production of language can encourage learners into deeper awareness of structures and forms of language and thereby help them to make comparisons between the target language and their current language proficiency. Long's Interaction Approach (1996) is a hypothesis on second-language acquisition which posits that negotiations for meaning are stimulated by communication problems, and through collaboration with others, problems may be resolved and new language forms and rules acquired.

The negotiation of writing in the analysis chapters suggests that this is exactly what the learners were doing. Significant focus was given to the accuracy of language output and negotiation of meaning. Swain's (1985) challenge to Krashen's *comprehensible input* (1985) was later revised (1995) to include three features of *comprehensible output*. Philp et al revisit those three features and summarise them. A focus on output 'promotes (a) a noticing function (b) a hypothesis-testing function and (c) a metalinguistic awareness function' (2014: 20-21). These three features of output are here discussed in relation to the learners' interaction with writing.

7.6.1 noticing function

Schmidt (2001, 2010) has long theorised the importance of learner-attention towards form, because by being attentive to form, a learner is better placed to notice problems and address them. What writing tasks can do more easily perhaps, which speaking can do but with more difficulty, is focus attention on output; perhaps because the visual permanence of language on-screen encourages accuracy. Here, output becomes input because the learners' own writing becomes the focus of potential intake. A 'learner must

attend to and notice linguistic features of the input that they are exposed to if those forms are to become intake for learning' (Schmidt, 2010: 4). A student's writing can draw attention to their output, and by being attentive, position them cognitively to better notice input issues which in turn can be learned and become intake. Being conscious to one's learning, as opposed to subconscious, is central to Schmidt's hypothesis. The analysis chapters revealed a number of language-related episodes where the learners negotiated the accuracy of output as itemised above at (A). In section 7.6 it was seen how much concentration was given at phoneme, syllable, lexeme and chunk level to the accuracy of form. Noticing problems led to changes at all these lexical levels and what is important in this context is how peer-interaction allowed for greater *noticing* because distributed cognition enabled a wider pool of linguistic forms to draw on. Darras for example notices a problem which is actually a misreading of 'show' instead of 'shower.'

Transcript 7-4 (noticing function of language)

03:12	Darras	what do you mean↑ getting show	← confirmation request
03:14	Shourok	shower shower	← clarification
03:16	Darras	shower haha	← comprehension
03:17	Shourok	shower it's when the ()	← clarification
03:25	Darras	take (1) take shower	← comprehension check
03:27	Shourok	بگیری! {take a shower}	← resolution

Correcting of the problem goes through a process of negotiation: confirmation checks, clarification requests and comprehension. Philp et al (2014: 21) draw attention to similar 'interlocutor signals' when noticing problems and also discuss 'segmentation' of language. We see segmentation above through echo utterance, that 'show' and 'shower' are repeated through their turn-taking which leads to resolution. It might be that Darras's confusion stems from the incorrect verb 'getting' in relation to showering as he goes on to suggest the more appropriate verb form 'take.' Noticing an issue at lexeme level led to an improvement at lexical chunk level. The accumulative effect of Darras's noticing is that Shourok edits the text and refines the language to improve it.

7.6.2 hypothesis testing function

After noticing potential problems in second-language use, Swain suggests there may be a hypothesis process in which a learner will try out and experiment with different forms. A learner may draw on what they already know in their first-language to test out possibilities in the second language. The hypothesis may be accepted or rejected

following experimentation. In this example, the words 'separately' and 'degradation,' 'littering' and 'pollution' go through a sequence of hypothesising between the two learners.

Transcript 7-5 (hypothesis testing function)

02:50 Gamda	to sort them out separately		← hypothesis A presented
02.56 Sakia	aha		← hypothesis A accepted
03:14 Sakia	s e pa		← hypothesis A tested
03:15 Gamda	ra sepa <separately>		← hypothesis A tested
07:56 Sakia	it takes only few minutes		
07:58 Gamda	to put the litters in order	← hypothesis A rejected and B presented	
08:10 Sakia	the earth from degradation (08) from what↑		← hypothesis C presented
08:20 Gamda	from (1) zaśmiecenie	{littering}	← hypothesis D presented
08:27 Sakia	from pollution (08)	jeden odcinek {that's one part}	← hypothesis E presented
08:42 Sakia	zaśmiecenie to ok a degradacja {littering is ok but what about degradation}		← hypothesis C & D tested
08:43 Gamda	mhm		
08:44 Sakia	degradacja	{degradation}	← hypothesis C tested
08:49 Gamda	myśle że to (pollution)	{I think that's it}	← hypothesis E accepted; C & D rejected

In section 7.6 the term *lexical mutation* was introduced to describe the process by which learners experiment with language by hypothesising alternative vocabulary. This is a benefit of peer-interaction. With a teacher, or significant other who is fluent in the target language, there can be less risk-taking with language for fear of being wrong in the eyes of the other. Peers can create a 'safer' environment to experiment with form, lexis, mechanics, and this activity was seen in both pairs of learners. As can be seen from the transcript above, suggestions (or hypotheses) are presented, tested and either rejected or accepted. The challenge for them was to find the most semantically appropriate word for their sentence in English but to do this they swapped to Polish to talk and look online. From 08:10 Sakia presents hypothesis C 'degradation' but she seems unconvinced. Perhaps because of this Gamda presents hypothesis D 'zaśmiecenie' {littering} at 08:20. Hypothesis E 'pollution' is presented by Sakia at 08:27. All hypotheses are then tested before Gamda accepts hypothesis E, and in doing so, rejects the alternatives. At this point Sakia types the words and the task is completed. 'In the case of explicit learning, attended and noticed instances become the basis for explicit hypothesis formation and testing' (Schmidt, 2010: 6). As in the example above, we see opportunities through peer-interaction for explicit, active learning as opposed to

implicit, passive learning. It is an effective process achieved through the co-cognition of lexical forms.

7.6.3 metalinguistic awareness function

The third function in Swain's (1995) output hypothesis is the extent to which learners are able to metacognitively assess their own linguistic output. Language is both the tool for thinking about language and the means through which we learn a language. Languaging is used here to refer to the process through which learners reflect on their language, by using language, to metacognitively think about their language and where appropriate analyse and modify output. 'The verb languaging forces us to understand language as a process rather than as an object' (Swain & Watanabe, 2013: 1; see also García, 2007: xii). 'Translanguaging' is also a verb to describe 'the act performed by bilinguals of accessing different linguistic features … in order to maximize communicative potential' (Garćia, 2009b: 140).

The two terms *languaging* and *translanguaging* are useful for exploring the metalinguistic conversations learners have during collaborative peer-interaction. The evidence from the analysis chapters suggests that the learners are using both strategies. As discussed in transcript 02:50-08:49, the Polish learners depend heavily on translanguaging and use hypothesising to span the two languages. They also use languaging in English to discuss what English language to use, hypothesising words and chunks; some of which are rejected and some accepted so that they linguistically scaffold their sentences together, as seen in section 7.6 and through such words as:

- we can speak about 00:40
- maybe 01:35
- or 02:24
- but 02:42
- we can write is as well 02:42
- like 00:46, 02:42
- from what 08:10

The languaging uses conditionals to make suggestions and offer alternative ideas. Whilst the Kurdish learners also use languaging and translanguaging to negotiate their text, there are less conditionals and more use of direct instructions and imperatives:

- what do you mean↑ 03:12

- بگیری! {take a shower} 03:27
- last one you should put and here 04:41
- no there's new 04:44
- دیگه؟ {finished what else} 06:36
- ننویسیم؟ بنویسیم؟ بیشتر still {shouldn't we still continue writing} 06:36
- if you change this to 07:40

The evidence here is that translanguaging, when used to talk in the first-language to improve writing in the second language, is an effective strategy for language development. This is the finding of other research on bilingual learners: 'learning is maximized when they are allowed and enabled to draw from across all their existing language skills (in two+ languages), rather than being constrained and inhibited from doing so by monolingual instructional assumptions and practices' (Hornberger, 2005: 607). As Hornberger contends, translanguaging does raise questions about the 'English Only' rule. Common practice amongst many language teachers is to 'demand' students use the target language as the only means of communication, thereby reducing reliance on the first-language and not risking cross-contamination.

Reviewing the literature on translanguaging, Creese & Blackledge discuss how: 'moving between languages has traditionally been frowned upon in educational settings, with teachers and students often feeling guilty about its practice. Research shows that codeswitching is rarely institutionally endorsed or pedagogically underpinned' (2010: 103). The evidence from the analysis chapters suggests that the first-language is an indispensable cognitive tool for meta-discussion about the target language and might actually improve the accuracy and range of vocabulary in second-language writing; and vice-versa, by noticing and improving errors in output, feed those improvements back into speaking. A wider discussion is whether or not translanguaging should become pedagogic practice, moving from monolingual historical practice to polylingual in language learning contexts. The review by Creese & Blackledge (2010) suggests that it should. The benefit of translanguaging for collaborative writing tasks was also made by Antón and DiCamilla (1998) – 'Qualitative analysis of five dyads working collaboratively on writing tasks demonstrated how the L1 mediated intersubjectivity and externalization of inner speech (i.e., private speech) during cognitively difficult activities' (Swain & Watanabe, 2013: 4). As in this research, the switch to L1 in the analysis chapters frequently happened during 'cognitively difficult activities' and the translanguaging function worked well to correct problems in the L2.

7.7 conclusion

In this chapter the term *transmodal talk* was identified as a finding to name the process of collaborative spoken language becoming print, phonemes being temporally mapped onto graphemes. The sequencing of that process was structured as cognitive orientation, off-screen drafting, on-screen writing, off-screen noticing, on-screen correcting. The learners used a range of scaffolding techniques at phoneme, syllable, lexeme and chunk level to complete the task; with evidence of languaging and translanguaging. It is conceivable that peer-interaction to collaboratively produce L2 writing with pen and paper would have a similar structure.

The concluding chapter will discuss the peer-interaction frameworks of the learners to identify the structure of learner collaboration. As seen in chapter 7, such frameworks are characterised by inclusionary and exclusionary behaviours evidenced across modal alignment and misalignment. These forms of alignment were evident linguistically, proxemically, spatially and cognitively. Chapter 8 will provide terminology to identify the features of alignment in interaction, comparing Goodwin's five stances of organisation (2007a) to evidence correlation between the features of his participatory framework and the peer-interaction framework presented here. The findings and terminology could support future classroom research into peer-interaction and embodied interaction more broadly.

8 Discussion

8.1 introduction

In the previous section the learners' peer-interaction was broken down to modal level to demonstrate how participation frameworks are structured. The term *transmodal talk* was then used to name the process of coordinating off-screen talk with on-screen writing. The research questions in relation to these two topics of talk and participation are foregrounded here:

- *How are off-screen talk and on-screen text coordinated?*
- *How are peer-interaction frameworks for learning structured?*

To work towards answering the two research questions, this chapter will start by introducing terminology and diagrams on collaborative writing, in section 8.2, and peer-interaction, section 8.3.

8.2 research question 1 (the coordination of talk and text)

This section will serve as a summary towards answering the first research question: *How are off-screen talk and on-screen text coordinated?* In many instances the two pairs of learners used similar strategies to complete their writing tasks; in other instances there were very contrasting differences. These similarities and differences will be discussed in further detail in the next section under peer-interaction frameworks. A key outcome of the research, in relation to question one, is the term *transmodal talk* and the findings of how language, literacy and technology are structured in the multimodal context of peer-interaction in F2FCCW. Table 8-1 collates all the features of transmodal talk discussed in the previous sections to list the terminology used when thinking about the coordination of talk and text. All these terms were first introduced in the analysis chapters and then discussed in more detail in the previous sections 7.3 to 7.6. Based on the discussion so far, the sequentiality of transmodal talk has the following overarching features in Figure 8-1, and Table 8-1 provides further detail. It is sequential in appearance but recursive in execution, temporally mapped and polyvocal:

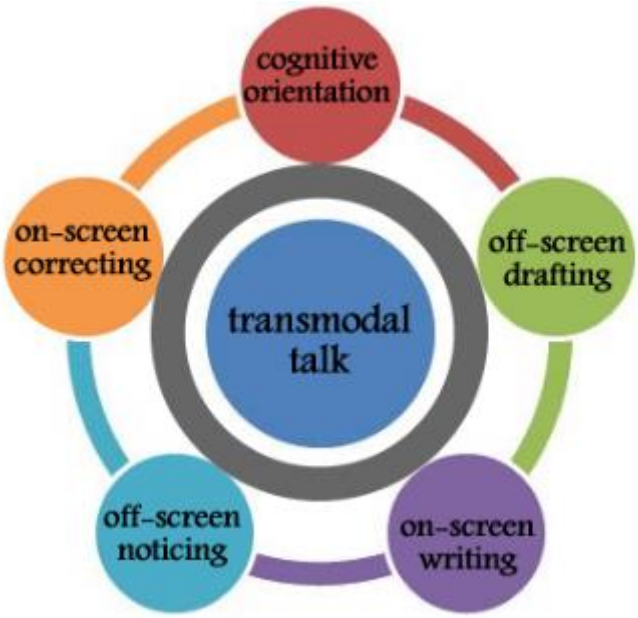


Figure 8-1 (transmodal talk diagram)

Table 8-1 (transmodal talk features)

Transmodal talk	*Features*
cognitive-orientation	a) mutual cognitive-orientation & divided cognitive-orientation b) modal alignment & misalignment c) mutual coherence & divided coherence d) reciprocal actions & singular actions e) symmetrical configurations & asymmetrical configurations
off-screen drafting	f) echo utterance (or latching) g) cognitive batons h) linguistic scaffolding i) polyvocality (linguistic equality & linguistic inequality) j) cognition (alignment & misalignment)
on-screen writing	k) temporal mapping • pausing, slowing speech, syllabification, repetition, stressing sounds l) phonological scaffolding • phoneme scaffolding, with grapheme, digraph, trigraph correlation • syllables and morphemes • lexeme scaffolding • lexical chunk scaffolding (or lexical bundles) m) lexical mutation and experimentation n) process • monadic temporal mapping • dyadic temporal mapping o) voice • polyvocal, monovocal
off-screen noticing ***and*** ***on-screen correcting***	p) process • synchronous • asynchronous q) language-related episodes • focus on form (correct grammatical forms) • focus on mechanics (spelling, punctuation, pronunciation) • focus on lexis (words and meaning) r) typography-orientated • focus on font size, style, alignment s) comprehensible output and input • noticing function • hypothesis testing • metalinguistic awareness o languaging o translanguaging

The precursor stage to transmodal talk involved the learners trying to understand what was required of them in the learning task. It is an important step evident in both pairs of learners and was termed *cognitive-orientation*. There was mutual coherence in this

preliminary stage for the Polish learners who modally aligned to co-cognitively translanguage the instructions to ensure they both understood what was required. For the Kurdish learners there was divided coherence and misalignment. Darras tried to cognitively align with Shourok by explaining what was required but his actions were not reciprocated so both learners had an asymmetrical configuration to the activity which impacted on their continuing collaboration. Darras was excluded because Shourok had control of the cognitive technologies and he could continue without Darras's input.

In *off-screen drafting*, the Polish learners worked successfully to scaffold their first sentence. Linguistically they used echo utterance of each other's words to structure the sentence. We therefore see evidence of distributed cognition because the words functioned like cognitive batons between each learner. The completed sentence was a polyvocal construction which in turn indicated linguistic equality in the task. Shourok skipped this stage altogether. Because of this the beginning of their first sentence could be said to be a monovocal construction, which in turn indicates linguistic inequality in the task; until Darras later adds his voice to the sentence by making suggestions.

In *on-screen writing* both pairs of learners demonstrated a similar temporal mapping process to shift talk to text. Speaking was slowed down, paused, repeated and some words were syllabified and stressed. Mapping phonemes to print required scaffolding at a number of linguistic levels: grapheme, syllable, lexeme and bundles. In discussing communication strategies Gullberg notes that: 'L2 use is characterized by difficulties caused by phonological, lexical, grammatical, and pragmatic shortcomings' (2011: 137). The same 'shortcomings' at different linguistic levels were seen in the learners' 'talk' where they worked collaboratively to resolve and co-construct language. The process through which this happened was either monadic temporal mapping (which was one learner typing without input from the second learner) or dyadic temporal mapping (which was both learners contributing to the writing such as one speaking and one typing). This process evidenced experimentation with form and lexical mutation with some words and chunks being dropped or morphing into alternatives.

In *off-screen noticing* and *on-screen correcting* the students identified errors in their writing. The process through which this occurred was either synchronous (identified in real-time with the writing) or asynchronous (identified after or during a pause in the

writing). *Noticing* ran in conjunction with *on-screen correcting* which involved one of the students deleting, adding and re-typing the text on-screen but this could be accompanied by verbal input from the other learner which resulted in temporal mapping of speech and writing. In the language-related episodes where the leaners focused on form, mechanics and lexis, *languaging* and *translanguaging* were two forms of language strategies used by both pairs of learners across the process.

None of these stages occurred in a neat sequential scale of time because the learners shifted iteratively from writing to noticing to correcting as they proceeded through the task. In answering the question: *How are off-screen talk and on-screen text coordinated?* - it can be seen that the coordination of talk and text has identifiable features across both pairs of learners. Transmodal talk, within the context of F2FCCW, is an outcome of the intersection between language, literacy and technology. Figure 8-1 provided a headline visualisation of the stages of transmodal talk and Table 8-1 listed the features of those stages.

8.3 research question 2 (peer-interaction frameworks)

This section will serve as a summary towards answering the second research question: *How are peer-interaction frameworks for learning structured?* Table 8-2 collates all the features of peer-interaction discussed in the previous sections to list the terminology used when thinking about the multimodal coordination of learner interaction within the context of F2FCCW. All these terms were first introduced in the analysis chapters and then discussed in more detail in the previous sections 7.3 to 7.6. These actions (or 'behaviours' / 'features') were evident throughout the five stages of transmodal talk, discussed in the previous sections, so there is some overlap with Table 8-1. Here these features are itemised to illustrate inclusionary and exclusionary differences.

Table 8-2 (features of an inclusionary and exclusionary framework)

inclusionary peer-interaction framework	*exclusionary peer-interaction framework*	*Goodwin's five stances of organisation (2007a)*
1. *mutual cognitive-orientation* • *mutual coherence*	*divided cognitive-orientation* • *divided coherence*	*cooperative, epistemic*
2. *modal alignment and coordination* • *linguistic, proxemic* • *posture, gesture, gaze* • *density, complexity*	*modal misalignment*	*cooperative, moral*
3. *symmetrical configurations* • *mirroring / mimicry* • *reciprocal actions*	*asymmetrical configurations* • *dissimilarity* • *singular actions*	*cooperative, moral*
4. *spatial equality* • *shared space* • *layout* • *cognitive tools*	*spatial inequality* • *contested space*	*instrumental*
5. *proxemic distance* • *close, equidistant*	*proxemic distance* • *marginal*	*instrumental*
6. *linguistic equality, through:* • *polyvocality* • *echo utterance / latching* • *turn-taking* • *linguistic scaffolding* • *cognitive batons*	*linguistic inequality* • *monovocal*	*cooperative, moral*
7. *cognitive alignment* • *embodied* • *distributed / shared*	*cognitive misalignment*	*epistemic, affect*
8. *collaborative co-construction*	*competitive co-construction*	*affect*

Through the analysis of the Polish learners' collaboration, it was evident that the features in the left column were repeated and patterned throughout their task. These were all positive behaviours suggesting an inclusionary peer-interaction framework. Both learners were engaged throughout. Looking then at the Kurdish learners, through the lens of these 'positive behaviours,' it was clear that these features were either missing or minimal. In many of their actions, such as cognitive-orientation to the task and off-screen drafting, the features in the left column were completely absent. For the collaboration of the Kurdish learners this led to the terming of behaviours in the centre column. The sheer absence of a behaviour, or its opposite, in comparison to the Polish

learners, led to what could be construed as negative terms: inequality, misalignment, dissimilar and divided.

It is tempting to draw up binary stances between the two pairs of learners as the terminology used would suggest this is the case. However, there were occasional moments when the Polish learners demonstrated behaviours in the centre column, such as cognitive misalignment. Whilst identifying polar differences in learner actions can be useful for the purpose of describing and discussing similarities and differences, it can risk reductionist ways of thinking. Peer-interaction frameworks are fluid because of ongoing shifts in modal alignments. There were periods for example when Shourok's modal alignment was inclusive to the coordination of his own modal actions, thereby excluding Darras who sat in misalignment, but at other times they aligned linguistically, cognitively, proxemically and worked together. But more frequently they did not align as well as the Polish learners. Such polarisations of terms are useful for describing behaviour of learners in a given moment of time but actions will continue to change so caution should be taken not to prejudge any learner so they are placed in one binary position. There are differing scales of time in which actions occur and learning processes develop and there will be modal shifts in alignment between two or more people; sometimes that is conducive to peer-interaction and a shared learning outcome and sometimes it is not. Alignment can be found on a number of levels; facial mimicry (Blairy, Herrera, & Hess, 1999), postural (Bernieri, 1988), gestural (Bavelas et al., 1988), bilingual (Hartsuiker et al, 2004), vocalics (Neumann & Strack, 2000) and dialogue (Pickering and Garrod, 2004). Holly at al (2010) provide a summary of alignments across a range of literature as they try to inform alignment in computer interaction.

The right-hand column of Table 8-2 identifies the five stances of organisation identified by Goodwin (2007a). I have placed them adjacently on the rows to suggest possible correlation between my findings on peer-interaction frameworks and Goodwin's findings on participatory frameworks. 'The alignment of participants towards each other generates at least five different kinds of stance:

1. instrumental stance, the placement of entities in the ways that are required for the sign exchange processes necessary for the accomplishment of the activity in progress;

2. epistemic stance, positioning participants so that they can appropriately experience, properly perceive, grasp and understand relevant features of the events they are engaged in;

3. cooperative stance, the visible display that one is organizing one's body toward others and a relevant environment in just the ways necessary to sustain and help construct the activities in progress;

4. moral stance, acting in such a way as to reveal to others that the actor can be trusted to assume the alignments and do the cognitive work required for the appropriate accomplishment of the collaborative tasks they are pursuing in concert with each other, that is to act as a moral member of the community being sustained through the actions currently in progress; and

5. affective stance, emotions by the individual and toward others that are generated, in the situations being examined here, by the organization of participation in interaction' (Goodwin, 2007a: 70-71).

Goodwin's five stances of organisation will be discussed shortly in conjunction with Table 8-2 and Figure 8-2. The following overarching features can help to visualise the organising structure of peer-interaction, as discussed in the context of F2FCCW.

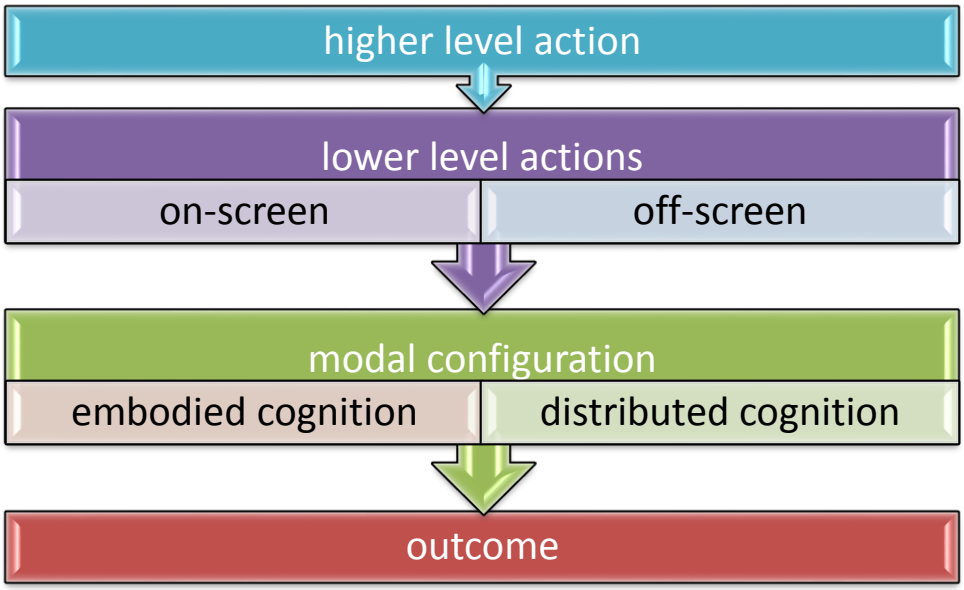


Figure 8-2 (peer-interaction framework)

An action was previously defined as having 'purpose, intentionality; there is agency and cognitive engagement with goal-orientated outcomes which are realised through physical means. In this sense actions are observable phenomena, the embodied manifestations of intention.' The interaction framework at Figure 8-2 visualises a

possible structure for the organising features of an action's outcome. A peer-interaction framework could be explained in terms of two broad levels of interaction: (1) a *higher-level action*, what the learners are trying to achieve, and (2) *lower-level actions*, what the learners do, their modal configuration of these to help them complete the action. The analytical approach to deconstructing actions was based on Norris' methodological framework (2004). In this context, the higher-level action started when the learners read what was required of them ('write a few sentences about Greenpeace or environmentalism') and was considered to be complete when the learners finished the writing and accepted it. The completed 'end product' in this context (i.e. the outcome of the higher-level action) was a piece of writing. At other times it was the design of an image or the design of a website. *Actions* across different scales of time (i.e. chronemic configurations) became the units of analysis for re-framing and analysing past activity.

The analysis chapters delineated the higher-level actions down to lower-level actions to understand their coordination and accumulative contribution towards achieving the higher-level action. The lower level actions within collaboration were found to be constituted on an ongoing flux of multimodal coordination where off-screen actions (gesture, languages, posture, proxemics, layout, cognitive tools) were interwoven in conjunction with on-screen modes (software, websites, languages, typography). The embodied and distributed cognition within this multimodal coordination led to the successful outcome of the higher action. Each learner contributed something to the process, albeit in differing amounts. An on-screen piece of writing was the realisation of their off-screen actions.

Figure 8-2 provided a headline visualisation of the organisation of interaction and Table 8-2 listed the common features in that organisation, with reference to Goodwin and the interactive organisation of stance within participatory frameworks. I now want to bring together this introductory discussion on stances and interaction into a final summary to answer how the peer-interaction frameworks for learning were structured in the classroom. Many of Goodwin's five stances do overlap but they are separated here to briefly show how each relates to my own findings.

instrumental stance (items 4 and 5 at Table 8-2)

A significant concern for an educator, when using peer-interaction, is to consider the way in which learners position themselves to each other and their environment. Particular learning outcomes require individuals and cognitive tools to be structured in such a way so that each is mutually beneficial. In F2FCCW, the spatial layout of the computer with peers creates a triangular configuration which destabilises the standard face-to-face communication stance so that the screen becomes the central focal point rather than the face of the person to whom one is speaking; though frequently the learners did turn to face each other when directly speaking to another. There is also only one set of tools (mouse, keyboard) which must be shared. Keyboard and mouse switches seemed to go together. Instances when one learner controlled the mouse, and the other controlled the keyboard, were rare. Across the completed lesson to finish all the learning outcomes, considering the mouse-switches where one learner had greater 'control' of the computer compared to another, the switches and overall mouse control time were very different across the two pairs. The Kurdish learners had seventeen mouse-switches between them with nine to Shourok and eight to Darras. Of these switches to Darras, seven out of eight were failed attempts to take control where he held the mouse but never used it; Shourok took it back. Darras held the mouse for a total of 4 minutes 39 seconds. Shourok had control for 2 hours 11 minutes. The Polish learners had thirty mouse-switches between them with fifteen each. Sakia held the mouse for a total of 1 hour 50 minutes. Gamda had control for 1 hour 1 minute. From a numerical perspective alone, it is clear that one pair of learners had a more equitable peer-interaction than the other.

Cognitive tools (be it pen and paper, hammer and chisel, measuring tape and scissors) are material artefacts for achieving goal-orientated outcomes. Through our actions they embody individual cognition (i.e. taking notes in a classroom, using a tape measure on cloth to cut to shape) and enable the distribution of cognition to others; i.e. showing someone how to spell a word or how to measure a piece of cloth. Gesture and language configure heavily within that sociomaterial process so for the learners there was frequent pointing at the screen in conjunction with deictic language such as 'this' and 'that' and even the white cursor arrow to point at items on the screen, as an embodied extension of a figure, to draw attention to something that was unknown by one participant or to show something that was already known. In this way, the learners embodied and distributed their understanding, or lack of, through a wider range of semiotic modes than language alone. The environment structured their learning but this

was dependent on how each learner placed themselves *instrumentally* to engage in that sociomaterial process. To accomplish an action, Goodwin stresses the importance of 'positioning for perception by taking up appropriate stances towards a world structured by both objects being scrutinized and other actors' (2007a: 61). In this research it was shown how each pair of learners positioned themselves for perception through their proxemic alignment to each other and the computer. Adopting an *instrumental stance* might be considered a fundamental component of successful learning. In most instances the learners adopted a close or equidistant position to each other and the tools and this was instrumental to completing the actions. When a learner marginalised themselves proxemically, or the learning space was contested, the peer-interaction framework became less stable. This can be seen for example in section 6.6 when Shourok took exception to Darras's accusation that he 'is doing it all himself' and Shourok pushed himself away from the interaction space around the computer and said 'done all yours.' Shourok positioned his body in such a way that he could no longer engage with Darras nor the learning. His gaze was directed elsewhere. His body posture was closed and positioned away from Darras. There was silence and when he did speak is was in single word utterances.

epistemic stance (items 1 and 7 at Table 8-2)

This distancing of one learner from another is what Goodwin might call a failure of epistemic alignment, but what I called cognitive misalignment, which leads to mutual or divided coherence in the task. Prior to contesting at section 6.6, Shourok did not try to align cognitively (i.e. epistemically) with Darras; this meant there was no mutual coherence between the two learners, so when Shourok positioned himself at a distance from the interaction space, Darras did not know how to proceed. We see evidence of Darras cognitively 'reaching out' to understand what he needs to do through gesture, laughter, smiling, gaze, proxemics, tools and question forms; as a brief example: 'and here open↑ (6) this↑ (5) what is this↑ and↑ (1) it should be what↑'. Epistemic stance is evidenced through how successfully, or not, learners cognitively align with others. It is about embodied positioning so as to be instrumental in the learning environment. From a peer-interaction perspective as an educator, one needs to enable learners to position themselves so they add epistemic value to their own learning and that of others around them. In this research there is more evidence for successful epistemic alignment. As a stance it is not particularly noticeable, when learners are working cooperatively, but

when there is cognitive misalignment that is when the loss of epistemic stance becomes far more noticeable. 'Instrumental stances … can also provide the basis for consequential epistemic stances' (Goodwin, 2007a: 61). Cooperative alignment between learners has positive sequential benefits at a cognitive level.

cooperative stance (items 1, 2, 3 and 6 at Table 8-2)

'Cooperating, or failing to cooperate … provides an environment for the visible emergence of both moral and affective stances' (Goodwin, 2007a: 62). The extent to which learners *instrumentally* position themselves to each other and the environment, or do not, can have positive or negative consequences for the epistemic alignment for others in the interaction. For Goodwin, evidence of the previous two stances suggests a cooperative stance. Very simply, learners are working mutually together, cooperatively, and that is seen through the instrumental and epistemic stances. It was demonstrated in this research how cooperative learning was seen in the symmetrical configuration of modal alignments between two learners, including:

- language (polyvocal, echo utterance, turn-taking, linguistic scaffolding, cognitive batons),
- proxemics (close, equidistant),
- posture (mirroring),
- and evidence of reciprocated gaze, sharing tools, etc.

Where there was evidence of asymmetrical configurations, this is where non-cooperation was found. In asymmetrical configurations there were dissimilar actions and a lack of reciprocation. For example: an unanswered question, peripheral positioning, an unreturned gaze, etc. Returning to section 6.6 with the Kurdish learners, this was the most prevalent example of a failed cooperative stance. We saw the absence, and opposite, of the symmetrical configurations above. Learning was suspended as Darras struggled to continue alone with the task.

moral stance (items 2 and 6 at Table 8-2)

This brings us to the moral stance. Goodwin calls it a 'moral failing … that affects not only the defaulter but also the current, immediate projects of co-present others' (2007a: 65-66). This can be seen with Darras as above, from section 6.6. Shourok's 'failure' to *cooperate* in the task is proven by his proxemic and *epistemic* withdrawal from the

cognitive tools and his peer so that he is no longer *instrumental* in the learning environment. In educational settings such behaviour might be punished by the teacher. Peer-interaction is different though as there is no official 'policing' of behaviour; this has to come from the peers. The LSA walks away following Shourok's 'done all yours.' This leaves Darras alone to redress Shourok's distancing of himself from the interaction. He could have chosen to call for the teacher or to continue alone but because there is no mutual coherence, no epistemic alignment, Darras is dependent on Shourok to proceed.

affective stance (items 7 and 8 at Table 8-2)

'Moral failings' within participatory frameworks have an *affective* impact on individuals. There is an emotional cost to non-cooperation. The LSA challenges Darras for not speaking enough. Darras challenges Shourok that he is 'only doing himself.' Shourok returns the challenge: 'come on say something' and then abandons the interaction:

01:01	Shourok	done all yours
01:07	Darras	she doesn't mean this
01:09	LSA	no(h) no(h) (h)no really (h)no (h)no I did(h)n't me(h)an it haha

Darras and the LSA become apologetic with defensive language and nervous laughter. The moral fabric which enables people to cooperate has broken and the impact is felt at an emotional level. Individuals become sad, angry, submissive, aggressive, etc. It becomes something of a 'battle' to re-engage Shourok with the interaction and Darras cognitively 'reaches out' to Shourok through questioning, gesturing inwards with his finger, smiling, making eye-contact. In Goodwin's research into non-cooperation, he observes that 'One way to resolve such a dilemma is to refuse to engage in further interaction' (2007a: 67-68). This is Shourok's approach. In peer-interaction learners need to be trusted to work cooperatively and it is this failure of trust, rather than non-compliance to an enforceable rule, which perhaps makes it a 'moral failing.' However, Darras skilfully uses powers of persuasion to re-engage Shourok. For the Polish learners there was no similar 'moral failing' in their interaction or negative affective stance towards their learning. This is probably because they maintained alignment between them across a number of modes and stances throughout their learning:

- learner alignment of an instrumental stance,
    - i.e. when Sakia translanguaged with a Polish website Gamda did the same with a pocket translator;

- learner alignment of an epistemic stance,
    - i.e. when Sakia and Gamda scaffolded their sentences through phonemes, syllables, lexemes and chunks;

- learner alignment of a cooperative stance,
    - i.e. ensuring there was linguistic equality in their writing through a consideration of the other's 'voice';

- learner alignment of a moral stance,
    - i.e. reciprocation and recognition of each other's opinions and contributions;

- learner alignment of an affective stance,
    - i.e. using humour and positive body language with each other through mirroring.

Whilst some of these features were evident in the Kurdish learners, they were not as consistent and there were many opposite behaviours which would suggest that for the Kurdish learners their peer-interaction framework was predominantly exclusionary. In item 8, Table 8-2, I placed Goodwin's 'affect' alongside collaborative co-construction and competitive co-construction for it is perhaps here where there is most evidence of positive and negative emotional impact on learners. For the Polish learners it was clear that they had an inclusionary peer-interaction framework and that there was *collaborative co-construction*. They appeared to enjoy the task. For the Kurdish learners I applied the term *competitive co-construction* because their interaction often looked more like a battle than collaboration. There was more discord than harmony with the learners competing for tools and spatial control and voice.

Goodwin observed from his research that refusal to align can initiate 'a cascade that undercut(s) all of these forms of stance' (2007a: 71). This was clearly evident here with the Kurdish learners. Shourok's *moral* failure to *cooperate* with Darras resulted in *epistemic* uncertainty to the point where neither of them were *instrumental* and both suffered the *affective* consequences of this at an emotional level. Alignment across multiple modes and stances (as itemised in the left column of Table 8-2) is possibly one of the key components of a successful peer-interaction framework because without those alignments the whole structure of embodied, shared learning weakens and can even fall apart.

8.4 recap of purpose and findings

The research sought to understand the participation frameworks of second-language adults as they designed multimodal media, in one classroom session, collaborating in pairs to create an environmental booklet with associated text, imagery and a website. The primary motivation was to understand how talk and action are configured at a shared computer when language, literacy and technology combine. In schools and colleges, with the number of students outnumbering available computers, a common tactic is to pair people around a single computer. In second-language learning contexts, a pedagogical justification for pairing learners at a computer is often to encourage authentic language production; see section 2.5. The workplace is a similar site where two or more people may congregate around a single viewing space with a single set of tools. 'Collaborative computing is a significant and pervasive social phenomenon in education and the workplace, from early childhood to adulthood' (Gardner & Levy, 2010: 1). Whilst there are many social and pedagogic benefits to pairing learners, there is also a risk of conflict, unequal collaboration and competition between participants who are collaborating on a joint task. Such conflict is more frequent in school settings (Crook, 1994: 135) but can occur in adult settings.

The aims of this research were: (1.) understand the coordination of talk and text as a transmodal event in peer-interaction, and (2.) through the analysis of transmodal collaborative writing, work towards an understanding of peer-interaction, one which considers learning to occur within a multimodal participatory framework of embodied and distributed cognition. In the context of this research, the focus on cognition was specific to how learners exhibited (i.e. *embodied*) and shared (i.e. *distributed*) skills and knowledge through language and embodied interaction. Many cognitivist studies in second-language acquisition frame learning and understanding as internalised processes which can be measured quantitatively by the volume of accuracies and errors in output, such as grammar (Doughty & Long, 2003). In this research the approach was sociocognitive, where there was still significant interest in the demonstration and cascading of knowledge, but learning and understanding were framed qualitatively by what learners said and did and how they organised themselves (e.g. Atkinson, 2010).

Analysing transmodal talk revealed the fluid structure of peer-interaction frameworks, within which, embodied cognition and multimodal actions occur. Goodwin explains the importance of understanding this: 'a theory of action must come to terms with both the details of language use and the way in which the social, cultural, material and sequential structure of the environment where action occurs figure into its organization' (2000a: 1489). Episodes of interaction were found to be fluid frameworks, constantly fluctuating in response to how learners aligned themselves towards each other and the semiotic resources they employed to complete the tasks. Though modes and alignments shifted constantly in their interaction, and the configuration of the framework flexed and fluxed with those shifts across different scales of time, there still remained a framework which could be named and described and participant roles identified. The findings from this research, in relation to the field of multimodal interactional analysis, were summarised in section 8.2 as a peer-interaction framework (PiF) with features identifying inclusionary and exclusionary behaviours. The features of transmodal talk were identified at section 8.1.

In education, as in many collaborative contexts of joint activity, frameworks of interaction can be co-constructed and hence relatively equal in participation and co-cognition. Other frameworks are less equal because they can be exclusionary. As an educator, a concern is the potential loss of learning opportunities for one or more participants in peer-interaction. Who does what, how do they do it, what is the impact, how long does it take, are all reasonable questions to be asking of peer-interaction. To identify the organisation of peer-interaction frameworks, video technology was used to record the learner's interactions off-screen and on-screen. This exposed the significant features of learner alignment and modal configuration in the material and sequential structure of their interaction. It is important to understand this because 'all cognition takes place in human bodies embedded in sociomaterial worlds' (Atkinson, 2010: 619). Understanding the embodied dimensions of learning can help educators make informed decisions.

8.5 connections with previous research

The findings of this research connect primarily to three fields of research: (1.) second-language collaborative writing; (2.) embodied peer-interaction within multimodal interactional analysis; (3.) new literacies within the context of digital technologies as

part of the third generation interest in empirical literacy research. These are discussed in the following sub-sections.

8.5.1 second-language collaborative writing

The process of second-language writing in this research is broadly in line with similar research. *Language-related episodes* (Swain and Lapkin, 1998) were found when learners demonstrated an uncertainty with the second-language they were using. This was identifiable in their conversations when learners *languaged* (Swain, 2006); that is, they metacognitively used language to discuss language and they also *translanguaged* (Garćia, 2007) when they used their first-language to discuss the second-language. Language-related episodes (LREs) were categorised as attempts to improve accuracy in *form*, *lexis* and *mechanics*, for example:

- Form-focussed: e.g. choosing the right tense; using the correct grammatical forms.
- Mechanics-focussed: e.g. discussing pronunciation, spelling, punctuation.
- Lexis-focussed: e.g. deciding on the best word, or chunk, to use and its correct meaning.

A mechanism to trigger *languaging* was for one of the learners to *notice* (Schmidt, 2010) an error or opportunity for improvement. This was mostly done through *negotiation* (Long, 1996) in their peer-interaction. Phonological s*caffolding* between the learners was evident at phoneme, syllable, lexeme and lexical chunk level. In this research the learners were of equal proficiency. Further research in similar contexts might pair learners as high-high, low-low and high-low to measure phonological s*caffolding*.

Krashen's *comprehensible input* (1985) sits within the discussion on scaffolding as does *comprehensible output* (Swain, 1985) in which a focus on output promotes (a) a *noticing* function (b) a *hypothesis-testing function* and (c) a *metalinguistic awareness function* (Swain, 1995). 'Input' was not considered here but a focus on output was clearly evident and this was discussed in section 7.6. Rouhshad states that 'one way to enhance negotiations, particularly negotiations for form, in learner interaction may be to require learners to write collaboratively. This is because affordances of writing (i.e. extra processing time, greater demand for accuracy and permanence of the text) may enhance attention to form' (2015: 16; see also Williams, 2012). Collaborative writing tasks are now being recognised as one of the most beneficial pedagogic methods for

improvements in overall L2 accuracy (Rouhshad, Wigglesworth, Storch, 2015) because it encourages *noticing* of issues with form, mechanics and lexis through peer-negotiation. *Languaging* and *translanguaging* strategies are evident in such *language-related episodes* and were equally found here in a digital writing context. Microanalysis of videoed interaction was able to demonstrate how that process unfolded as transmodal talk; something which is missing from existing research in collaborative L2 writing.

8.5.2 embodied peer-interaction within multimodal interactional analysis

The works of Charles Goodwin and Sigrid Norris have been pivotal. The core idea in this research has been the perception of 'action' as embodied, cognitive, temporal, multimodal and co-constructed. Their analytical methods towards describing the minutiae of interaction allowed me to apply similar methods. Section 8.3, as a concluding summary of the findings of embodied peer-interaction, was able to demonstrate a link between forms of alignment found in this study with the five stances of organisation in Goodwin's (2007a) participatory framework. Taking from Norris the perception that actions can be thought of as higher and lower, and comprised of modal density, helped me to deconstruct a full classroom session from its totality of three hours down to split-second utterances and gestures. Actions from the perspective of fluctuating chronemic arrangements was (and is) a practical approach to the analysis of interaction (Lemke, 2009). Goodwin and Norris in-turn acknowledge their debt to the research undertaken in *interactional sociolinguistics* and similar *situated activity systems* (Goffman, 1961; Gumperz, 1982; Levinson, 1992; Scollon & Scollon, 2003) and to which this research has links.

Goodwin acknowledges: 'Such systems constitute an environment within which the analyst can investigate in detail how participants deploy the diverse resources provided by talk … sequential organization, posture, gaze, gesture, and consequential phenomena in the environment that is focus of their work in order to accomplish the courses of action that constitute their lifeworld' (2000a: 1519). The same approach has been applied in this research. In this research the analysis of language in action was prioritised to give a detailed description of 'language as social practice to accomplish social action and … alignment, fitting oneself to one's environment' (Atkinson, 2010).

8.5.3 new literacies within the context of digital technologies

To a lesser extent there is a link with new literacies as this was a background context to the weekly learning activities of the group: reading and producing English in a range of electronic media. Baynham & Prinsloo explain that there has been 'a growing concern with the multiple communicative modalities that underwrite literacy practices in contemporary times, besides print, including sound and movement, particularly with regard to screen-based multimedia literacies' (2009: 5). The research methodology in this study used audiovisual methods of data collection in an attempt to capture as much as possible of the multimodal literacy events, on-screen and off-screen, seeking to identify how the different modes align and configure when two people interact with a single computer, including linguistic and non-linguistic modes. As discussed by Prior, multimodality is broader than the traditionally researched *sites of display*, such as a book, screen, posters, etc. 'I do not believe that we can account for multimodality and affordances without a focus on the whole of practice – on artefacts, activity and people alike' (Prior, 29: 2005). Research into digital literacies in education has favoured a product approach to discussion and analysis with less consideration of the process; as discussed in section 2.6. Literacy artefacts already produced and how learners interact with these, rather than how learners produce multimodal designs, is a typical characteristic of the analysis. This research has redressed this in a small way.

8.6 limitations of research

An ethnomethodological approach broadly seeks to make sense of order in local, situated contexts, one that insists 'on the adequacy of description and a focus on contingent empirical detail' (Rawls, 2002: 4). This inevitably means breadth is sacrificed for depth. A large amount of descriptive detail was provided on the coordination of talk, text and action in chapters 5 and 6. Analysis of that in chapter 7 led to a framework for identifying observable behaviours in peer-interaction (positive and contested; section 8.3) and a framework for identifying the structure of transmodal talk, section 8.2. Only two pairs of learners were analysed and each of them for only 10 minutes of interaction each. They were also second-language learners and adults in a further education college. Clearly this means the findings are not easily generalisable given the very narrow sample and type of participant. Different environmental contexts, different age groups and different sociocultural backgrounds might provide alternative behaviours. However, generalisability was not the purpose of a detailed, small-scale case study.

In addition, this detailed level of microanalysis did sacrifice wider discussion which could have been had on the wider connections of digital literacies intersecting with the learners' classroom practice. The learners in this study for example discussed how literacy and language intersects with technology and paper in a number of their lifeworlds: work, family and education. In section 3.2 it was discussed how the New London Group (2000) called for a pedagogy of multiliteracies in response to cultural and linguistic diversity. Interviews with the learners (Appendices D to G) evidence learners' concerns with essay-writing, website designing, reading books, chat rooms (Darras), writing poems (Gamda) and a desire 'to have normal English' says Sakia. Literacy mediated by digital and traditional means is evident across: (1) their working lives (2) public lives and (3) personal lives (Cope et al, 2000: 10-17). The interconnection of these might have provided a more holistic account of each individual's engagement with literacy and technology; such as in the research of Bhatt (2014) who investigated literacy 'events' at the intersection of traditional and cyber spaces.

Looking at the literacy practices of learners in FE and HE settings, Satchwell et al note 'the prevalence of *digital* literacy practices in students' everyday lives, and the prevalence of *paper-based* practices on their courses. However, this was not the only difference, and we came to realize that we needed to analyse the whole range of aspects of any literacy practice in order to identify particular elements that could travel across boundaries' (2013: 45-46). The same perception of *boundaries* could be applied here. For example, recent research into academic text trajectories offers a sociomaterial approach to understanding the social organisation of writing across a range of technologies in space and time: e.g. the intertextual weaving of an academic's blogs, emails, tweets, journal articles, lectures, etc (Tusting et al, 2015). A technobiographical approach to a learner's digital writing in multiple domains and contexts is a potential development of this research. Bridging the *inside* of the classroom with the *outside* of the learners' lifeworlds can help to generate relevance and authenticity in pedagogy (Woulds & Simpson, 2010). An ethnography (or technobiography) into the sociomaterial writing habits in the wider digital 'worlds' of the learners could be informative for practitioners and researchers. Some of the learners in this study for example talked about their website designs in English as they became *digital scribes* for their local communities, compared to traditional literacy scribes (Baynham, 1993). One

designed a website at the request of her local church. Another, a sickle cell sufferer, designed a website on the disease for a charity group based in Leeds; another a bilingual website for his Kurdish community; another a website for the employer of the tanning salon she worked in. The *boundaries* of interest in the current study were narrowed to the classroom in a collaborative digital writing context. This presents limitations in the research from the perspective of second-language writing and digital literacies.

A further limitation is the methodology. The *videography* (Knoblauch, 2012) of data collection required a level of technological skill which is not easily replicable. First, there was the practical problem of orchestrating so many video and audio recording devices in the research environment; including a macro lens on global actions, a meso lens on localised embodied interactions and a micro lens on the sequential development of the object at the core of the interaction: here the on-screen writing and software. An analogy of the complexity might be to think about trying to analyse surgery at similar levels of visual detail: capturing interaction across the room by all individuals, the close-up actions of the surgeon with the patient's body and then the micro-detail of the surgical tools being used to undertake the operation. Because this was my own teaching classroom I was able to repeatedly pilot the methods and place the technology in a way which is not easily replicable in other social settings. Second, gaining access to so many camcorders and voice-recorders is not easy and is a financial cost to consider. Because I work in an educational setting I was able to borrow the technology from a common resource pool. Third, integrating multiple audio and video files into video-editing software for synchronisation purposes adds an additional level of complexity and a requirement for extreme levels of patience. This also requires familiarity with software for video analysis and audio editing; here Transana and Audacity. Fourth, presenting the findings of multimodal data requires the designing of bespoked systems of representation, which in previous research has included the use of photographs, drawings, tables, screenshots, etc. Communicating the findings of multimodal data requires transposing sounds to print, three-dimensional space to a blank page, gestures to static description and sequences of moving time to frozen time-stamps. No single means for doing this has yet been developed.

8.7 problems arising during the research

As discussed above, and at section 3.3, the most significant problem was collecting and managing the data. The piloting stages required experimentation with where to place the

cameras and how to record the on-screen activities. Hours and days of manipulation generated hours of useless audio-visual data. Lessons learned however did allow for fine-tuning over a number of weeks to the point where I was able to record every gesture, every word and every action of everybody. The synchronisation issue discussed above, an outcome of having so many recording devices, was because the classroom I was in had older computers and slow processors. A way around these synchronisation issues is to use hardware and software with integrated recording systems. New computers now have built-in cameras which can capture the meso level of detail around the screen. Adding a high quality microphone could capture all the audio around the screen. Using a program like Camtasia can simultaneously record the external audio, the on-screen actions of the mouse and keyboard and the off-screen actions of individuals through the built-in webcam. The software would generate one video file, though the quality of the built-in recording devices would need careful consideration.

Working with second-language students means there will be code-switching at some, and usually many, points in a classroom session. Transcribing audio data can thus be complicated by multiple languages. In this research work-colleagues were able to translate these for me. Consideration should be given to the languages of individuals and later complications with translation.

The secondary method of data-collection was interviewing, and as discussed at section 3.4, revealed issues with the reliability of my interviewing skills, the process and thus the validity of my interpretations. I view *reliability* as synonymous with the *replicability* of the researcher's data collection processes and *validity* as synonymous with the *accuracy* of the knowledge claims made about what has been discovered (Cho & Trent, 2006; Eisner & Peshkin, 1990). The reliability of data collection methods, how they are checked and how the data is interpreted seem to me to be a solid foundation for any claims to validity, regardless of quantitative, mixed-method or qualitative paradigms. Member-checking (Morse et al, 2002) was introduced so that the students could approve the accuracy of the transcribed interviews. The addition of triangulating (Bryman, 2007) audio-recording devices to improve transcription and self-reflexivity in the writing-up (Richardson, 1997) were also used in the hope that the data would be more reliable and hence the interpretations more valid.

8.8 implications of findings

Whilst transmodal talk and peer-interaction were framed and features named, and this is considered a positive contribution, there were some complications for the learners involved. Sharing technology such as a mouse and keyboard can result in unequal task-collaboration. 'Higher order cognitive functions emerge as the result of interpersonal interactions' contend Philp et al (2014: 159). If this is true, then as educators we need to think and plan carefully on how we maximise learning opportunities with the pairings we impose, and on the pairings learners choose for themselves. Across the two pairs there were many instances of *competitive interaction* with the two males rather than *collaborative interaction*. Temporal mapping of talk to text was still co-constructed but rather than *collaborative co-construction* of meaning, as with the Polish learners, there was *competitive co-construction*. This was evident by Darras seizing moments to use the keyboard and mouse when Shourok was distracted and by Shourok refusing to cooperate when challenged. Section 2.4 problematised what is meant by collaboration, and whilst this is a generalisation based on experience, a great deal of paired and group interaction is based on the premise of encouraging cooperative learning. The social aspect is laudable but the actual distributed cognitive payoff is uncertain.

However, in the post-task interview (Appendix I) Darras acknowledged that he learned new things from their peer-interaction saying: 'I didn't know something but when he did it I learn it' and 'I learned from him yes.' He talked about the experience of peer-interaction as one that was 'not easy' but recognised the value of peer-interaction: 'sometimes you forget something and your partner can help.' Shourok said he found peer-interaction 'difficult,' which possibly says more about learning preferences between the two learners. In the post-task interview of the Polish learners (Appendix H) Gamda thought sharing a computer in peer-interaction 'was a good idea' and they tried to 'be creative together' but they both recognised the challenge of balancing different choices. Collaborative support is something they regularly do, as Sakia says: 'We usually support ourselves. Sometimes I look at Gamda's screen and sometimes she looks at mine.'

One would expect that most educators pair their learners accordingly. However, thought could be given to the named delegation of tasks and explicit discussion provided on cooperative peer-interaction. Sharing a computer for example, learners could be given a list of instructions with each learner named against each task, or time-slots given when

one learner has control of the mouse and keyboard before releasing it back to their peer. Educators still have an obligation to manage peer-interactions by observing behaviours and correcting, as Philp et al identify: 'the teacher is an ever-present resource, even when group work is the dominant pattern of interaction' (2014: 192). Observing the peer-interaction in this research exposed *alignment* and *misalignment* as key indicators of learner engagement in the learning process. Noting the observable behaviours at Table 8-2 in section 8-3, tutors could seek to manage peer-interaction pairs by looking for alignment or misalignment across language, proxemic distance from each other, postural direction, gesture and gaze between each other, and an equitable sharing of tools and resources. This includes consideration of modal density: are some learners showing preference for one mode over another and if need be, redressing that; for example, avoiding writing tasks or reading or speaking. Noticing forms of modal misalignment between learners and modal preference could be an opportunity to help individuals.

Collaborative enterprise with new literacies is a second-language learning opportunity which could be employed more widely by educators. The learners in this study talked positively about the multimodal construction of websites and images and the affordance of these over other modes, as Shourok says at Appendix G: 'If you look at my website its name is iran4all. It gives me a different way to *talk* to others. If you can't say it in one way you can say in a different way. It's like Rimi said, if you can't speak in good English you can say it in another way, for example, designing a protest image and saying it that way. If we don't like something and we can't say it we can communicate through images or file; something like that.' The multiliteracies strategy of Cope et al (2000) is highly relevant here. As discussed at section 8.2, the focal point of this research was bounded by pedagogy only and did not consider the wider sociomaterial practices of digital writing in the students' personal spaces. It is therefore difficult to make empirical connections between the micro of the research questions and the macro of the broader sociocultural and political contexts within which the participants lead their lives. As mentioned in section 3.2, the learners in this study belong to a superdiverse urban area and as such are representative of the 'type' of learner discussed under the pedagogy of multiliteracies, where increased levels of migration have led to increasing levels of cultural and linguistic diversity in densely populated urban areas in many UK cities. To be literate means to be able to engage in the multiple spaces which populate the domains of work, public lives and personal lives (Cope et al, 2000, 10-17).

‘How adult students engage with English outside class, in the broader sociolinguistic setting, is relevant for teachers. Approaches to language teaching generally should encompass a concern with students’ needs’ (Simpson, 2016, forthcoming). For instance, completing a form on a website creates different barriers to completing a paper-based form with a pen. Interactions with institutional others (e.g. welfare and employment offices) are more frequently mediated by English as a second-language in a digital format. A recent personal example is a young student told he was no longer eligible for income support so had to apply for jobseeker’s allowance. He arrived at the college with a typed web address (given to him by the jobcentre) to locate the form he had to complete. Whilst web-forms are now commonplace, for some they are still mystery and complicated further by a second-language. His confusion was palpable. With a teacher, he sat at a computer and they worked together in a manner similar to the collaboration of shared computer use seen in this research. The classroom can bring controlled and ‘safe’ exposure to multiple text types in multiple media; an opportunity to experiment and get it wrong in practice prior to textual and digital performance in the ‘real’ world.

Finally, learners could be encouraged to share technology, and allowed to translanguage rather than demanding an ‘English-only-rule.’ The structure of transmodal talk, at Table 8-1 in section 8-2, suggests learners have a number of strategies for languaging with language and the internet as corpus has potential. Working in isolation at a computer and with English only should be complemented with opportunities for dual computer sharing with same-language learners. All learners in this study used translanguaging as a method to cognitively work with English as a second language, and by negotiating off-screen talk to transpose to on-screen text, used a range of phonological strategies, scaffolding language through phonemes, syllables, lexemes and lexical chunks. In text construction words became cognitive batons between each learner, in a range of languages and from multiple sources, increasing lexis and enabling polyvocal practice. The implications for teachers is a pedagogic strategy with positive potential for second-language learning. Outside the classroom, the TLang project is currently undertaking a number of linguistic ethnographies in superdiverse urban areas in four UK cities (TLang, 2014-2018). Translanguaging is a strand of that research and promises interesting insights into the role of multilingualism, both in the class and outside the class. Simpson equally contends that ‘Approaches to pedagogy which draw upon notions such as translanguaging for their theoretical bases would seem to be particularly

fitting in educational settings in the global cities of today, where students may well be developing their competence in English as part of a multilingual repertoire' (forthcoming, 2016). This would suggest that a demand to conform to monolingual conventions in ESOL classrooms is incommensurate with how many bilingual learners actually use language(s) in their work, public and personal spaces.

8.9 recommendations

Research into peer-interaction to improve speaking demonstrates that there are clear advantages, with learners paired as low-low, low-high and high-high. The same can be said of peer-interaction to improve writing, though to a lesser extent, so more research is needed here. It is unclear of the extent to which peer-interactive writing has an additional beneficial impact on speaking. In collaborative writing, learners are inevitably speaking to each other and adapting, suggesting, correcting each other's language prior to it being rendered into text. Research on the dual benefits of collaborative L2 writing is an opportunity for further research, particularly as a longitudinal study. Additionally, whilst research appears to confirm that collaborative writing produces text with more accuracy as an outcome of the pair or group, it is unclear if that improvement in form becomes individual intake. A longitudinal study is required to assess the transfer of skills from group, or pair level negotiation, to individual improvement.

As polyvocal practice, educators could *stretch and challenge* a learner's vocabulary by explicitly teaching how to use 'another's speech in another's language' (Bakhtin, 1981). The learners' translanguaging strategies with each other, and use of websites, successfully enabled them to produce writing with vocabulary they might not have had or could not easily remember. An obvious caveat is the risk of plagiarism but educators could introduce the internet as a corpus for language development opportunities, with the different sites as corpora (Cheng, 2012; Hunston, 2002). Discussion around strategies and acceptable 'borrowings' from websites could help students; for example, identify individual lexemes and lexical bundles, but not sentences, and integrate them into their own writing. Then encouraging learners to morph 'found' words into lexeme families and suitable synonyms. The Polish learners for example explored synonyms in their own language to find a semantic correlation with an English equivalent. Learners sometimes need to 'play' with language, experiment, and as explained, peer-writing can

be a significant pedagogical strategy for language development: collaboration and experimentation rather than exclusive tutor correction.

Van Leeuwen states that a future direction of multimodal research is to engage further with more everyday technologies: 'writing softwares such as Word, Excel, PowerPoint … are themselves semiotic resources which build in constraints and affordances that deeply influence not only what can be said and how in these media, but also how the different semiotic modes they include can combine' (2011: 680). Further research is needed on how these are used as inter-semiotic resources for building meaning and as artefacts for distributed cognition, both in the immediate environment and wider. Such softwares are now the common tools for learners in most educational settings. In this research there was detailed analysis of the physical engagement of learners with each other and the immediate tools of the environment but very little discussion on the affordances of the software itself and the choices available to the students. Learners' embodied engagement with these individually and collectively, as classroom practice, and the affordances of the different software functions, merits further discussion. With increased cloud computing and remote networking it is becoming easier for individuals to work on centralised software documents from anywhere and on any technology.

For example, a student can create a PowerPoint on a standard computer at school with another student, save it and open it at home on a laptop with family, work further on it using a smartphone on the bus back to school the following day, then upload it to an e-Portfolio website for a teacher, which is then checked by a school internal moderator and then an external moderator. The lifespan and the journey and the audience of a single digital document is potentially huge. Bhatt for example notes that 'assignments, as with all entities, are sociomaterially constructed through practices and have a life prior to and beyond the classroom' (2014: 270). Gourlay (2014) explores the dimension of time across material domains to understand the co-agency of devices and technologies in education. Tusting et al (2015) seek to understand the writing spaces and text trajectories of digital technologies for academics in university settings. There has been a research propensity to explore 'innovative' web-based technologies (e.g. wikis, forums, blogging, vlogging, virtual worlds) but there is a profusion of the 'everyday software' around us in need of further exploration; particularly, the chronemic arrangement of polyvocal writing in simple software across space, time and technologies.

Section 7.2 introduced a peer-interaction framework with identifiable features of inclusionary and exclusionary behaviours. This also included discussion on Goodwin's 'five stances' of organisation in participatory frameworks. The setting was further education with adult learners. For comparative future research, the findings here might be applied in other dyadic and triadic frameworks, where technology is present as a cognitive conduit in the interaction and where software are the portals through which collaborative meanings are made; for example: (1.) other formal learning contexts such as schools, HE, and alternative science and humanities curricula; (2.) interactions in alternative collaborative contexts where a computer is present: work, home, shops; (3.) interactions where the teacher is present with the learner interaction at a computer; (4.) interactions where the participatory framework uses a range of learner proficiencies with technology: low-low, high-high and low-high. Videoed interaction in such contexts might explore the extent to which alignment and misalignment is evident across multiple modes and technologies in different contexts.

8.10 contributions to research

There are potentially four strands of research which this thesis contributes to: (1) a research methodology was developed for the collection and analysis of video data; (2) an analysis of the *design-process*, as opposed to *product-analysis*, may contribute to the field of multimodal literacies; (3) a peer-interaction framework was presented which broadens our understanding of classroom interaction in the field of multimodal interactional analysis; (4) the specialism of second-language writing was extended through an analysis of peer-writing using a computer.

8.10.1 videography

The methodology of videography as research is relatively new and it is hoped that methods and lessons learned from this research has some contribution in what Knoblauch calls a 'promising and rapidly growing field' (2012: 253). His preference leans towards video methods in sociology but as he goes on to explain: 'In addition to their focus on social interaction in natural settings, many videographers also share a number of methodological convictions. On the basis of ethnomethodology, social constructivism and conversation analysis, a number of methodological suggestions for the analysis of video have already been made that tend towards a certain paradigmatic pattern' (2012: 253). In educational settings the use of the simple camcorder holds

promising opportunities for observing and analysing social interaction in qualitative research endeavours. The praxeology of doing that, reflecting on it, improving it, are complex but not overly so (Mondada, 2009). Videography as research methodology is defined here as one which prioritises the video-documenting of learning in natural social interactions. Four sequential strands were identified and discussed: *collection* (section 3.6) *preparation* (section 3.7) *transcription* (section 4.5 and Appendix A) and *analysis* (chapter 4). In this research the mechanisms for videography have been trialled, researched and improved upon. In this sense the practical 'doing' of video research has a contribution.

8.10.2 multimodal literacies

In a smaller way, there may be some contribution to the *third generation* of empirical work into literacy practice. Literacy as reading was not part of this research but literacy as writing in a digital context was. Multiliteracies as a pedagogic strategy (Cope & Kalantzis, 2000) was evident as learners engaged in multiple forms of representation, writing in leaflet designs and booklets, creating visuals in image software, designing websites. All learners voiced a sense of satisfaction with learning in this way and the collected data demonstrated how they actually went about it. The totality of their engagement across a single session was captured as higher-level actions (Appendices B and C) but only a tiny part of that could be analysed because of the constraints of the PhD. Future opportunities to analyse the remaining data could offer additional findings on the learners' engagement with multiple literacies and modes. This research sought to identify how learners of English produce multimodal texts in real-time and the findings evidenced interesting polyvocal practice in their completed literacy artefacts. It is hoped that the current research has contributed something new to the field of literacy, in terms of additional literacy events for exploration, and in the methodology of how one goes about doing that. In this study, polyvocality was evident as an outcome of the learner interconnection between digital and traditional spaces, but not the wider literacy and technology spaces outside the classroom. This is an area for future development.

8.10.3 embodied peer-interaction framework

The development of a visual transcription notation (Appendix A) and method of analysis (chapter 4) are possible contributions to multimodal interactional analysis. Looking at peer-interaction through the lens of embodied analysis required the production of unique methods for analysing and presenting the data. Language was

found to be the dominant mediating, communicative mode in most of the learner interaction. It is assumed that the features of peer-interaction evidenced here would be more-or-less similar in other contexts where individuals collaborate together with a computer. The embodiment of language with gesture and the environment had similarities with the findings of Norris (2004) and Goodman (2007), suggesting the findings from section 7.8, including the videographic methodology, could be applied to other contexts in which individuals interact. Essentially, from the multimodal transcription and analysis, there was evidence of co-construction of meaning suggesting co-cognition through peer-interaction. For researchers and educators, the terms and ideas in that framework might prove useful in similar learning contexts.

8.10.4 second-language writing

The identification of transmodal talk and the features of how language is scaffolded and sequenced in collaboration is a possible contribution to the field of second-language writing. The structuring of talk at a computer had similarities with the findings of Gardner & Levy (2010) and Ulhirová (1994), suggesting the findings at section 8.2 could be usefully applied in other collaborative encounters with shared technology. For the learners in this case study, much time was spent worrying about the accuracy of their writing. In terms of 'errors in output' this was significant to them so discussion in sections 7.5 and 7.6 explored how they negotiated issues with accuracy of form (their language-related episodes) and related this to the literature. Learner focus on accuracy correlates with the findings of similar research (e.g. Shehadeh, 2011; Dobao, 2013; Ajmi & Ali, 2014; Sajedi, 2014). The findings from this research, in relation to the field of collaborative L2 writing, were summarised in section 8.2, evidencing how transmodal talk was scaffolded across a number of modal configurations.

8.11 autobiographical conclusion

Starting this PhD for me was borne out of curiosity and professional engagement as an educator. A teacher walks through a classroom door. The event is fixed in time and space and duration. A couple of hours later, the same teacher walks out the same door; usually! Somewhere in between some learning happens; hopefully! It is likely we will do this tomorrow and tomorrow and tomorrow. Familiarity can breed complacency. Formalised research into one's own practice can bring about new ways of seeing … and heightened levels of anxiety. It is worth it. People are fascinating; learning is amazing. See through the eyes of a child.

List of references

Agar, M. (1996) *The professional stranger: An informal introduction to ethnography* (2nd edition) San Diego, CA: Academic Press.

Ajmi, A. & Ali, H. (2014) Collaborative Writing in Group Assignments in an EFL/ESL Classroom. *English Linguistics Research*, Vol 3, No. 2.

Altheide, D. & Johnson, J. M. C. (1998) Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (eds.) *Collecting and interpreting qualitative materials*. (283-312) Thousand Oaks, CA: Sage.

Amirkhiz, S. Y. Y., Bakar, K. A., Samad, A. A., Baki, R. & Mahmoudi, L. (2013) EFL/ESL Learners' Language Related Episodes (LREs) during Performing Collaborative Writing Tasks. In *Journal of Language Teaching and Research*, Vol 4, No 3, 473-479.

Andersen, P. A. (1999) *Nonverbal Communication: Forms and Functions*. Mountain View, CA: Mayfield.

Anstey, M & Bull, G. (2006) *Teaching and Learning Multiliteracies: Changing Times, Changing Literacies.* USA: International Reading Association.

Antón, M. & DiCamilla, F. (1998) Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54, 314–42.

Appleton, J. V. (1995) Analysing qualitative interview data: addressing issues of validity and reliability.' *Journal of Advanced Nursing*, 1995, 22, 993-997.

Archer, A. & Breuer, E. (2015) *Multimodality in Writing (Studies in Writing).* BRILL, in press.

Arndt, V. (1987) Six writers in search of texts: A protocol-based study of L1 and L2 writing. *ELT Journal*, 41/4 257-267.

Atkinson, D. (2010) Extended, Embodied Cognition and Second-language Acquisition. *Applied Linguistics 31/5*, 599-622.

Bachman, L. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bakhtin, M.M. (1981) *The Dialogic Imagination: Four Essays*. Michael Holquist. (ed) Trans. Caryl Emerson and Michael Holquist. Austin and London: University of Texas Press.

Barton, D. Ivanič, I (eds) (1991) *Writing in the community* (1–13) Newbury Park, CA: Sage Publications, Inc.

Barton, D. & Hamilton, M. (1998) *Local Literacies: Reading and Writing in One Community.* London: Routledge.

Barton, D. (1991) The social nature of writing. In Barton, D. & Ivanič, R. (eds) *Writing in the community* (1–13) Newbury Park, CA: Sage Publications, Inc.

Barton, D. Hamilton, M & Ivanič, R. (2000) *Situated Literacies: Reading and Writing in Context.* London: Routledge.

Bashir Ali (2006) Language Learning and the Definition of One's Social, Cultural, and Racial Identity. *TESOL Quarterly* Vol. 40, No. 3, September 2006.

Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R., & Mullet, J. (1988) Form and function in motor mimicry: Topographic evidence that the primary function is communicative. *Human Communication Research*, 14, 275-299.

Bayham, M. (1993) Code switching and mode switching: community interpreters and mediators of literacy. In Street, B. V. (ed) (1993) *Cross-cultural approaches to literacy*.

Cambridge: Cambridge University Press.

Baynham, M. & Prinsloo, M. (2009) Introduction. In *The Future of Literacy Studies*, (eds) M. Baynham & M. Prinsloo, Palgrave Macmillan, Basingstoke, 1-20.

Baynham, M. (1993) Code switching and mode switching: community interpreters and mediators of literacy. In Street, B. (ed) *Cross-cultural approaches to literacy*. Cambridge: Cambridge University Press.

Baynham, M. (1995) *Literacy Practices*. Harlow: Longman.

Baynham, M. Roberts, C. Cooke, M. Simpson, J. Ananiadou, K. Callaghan, J. McGoldrick, J. & Wallace, C. (2007) *Effective Teaching and Learning ESOL*. London: NRDC. Available online from http://www.nrdc.org.uk/uploads/documents/doc_3341.pdf

Bell, J.S. (1995) The relationship between L1 and L2 literacy: some complicating factors. *TESOL Quarterly*, 29 (4):691702 cited in Pitt, K. (2005) *Debates in ESOL Teaching and Learning, Culture, communities and classrooms*; Oxon: Routledge.

Berger, J. (1972) *Ways of Seeing*. London: Penguin books.

Berlo, D. (1960) *The Process of Communication: an Introduction to Theory and Practice*, New York: Holt Rhinehart and Wilson.

Bernard, H. R. (ed) (1998) *Handbook of methods in cultural anthropology*. Walnut Creek, CA: AltaMira.

Bernieri, F. J. (1988) Coordinated movement and rapport in teacher-students interactions. *Journal of Nonverbal Behavior*, 12, 120-138.

Bernstein, B. (1999) Vertical and horizontal discourse: an essay. *British Journal of Sociology of Education*, 20/2, 157-73.

Bezemer, J. and Mavers, D. (2011) Multimodal transcription as academic practice: a social semiotic perspective. In *International Journal of Social Research Methodology*, 14(3): 191-206.

Bhatt, I. (2014) *A sociomaterial account of assignment writing in Further Education classrooms*. PhD thesis. University of Leeds.

Bhatt, I. and Roock, R. (2014. Capturing the sociomateriality of digital literacy events. In *Research in Learning Technology*, Vol. 21.

Bhattacharya, R., Gupta, S., Jewitt, C., Newfield, D., Reid, Y., & Stein, P. (2007) The policypractice nexus in English classrooms in Delhi, Johannesburg, and London: Teachers and the textual cycle. *TESOL Quarterly* 41, 465–487.

Birdwhistell, R. L. (1952) *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture*. Washington, DC: Department of State, Foreign Service Institute.

Björkvall, A. & Karlsson, A. (2011) The materiality of discourses and the semiotics of materials: A social perspective on the meaning potentials of written texts and furniture. *Semiotica*, 187:1, 4, 141–165

Black, J. A., & Champion, D. J. (1976) *Methods and issues in social research*. New York: Wiley.

Blairy, S. Herrera, P. & Hess, U. (1999) Mimicry and the judgment of emotional facial expressions. *Journal of Nonverbal Behavior*, 23(1), Spring 1999.

Block, D. (2007) *Second-language Identities*, London: Continuum International Publishing Group Ltd.

Brandt, D. & Clinton, K. (2002) Limits of the local: Expanding perspectives on literacy as a

social practice. *Journal of Literacy Research*, 34(3), 337-356.

Brewer, J. (2000) *Ethnography*, Buckingham: Open University Press.

Brookes, A. & Grundy, P. (2002) *Beginning to Write*, 4th edition. Cambridge: Cambridge University Press.

Brown, P. & Levinson, S. C. (1978) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bruner, J. (1996) *The Culture of Education*. USA: Harvard University Press.

Bruner, J. S. (1973) *The relevance of education*. New York: Norton.

Bruner, J. S. (1978) The role of dialogue in language acquisition. In A. Sinclair, R., J. Jarvelle, and W. J.M. Levelt (eds) *The Child's Concept of Language.* New York: Springer-Verlag.

Bryman, A. (2006) Integrating quantitative and qualitative research: how is it done? *Qualitative Research*, Vol. 6, No. 1, 97-113 (2006).

Bryson, M. (1988) The Gaze in the Expanded Field. In *Vision and Visuality*. Foster, H. (ed) Seattle, WA: Bay View Press/Dia Art Foundation, 91-94.

Bourdieu, P. (1986) The forms of capital, reprinted in S. Ball (ed) The Routledge Falmer Reader in *Sociology of Education.* London: RoutledgeFalmer, 15-29.

Burgoon, J.K. & Saine, T. (1978) *Unspoken Dialogue: An Introduction to Nonverbal Communication*. Boston: Houghton Mifflin.

Burn, A. & Durran, J. (2007) *Media Literacy in Schools: Practice, Production and Progression* London: Paul Chapman Educational Publishing.

Burrell, M. (2005) The Efficacy of Computer-Activated Tasks (CATs) for Teaching E-Literacy in the ESOL Context. *Journal of eLiteracy*, Vol 2 (2005) 42-62.

Burton, N. Brundrett, M. & Jones, M. (2008) *Doing Your Education Research Project.* London: Sage Publications Ltd.

Bynner, J. (2007) A trajectory of disadvantage. In *Reflect*, Issue 9, November 2007. NRDC: London. Available online at: http://www.nrdc.org.uk/uploads/documents/doc_3840.pdf

Canale, M. & Swain, M. (1980) Theoretical bases of communicative approaches to second-language teaching and testing. *Applied Linguistics 1*, 1-47.

Carney, R. N. & Levin, J. R. (2002) Pictorial Illustrations Still Improve Students Learning from Text. Springer Netherlands: *Educational Psychology Review*. Vol 14, No. 1 / March, 2002, 5-26.

Chapelle, C. (2004) Technology and second-language learning: Expanding methods and agendas. *System*, 32, 593-601.

Chen, H. (2007) Reading authentic EFL text using visualization and advance organizers in a multimedia learning environment. *Language Learning & Technology*, Vol.11, No.3, October 2007, 83-106. [Accessed online 15/10/07 from http://llt.msu.edu/vol11num3/linchen/]

Cheng, W. (2012) *Exploring Corpus Linguistics: Language in Action*. London: Routledge.

Clark, A. (2001) *Mindware: An Introduction to the Philosophy of Cognitive Science.* Oxford: Oxford University Press.

Clark, J. M. & Paivio, A. (1991) *Dual coding theory and education*. Educational Psychology Review, 3(3), 149-170.

Cole, M. (1996) *Cultural Psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.

Cook, G. (2003) *Applied Linguistics*. Oxford: Oxford University Press.

Cook, G. (2011) Discourse Analysis. In Simpson, J. (ed) (2011) *The Routledge Handbook in Applied Linguistics*. New York and London: Routledge.

Cooke, M. & Simpson, J. (2008) *Understanding ESOL*. Oxford: Oxford University Press.

Cooke, M. Roberts, C. Campbell, S. & Stenhouse, J. (2007) *Frequently Asked Questions: the Great British Job Interview.* Kings' College: London. No publication available at time of writing. For further information see http://www.kcl.ac.uk/phpnews/wmview.php?ArtID=2100

Cope, B. & Kalantzis, M. (1996) A Pedagogy of Multiliteracies. In *Harvard Education Review*. Volume 66 Number 1. [Accessed online 30/07/07 from: http://l-by-d.com/Multiliteracies%20HER%20Vol%2066%201996.pdf]

Cope, B. & Kalantzis, M. (eds) (2000) *Multiliteracies: literacy learning and the design of social futures*. London & New York: Routledge.

Cornuejols, C. (1991) The Meanings of Images Across Cultures. International Conference on Hypermedia & Interactivity in Museums. *Archives & Museum Informatics*. [Accessed online 07/08/07 from http://www.archimuse.com/publishing/hypermedia/hypermedia.Ch23.pdf]

Crawford-Thomas, A. and Bloxham, J. (2015) *Five Reasons Why You Should Go Digital*. [Accessed online 01/07/15 at https://www.jisc.ac.uk/blog/five-reasons-why-you-should-do-digital-01-may-2015.]

Creese, A. & Blackledge, A. (2010) Translanguaging in the Bilingual Classroom: A Pedagogy for Learning and Teaching? *The Modern Language Journal*, 94, i, 103–115.

Crook, C. (1994) *Computers and the Collaborative Experience of Learning.* London & New York: Routledge.

Crowther, J. Hamilton, M & Tett, L. (2001) *Powerful Literacies*. NIACE: Leicester.

Currie, P. & Cray, E. (2004) ESL literacy: language practice or social practice? *Journal of Second-language Writing* 13 (2004) 111–132.

Danesi, M. (1990) *Thinking is seeing: Visual metaphors and the nature of abstract thought.* Semiotica 1990; 80 (3/4): 221–37.

Danesi, M. (2000) *Semiotics in Language Education.* Berlin: Mouton de Gruyter.

Debes, J. (1969) The loom of visual literacy. *Audiovisual Instruction*, 14(8), 25-27.

Degener, S. (2001) Making Sense of Critical Pedagogy in Adult Literacy Education. Volume 2: Chapter Two. In *Review of Adult Literacy and learning.* NCSLL. [Accessed online 05/07/08 from http://www.ncsall.net/?id=562]

Denny, K. Harmon, C. & Redmond, S. (2000) Functional Literacy, Educational Attainment and Earnings: Evidence From the International Adult Literacy Survey. [Accessed online 30/10/07 from http://www.ifs.org.uk/wps/wp0009.pdf]

Department of Education and the Arts. [Accessed online 01/09/07 from http://www.trainandemploy.qld.gov.au/resources/about_us/pdf/dest2010.pdf]

DfEE (2001) *Skills for life: The national strategy for improving adult literacy and numeracy skills*, Nottingham: DfEE Publications.

DfEE (Department for Education and Employment) (1999) *A fresh start – Improving literacy*

and numeracy (Moser report) London: Basic Skills Agency.

DfEE. (2000) *Breaking the language barriers: The report of the working group on English for speakers of other languages (ESOL) London: Basic Skills Agency.*

DfES (2002) Update 2 DfES, Winter 2002: 8.

DfES (2003) *21st Century Skills: Realising Our Potential.* Available at http://www.dfes.gov.uk/skillsstrategy/uploads/documents/21st%20Century%20Skills.pdf

DfES (Department for Education and Skills) (2001) *Adult ESOL core curriculum.* London: Basic Skills Agency/DfES.

Dittman, A.T. (1987) The Role of Body Movement in Communication. In A. W. Siegman and Stanley Feldstein (eds) *Nonverbal Behaviour and Communication*, New Jersey: Lawrence Erlbaum Associates.

Donato, R. (1994) Collective scaffolding in second-language learning, in J.P. Lantolf and G. Appel (eds) *Vygotskian Approaches to Second-language Research*, 33-56. Norwood, NJ: Ablex.

Dondis, D. (1973) *A Primer of Visual Literacy*. Cambridge, MA: MIT Press.

Donnelly, C. (2004) Constructing the ethos of tolerance and respect in an integrated school: the role of teachers. *British Educational Research Journal Vol. 30, No. 2, April 2004.*

Doughty, C. and Long, M. (2003) The scope of inquiry and goals of SLA. In Doughty C. and M. Long (eds): *Handbook of Second-language Acquisition*. Blackwell, 3–16.

Ecclesfield, N. (2013) *Tackling the lack of FE research: Technology in FE and Skills, on location at the ALT 20th anniversary conference (September 2013)* [Accessed online 12/07/15 from http://lsect.co.uk/Tech-supplement-2013.pdf]

Eggins, S. (2004) *An Introduction to Systemic Functional Linguistics*, 2nd edition. London: Continuum International Publishing Group.

Ehlers-Zavala, F. (2005) *Bilingual Reading from a Dual Coding Perspective.* Proceedings of the 4th International Symposium on Bilingualism, (ed) Cohen, J. & McAlister, K. & Rolstad, K. & MacSwan, J. 656-662. Somerville, MA: Cascadilla Press. [Accessed online 02/08/07 from www.cascadilla.com/isb4.html]

Eisner, E. & Peshkin, A. (eds) (1990) *Qualitative Inquiry in Education: The Continuing Debate*. New York: Teachers College Press.

Ellis, R. (1994) *The study of second-language acquisition*, Oxford: Oxford University Press.

Elsom-Cook, M. (2001) *Principles of Interactive Multimedia.* Berkshire: McGraw-Hill Publishing Company.

Esienhart, M. (2001) Educational Ethnography Past, Present, and Future: Ideas to Think With. *Educational Researcher,* Vol. 30. No. 8, 16–27.

FELTAG (2013) *Paths forward to a digital future for Further Education and Skills: Recommendations for the Minister of State for Skills and Enterprise*, Matthew Hancock MP. [Accessed online 02/06/15 from http://feltag.org.uk/wp-content/uploads/2012/01/FELTAG-REPORT-FINAL.pdf].

Fenwick, T. J. & Edwards, R. J. (2010) *Actor-network theory in education*. London: Routledge.

Fenwick, T. J., Edwards, R. & Sawchuk, P. (2011) *Emerging approaches to educational research: tracing the sociomaterial.* London: Routledge.

Fenwick, T. J. & Edwards, R. J. (2012) *Researching education through actor-network theory.*

Malden, MA: John Wiley & Sons.

Fernández Dobao, A. & A. Blum. (2013) Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, 41/2, 365-378.

Fernández Dobao, A. (2012) Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second-language Writing*, 21/1, 40-58.

Fernández Dobao, A. (2014a) Attention to form in collaborative writing tasks: Comparing pair and small group interaction. *Canadian Modern Language Review*, 70/2, 158-187.

Fernández Dobao, A. (2014b) Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, 18/4, 497-520.

Fernández Dobao, A. (2015) In press. Peer interaction and learning: A focus on the silent learner. In M. Sato and S. Ballinger (eds) *Peer interaction and second-language learning: pedagogical potential and research agenda*. John Benjamins.

Ferris, D. (2003) *Response to student writing: Implications for second-language students*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fischman, G. (2001) Reflections about images, visual culture, and educational research. In *Educational Researcher*, 30 (8), 28-33. [Accessed online 01/10/07 from http://www.aera.net/uploadedFiles/Journals_and_Publications/Journals/Educational_Researcher/3008/AERA3008_RNC_Fischman.pdf]

Fletcher, J.D. & Tobias, S. The Multimedia Principle. In Mayer, R. E. ed (2005) *The Cambridge Handbook of Multimedia Learning*, New York: Cambridge University Press.

Flewitt, R. Hampel, R. Hauck, M. & Lancaster, H. (2009) What are multimodal data and transcription? In Jewitt, C. (ed) (2009) *The Routledge Handbook of Multimodal Analysis.* Oxon: Routledge, 40-53.

Freire, P. & Macedo, D. (1986) *Literacy: Reading the Word and the World.* Westport: Greenwood Press.

Garćia, O. (2007) Foreword. In S. Makoni & A. Pennycook (eds) *Disinventing and reconstituting languages* (xi–xv) Clevedon, UK: Multilingual Matters.

García, O. (2009a) *Bilingual education in the 21st Century: A global perspective*. Oxford: Wiley-Blackwell.

García, O. (2009b) Education, multilingualism and translanguaging in the 21st century. In A.K. Mohanty, M. Panda, R. Phillipson and T. Skutnabb-Kangas (eds) *Multilingual education for social justice: globalising the local*. New Delhi: Orient BlackSwan.

Gardner, R. and Levy, M. (2010) The coordination of talk and action in the collaborative construction of a multimodal text. *Journal of Pragmatics*, 42 (2010) 2189–2203.

Garfinkel, H. (1967) *Studies in Ethnomethodology*. Englewood Cliffs NJ: Prentice Hall.

Gee, J. P. (1990) *Social Linguistics and Literacies: Ideology in Discourses*, 2nd edition, London: Falmer.

Gee, J. P. (2004) *Situated Language and Learning. A critique of traditional schooling.* London & New York: Routledge.

Geertz, C. (1976) From the native's point of view: On the Nature of Anthropological Understanding. In K. Basso and H. Selby, (eds) Meaning in Anthropology, Albuquerque: University of New Mexico Press.

Geisler, C. & Slattery, S. (2007). Capturing the activity of digital writing: Using, analyzing, and supplementing video screen capture. In McKee, H. A. & Devoss, D. N. N. (eds) *Digital writing research: technologies, methodologies, and ethical issues*. Cresskill, N.J.:

Hampton Press. 185-200.

Gillies, R.M. (2008) Teachers' and students' verbal behaviours during cooperative learning. In R.M. Gillies, A. Ashman,& J. Terwel (eds) *The teacher's role in implementing cooperative learning in the classroom* (243–262) Brisbane: Springer.

Gittens, R. (1999) *An Introduction to Literacy Teaching*. Basic Skills Agency: London.

Glesne, C. & Peshkin, P. (1992) *Becoming qualitative researchers: An introduction.* New York, NY: Longman.

Goffman, E. (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press

Goffman, E. (1983) The Interaction Order. *American Sociological Review*, 48:1-17.

Goffman, E. (1961) *Encounters: Two studies in the sociology of interaction*. Indianapolis, IN: Bobbs-Merrill.

Golafshani, N. (2003) Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, Vol.8 No.4 December 2003, 597-607. [Accessed online 01/07/08 from http://www.nova.edu/ssss/QR/QR8-4/golafshani.pdf]

Goodfellow, R. & Lea, M. (eds) *Literacy in the Digital University: Critical Perspectives on Learning, Scholarship & Technology*. London: Routledge.

Goodwin, C. (1981) *Conversational Organization: Interaction between Speakers and Hearers*. New York: Academic Press.

Goodwin, C. (2000a) Action and embodiment within situated human interaction. *Journal of Pragmatics*, Vol. 32, (10) 1489-1522 (34).

Goodwin, C. (2000b) Practices of seeing: Visual analysis: An ethnomethodological approach. In van Leeuwen, T. & Jewitt, C. *Handbook of Visual Analysis.* London: Sage Publications.

Goodwin, C. (2003) The Semiotic Body in its Environment. In Coupland, J. & Gwyn, R. (ed) *Discourse, the Body, and Identity*. Hampshire: Palgrave Macmillan.

Goodwin, C. (2007a) Participation, stance and affect in the organization of activities. *Discourse & Society*, Vol 18 (1) 53–73.

Goodwin, C. (2007b) Interactive footing. In E. Holt & R. Clift (eds) *Reporting talk: Reporting speech interaction* (16-46) Cambridge: Cambridge University Press.

Goody, J. (ed) (1968) *Literacy in traditional societies.* Cambridge: Cambridge University Press.

Gorad, S. Rushford, K. & Taylor, C. (2004) Is There a Shortage of Quantitative Work in Education Research? *Oxford Review of Education*. Vol. 30, No. 3, September 2004.

Gorlée, D. (1994) *Semiotics and the Problem of Translation. With Special Reference to the Semiotics of Charles S. Pierce.* Netherlands: Approaches to Translation Studies.

Gourlay, L. Hamilton, M. & Lea, M. R. (2013) Textual practices in the new media digital landscape: messing with digital literacies. *Research in Learning Technology*, 21(4).

Gourlay, L. & Oliver, M. (2013) Beyond 'the social': digital literacies as sociomaterial practice. In Goodfellow, R. & Lea, M. (eds) *Literacy in the Digital University: Critical Perspectives on Learning, Scholarship & Technology*. London: Routledge. 79-94.

Gourlay, L. (2014) Creating time: students, technologies and temporal practices in higher education. *E-Learning and Digital Media*, 11 (2). 141-153.

Grabe, W. & Kaplan, R. B. (1996) *Theory and practice of writing: An applied linguistic perspective*, New York: Longman.

Graddol, D. (2006) *English Next*. British Council: London. Available online at:

http://www.britishcouncil.org/learning-research-english-next.pdf

Gregory, R. L. (1994) *Eye and Brain*, 4th edition. New York: Oxford University Press. First edition Weidenfeld and Nicolson, 1996.

Grosjean, F. (1982) *Life with Two Languages. An Introduction to Bilingualism.* USA: Harvard University Press.

Guba, E. G. & Lincoln, Y. S. (1981) *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.

Guerts, K. (2002) *Culture and the senses: Bodily ways of knowing in an African community.* California: University of California Press.

Gullberg, M. (2011) Multilingual Multimodality: Communicative Difficulties and Their Solutions in Second-Language Use. In *Embodied Interaction: Language and Body in the Material World.* Jürgen Streeck, Charles Goodwin, and Curtis D. LeBaron (eds) New York: Cambridge University Press.

Gulson, K. & Symes, C. (2007) Knowing One's Place: Space, theory, education. *Critical Studies in Education*, 48 (1), 97-110.

Gumperz, J. (1982) *Discourse Strategies*. Cambridge: Cambridge University Press.

Hall, E. T. (1963) A System for the Notation of Proxemic Behavior. *American Anthropologist*, 65 (5): 1003–1026.

Hall, E.T. (1959) *The Silent Language*. New York: Anchor Books.

Halliday, M. A. K. (1978) *Language as a Social Semiotic System*, London: Arnold.

Hammersley, M. (1987) Some notes on the terms 'validity' and 'reliability'. *British Educational Research Journal, 13*(1), 73-81.

Hammond, J.A., Bithell, C.P., Jones, L., Bidgood, P. (2010) A first year experience of student-directed peer-assisted learning. *Active Learning in Higher Education*, 11(3), 201-212.

Hartsuiker, R. J. Pickering, M. J. Veltkamp, E. (2004) Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science*, 15 (6), 409–414.

Hawisher, G. & Selfe, C. (2000) *Global Literacies and the World-Wide Web*. London: Routledge.

Hayes, J & Flower, L. (1980) Identifying the organization of writing processes. In Gregg, L. W. & Steinberg, E. R. (eds) *Cognitive Processes in Writing*. Hillsdale: Erlbaum, 3-30

Heath, C. Hindmarsh, J. & Luff, P. (2010) *Video in Qualitative Research: Analysing Social Interaction in Everyday Life.* Sage: Los Angeles, CA.

Heinich, R. Molenda, M. Russell, J. (1995) *Instructional media and the new technology of instruction*, 5th edition. New York: Wiley.

Hiippala, T. (2016) *The Structure of Multimodal Documents*: *An Empirical Approach (Routledge Studies in Multimodality)* London & New York: Routledge, in press.

Hodkinson, P. & Hodkinson, H. (2001) *The Strengths and Limitations of Case Study Research.* Paper presented to the Learning and Skills Development Agency conference Making an Impact on Policy and Practice, Cambridge, 5-7 December 2001.

Holly, H.P. Pickering, M.J. Pearson, J. McLeane, J. F. (2010) Linguistic alignment between people and computers. *Journal of Pragmatics.* Vol. 42, (9), 2355–2368.

Holstein, J. A. & Gubrium, J. F. (2004) The active interview. Silverman, D. (ed) *Qualitative*

Research: Theory, Method and Practice. London: Sage Publications Ltd.

Horn, R. (2001) Visual Language and Converging Technologies in the Next 10-15 Years (and Beyond) A paper prepared for the National Science Foundation Conference on Converging Technologies (Nano-Bio-Info-Cogno) for Improving Human Performance Dec. 3-4, 2001. [Accessed online 10/09/07 from http://www.stanford.edu/~rhorn/a/recent/artclNSFVisualLangv.pdf]

Hornberger, N. H. (2002) Multilingual language policies and the continua of biliteracy: An ecological approach.' *Language Policy*, 1, 27–51.

Hornberger, N. H. (2005) Opening and filling up implementational and ideological spaces in heritage language education. *Modern Language Journal*, 89, 605–609.

Hornberger, N. H. (2008) Continua of biliteracy. In A. Creese, P. Martin, & N. H. Hornberger (eds), *Encyclopedia of language and education: Vol. 9. Ecology of language* (2nd ed, 275–290) Boston: Springer Science+Business Media.

Hunston, S. (2002) *Corpora in Applied Linguistics.* Cambridge: Cambridge University Press

Hyland, K. (2002) *Teaching and Researching Writing*; London: Pearson Education.

Hyland, K. (2011) Learning to write: issues in theory, research, and pedagogy. In R. Mancho´n (ed) *Learning-to-write and writing-to-learn in an additional language* (18–35) Amsterdam: John Benjamins.

Hymes, D. (1964) Introduction: Toward Ethnographies of Communication. In *American Anthropologist*, 66 (6): 1–34.

Hymes, D. (1972) On communicative competence. In J. B. Pride, & J. Holmes (eds) *Sociolinguistics*. Harmondsworth: Penguin

Ice, G. H. (2004) Technological Advances in Observational Data Collection: The advantages and limitations of Computer-Assisted Data Collection. In *Field Methods*, Vol. 16, No. 3, 352-374.

Indices of Deprivation (2010) [Accessed online 15/04/2011 from: http://www.neighbourhood.statistics.gov.uk/dissemination/NeighbourhoodSummary.do?a=3&b=276810&c=LS7+3BG&g=384094&i=1001x1012x1013&j=296039&m=1&p=2&q=1&r=0&s=1302858798125&enc=1&tab=1&inWales=false]

Ivanič, R. & Moss, W. (1991) Bringing community writing practices into education. In D. Barton & R. Ivanič (eds) *Writing in the community* (193–223) Newbury Park, CA: Sage Publications, Inc.

Jamieson, H. (2007) *Visual Communication: More than meets the eye*, Bristol: Intellect Books.

Jefferson, G. (1972) Side sequences. In D. Sudnow (ed) *Studies in social interaction* (294–338) New York: Free Press.

Jefferson, G. (1986) Notes on 'latency' in overlap onset. *Human Studies*, 9, 153-183.

Jewitt, C. (2006) *Technology, Literacy and Learning. A multimodal approach.* Oxon: Routledge.

Jewitt, C. (ed) (2009) *The Routledge Handbook of Multimodal Analysis.* Oxon: Routledge.

Johnson, D.W., Johnson, R.T., Smith, K. (2007) The state of cooperative learning in post-secondary and professional settings. *Educational Psychology Review*, 19, 15-29.

Johri, A. (2011) The socio-materiality of learning practices and implications for the field of learning technology. *Research in Learning Technology*, 19(3), 207-217.

Kalantzis, M. Cope, B. & Slade, D. (1989) *Minority Languages and Dominant Culture: Issues*

of Education, Assessment and Social Equity, Sussex: Falmer Press.

Kaplan, R. (1966) Cultural thought patterns in inter-cultural education. *TESOL Quarterly*, 1, 10-16.

Kellner, D. (2000) *Critical Perspectives on Visual Imagery in Media and Cyberculture*. [Accessed online 02/07/07 from http://www.gseis.ucla.edu/faculty/kellner/ed270/VISUALLITcritical.htm]

Kendon, A. (1967) Some functions of gaze-direction in social interaction. *Acta Psycholigica*, 26: 22-63.

Kenning, M. (2006) Evolving concepts and moving targets: communicative competence and the mediation of communication. In *International Journal of Applied Linguistics*. Vol. 16. No. 3. 363-368.

Kerlinger, F. (1964) *Foundations of behavioural research*. New York: Holt.

Kessler, G. (2009) Student-initiated attention to form in wiki-based collaborative writing. Language, Learning & Technology, 13, 79-95.

Kist, W. (2005) *New literacies in action, teaching and learning in multiple media.* New York: Teachers College Press.

Knoblauch, H. & Tuma, R. (2011) Videography: an interpretive approach to video-recorded micro-social interaction. In Margolis M and Pauwels L (eds) *The Sage Handbook of Visual Meth*ods. Thousand Oaks, CA: Sage, 414–430.

Knoblauch, H. (2012) Introduction to the special issue of Qualitative Research: video-analysis and videography. *Qualitative Research*, 12(3) 251–254.

Knoblauch, H. René, T. & Bernt, S. (2015) *Videography Introduction to Interpretive Videoanalysis of Social Situations.* Peter Lang Publishing.

Knoblauch, H. Schnettler, B. & Raab, J. (eds) (2009) *Video Analysis-Methodology and Methods: Qualitative Audiovisual Data Analysis in Sociology*. Peter Lang Publishing.

Krashen, S. & Terrell, T.D. (1983) *The Natural Approach, Language Acquisition in the Classroom.* London: Prentice Hall Europe.

Krashen, S. (1985) *The input hypothesis: Issues and implications*. California: Laredo Publishing Co Inc.

Kress, G & van Leeuwen, T. (2006) *Reading Images, the grammar of visual design*, 2nd edition, Oxon: Routledge. First edition 1996.

Kress, G. & van Leeuwen, T. (2001) *Multimodal Discourse: the modes and media of contemporary communication.* London: Arnold.

Kress, G. (2000a) Multimodality. In Cope, B. & Kalantzis, M. (eds) (2000) M*ultiliteracies: literacy learning and the design of social futures.* London & New York: Routledge.

Kress, G. (2000b) Design and transformation: new theories of meaning. In Cope, B. & Kalantzis, M. (eds) (2000) *Multiliteracies: literacy learning and the design of social futures.* London & New York: Routledge.

Kress, G. (2003) *Literacy in the New Media Age*. London & New York: Routledge.

Kress, G. (2009) What is mode? In Jewitt, C. (ed) (2009) *The Routledge Handbook of Multimodal Analysis.* Oxon: Routledge, 54-67.

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

Kress, G. Jewitt, C. Bourne, J. Franks, A. Hardcastle, J. Jones, K & Reid, E. (2005) *English in*

Urban Classrooms. A multimodal perspective on teaching and learning. Oxon: Routledge.

Lankin, J. L. & Chartrand, T. L. (2003) Using nonconscious behavioral mimicry to create affiliation and rapport. In *Psychological Science*, 14, 334-339.

Lankin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003) The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. In *Journal of Nonverbal Behavior*, 27, 145-162.

Lankshear, C. & Knobel, M. (2006) *New Literacies. Everyday Practices and Classroom Learning*, 2nd Edition. Berkshire: Open University Press.

Law, J. (2009) Actor Network Theory and Material Semiotics. In Turner, R, B. S. (ed) *The new Blackwell companion to social theory.* Oxford: Wiley-Blackwell. 141-158.

Lave, J. & Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press: Cambridge.

Leininger, M. (1994) Evaluation criteria and critique of qualitative research studies. In J. M. Morse (ed) *Critical Issues in Qualitative Research Methods*. Newbury Park, CA: Sage.

Lemke, J. L. (2009) Across the Scales of Time: Artefacts, Activities, and Meanings in Ecosocial Systems. In *Mind, Culture, and Activity*, 7: 4, 273 - 290.

Levin, J. Kim, H. and Riel. M. (1990) Analyzing instructional interactions on electronic message networks, in Harasim, L. (ed) *Online Education: Perspectives on a new environment*. New York: Praeger.

Levinson, S. (1992) Activity types and language. In Drew, P. & Heritage, J. (eds) *Talk at work: Interaction in institutional setting*, 66-100. Cambridge: Cambridge University Press.

LeVine, P. & Scollon, R. (eds) (2004) *Discourse and technology: multimodal discourse analysis.* Washington: Georgetown University Press.

Lier, v. L. (2004) *The Ecology and Semiotics of Language Learning. A sociocultural Perspective.* USA: Kluwer Academic Publishers.

Lin, O. P. & Maarof, N. (2013) Collaborative Writing in Summary Writing: Student Perceptions and Problems. In *Procedia - Social and Behavioral Sciences.* Vol. 90, 599-606. 6th International Conference on University Learning and Teaching (InCULT 2012)

Lincoln, Y. S. & Guba, E. G. (1985) *Naturalistic inquiry*. Beverley Hills, CA: Sage.

Lowe, R. (1993) *Successful Instructional Diagrams.* London: Kogan Page.

Luckin, R. (2006) Understanding Learning Contexts as Ecologies of Resources: From the Zone of Proximal Development to Learner Generated Contexts. In T. Reeves & S. Yamashita (eds), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* 2006 (2195-2202) Chesapeake, VA: AACE.

Luckin, R. (2010) *Re-designing Learning Contexts. Technology-rich, learner-centred ecologies.* Oxon: Routledge.

Luke, A. &. Freebody, P. (2000) *Literate Futures: Report of the Literacy Review for Queensland State Schools*. Brisbane, State of Queensland (Department of Education and the Arts) [Accessed online 12/10/07 from http://education.qld.gov.au/curriculum/learning/literate-futures/pdfs/lf-review.pdf]

Luke, C. (2000) Cyber-schooling and technological change: Multiliteracies for new times. pp 69-91 in Cope, B. & Kalantzis, M. (eds) (2000) *Multiliteracies: literacy learning and the design of social futures*. London & New York: Routledge.

Machin, D. (2007) *Introduction to Multimodal Analysis.* London: Hodder Education.

Mackey, A. & Gass, S. M. (2005) *Second-language Research, Methodology and Design.* New Jersey: Lawrence AErlbaum.

Maoz, Z. (2002) Case study methodology in international studies: from storytelling to hypothesis testing. In F. P. Harvey and M. Brecher (eds) *Evaluating Methodology in International Studies*. University of Michigan Press: Ann Arbor.

Marita, A. (2000) Might and magic, lust and language – the eye as a metaphor in literature. In *Documenta Ophthalmologica*, 101**:** 223–232, 2000.

Martin, A. & Madigan, D. (eds) (2006) *Digital Literacies for Learning*. London: Facet Publishing.

Matsuda P.K. (2003) Changing currents in second-language writing research: A colloquium. In *Journal of Second-language Writing* 12 (2003) 151–179.

Matsuda, P.K. et al. (2003) Changing currents in second-language writing research: A colloquium. *Journal of Second-language Writing* 12 (2003) 151–179.

Mayer, R. E. (ed) (2005) *The Cambridge Handbook of Multimedia Learning*, New York: Cambridge University Press.

Mayer, R.E. (2000) The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction.* Vol 13 Is 2 pp125-139.

McDonough, K. & Sunitham, W. (2009) Collaborative Dialogue Between Thai EFL Learners During Self-Access Computer Activities. In *TESOL quarterly*, Vol. 43, No. 2.

McDonough, K. Crawford, W., & De Vleeschauwer, J. (2014) Summary writing in a Thai EFL university context. *Journal of Second-language Writing*, 24, 20-32.

McDonough, K. Crawford, W., & De Vleeschauwer, J. (2015) In press. Thai EFL learners' interaction during collaborative writing tasks and its relationship to text quality. In M. Sato & S. Ballinger (eds) *Peer interaction and second-language learning: Pedagogical potential and research agenda*. Amsterdam: John Benjamins.

McNeill, D. (ed) (2000) *Language and gesture*. Cambridge, England: Cambridge University Press.

Mellar, H. Kambouri, M. Betts, S. Nance, B. Moriarty, V. & Logan, K. (2007) *Effective Teaching and Learning: Using IT.* London: NRDC.

Mercer, N**.** Littleton, K. Wegerif, R. (2004) Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. *Technology, Pedagogy and Education,* 13, 2, 193-209.

Messaris, P. & Moriarty, S. (2005) Visual Literacy Theory. In Smith, K. (ed) (2005) *Handbook of Visual Communication*. New Jersey: Lawrence Erlbaum Associates.

Mills, K. (2006) Multiliteracies: A Critical Ethnography. pedagogy, power, discourse and access to multiliteracies. Doctoral Thesis. Queensland University of Technology. [Accessed online 21/11/07 from http://adt.library.qut.edu.au/adt-qut/public/adt-QUT20061003.145606/]

Mirzaei, A. & Eslami, Z. R. (2015) ZPD-activated languaging and collaborative L2 writing. In *Educational Psychology: An International Journal of Experimental Educational Psychology*, 35:1, 5-25.

Mises, L. (2008) *Human Action: A Treatise on Economics. The Scholar's Edition.* Mises Institute, Auburn, Ala. [Accessed online 20/04/2015 https://mises.org/system/tdf/Human%20Action_3.pdf?file=1&type=document]

Mondada, L. (2009) Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis. In Knoblauch, H. Schnettler, B. & Raab, J. (eds) (2009) *Video Analysis-Methodology and Methods: Qualitative Audiovisual Data Analysis in Sociology.* US: Peter Lang Publishing.

Morse, J. M. Barrett, M. Mayan, M. Olson, K. & Spiers, J. (2002) Verification Strategies for Establishing Reliability and Validity in Qualitative Research.' *International Journal of Qualitative Methods,* 1 (2) Spring 2002.

Mueller, A. and Oppenheimer, D.M. (2014) The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological Science* **25**(6): 1159-1168.

Negata, A. L. (2004) Promoting Self-Reflexivity in Intercultural Education. *Journal of Intercultural Communication*, No. 8, 139-167.

Nelson, G. L., & Carson, J. G. (1998) ESL students' perceptions of effectiveness in peer response groups. *Journal of Second-language Writing*, 7(2), 113–131.

Nelson, G. L., & Murphy, J. M. (1993) Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27(1), 135–141.

Neomy, S. (2005) Collaborative Writing: Product, Process, and Students' Reflections. In *Journal of Second-language Writing*, v14 n3, 153-173.

Neumann, R. & Strack, F. (2000) Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, 79, 211-223.

Nevile, M. (2015) The Embodied Turn in Research on Language and Social Interaction. In *Research on Language and Social Interaction*, 48(2), 121–151.

Newfield, D. (2013) Transduction, transformation and the transmodal moment. In Jewitt, C. (ed) *The Routledge Handbook of Multimodal Analysis*, 2nd edition. New York and London: Routledge. 100-114.

Norris, J. and Jones, R. H. (eds) (2005) *Discourse in Action: Introducing Mediated Discourse Analysis.* Routledge: Oxon.

Norris, S. (2004) *Analysing Multimodal Interaction: A Methodological Framework.* London: Routledge.

Norris, S. (2009) Modal density and modal configurations: multimodal actions. In Jewitt, C. (ed) (2009) *The Routledge Handbook of Multimodal Analysis.* Oxon: Routledge, 78-90.

Nunkoosing, K. (2005) The Problems With Interviews. *Qualitative Health Research.* Vol. 15 No. 5, May 2005, 698-706.

Nystrand, M., & Brandt, D. (1989) Response to writing as a context for learning to write. In C. M. Anson (ed) *Writing and response: Theory, practice, and research* (209–230) Urbana, IL: National Council of Teachers of English.

O'Donnell, A. M. (2006) The role of peers and group learning. In P. Alexander & P. Winne (eds) *Handbook of educational psychology*, 2nd Edition. Mahwah, NJ: Lawrence Erlbaum.

O'Halloran, K. L. & Smith, B. A. (eds) (2011) *Multimodal Studies: Exploring Issues and Domains*. New York & London: Routledge.

O'Halloran, K. L. (2005*) Mathematical Discourse: Language, Symbolism and Visual Images*. London and New York: Continuum. (reprinted 2008)

O'Halloran, K. L. (ed) (2004) *Multimodal Discourse Analysis*. London and New York: Continuum. (reprinted 2006)

Oliver, M. (2005) The problem with affordance. *E-Learning Journal*, 2 (4): 402-413.

Olson, D. R. (1994) *The world on paper, The conceptual and cognitive implications of writing and reading;* Cambridge: Cambridge University Press.

Ong, W. (1982) *Orality and literacy.* London: Methuen.

Pahl, K. & Rowsell, J. (eds) (2006) *Travel Notes from the New Literacy Studies*. Multilingual Matters Ltd: Cleveland.

Paivio, A. & Begg, I. (1981) *The Psychology of Language*. New York: Prentice-Hall.

Paivio, A. (1971) *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.

Paivio, A. (1986) *Mental Representations*. New York: Oxford University Press.

Papen, U. (2005) *Adult Literacy as Social Practice, More than Skills*; Oxon: Routledge.

Pelgrum, W. & Anderson, R. (eds) (2001) *IT and the Emerging Paradigm for Life-long learning.* IEA: Netherlands.

Petrie, G. (2003) ESL teachers' views on visual language: a grounded theory. *The Reading Matrix*, Vol.3. No.3, November 2003.

Philp, J, Adams, R & Iwashita, N. (2014) *Peer interaction and second language learning*. Taylor & Francis, New York.

Pickering, M. J. Garrod, S. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27 (2), 169–225.

Plass, J. & Jones, L. (2005) Multimedia Learning in Second-language Acquisition. In Mayer, R. (ed) (2005) *The Cambridge Handbook of Multimedia Learning*, New York: Cambridge University Press (pp 467-488).

Prinsloo, M. & Baynham, M. (2008) Introduction: renewing literacy studies. In *Literacies, Global and Local*, (eds) M. Prinsloo & M. Baynham, John Benjamins: Amsterdam.

Procter-Legg, E. (2013) *Tackling the lack of FE research: Technology in FE and Skills, on location at the ALT 20th anniversary conference (September 2013)* [Accessed online 12/07/15 from http://lsect.co.uk/Tech-supplement-2013.pdf]

Queensland Government (2002a) Literate Futures: Reading 1. Brisbane, State of Queensland (Department of Education and the Arts) [Accessed online 01/09/07 from http://education.qld.gov.au/curriculum/learning/literate-futures/pdfs/reading-part1.pdf]

Queensland Government (2002b) Literate Futures: Reading 2. Brisbane, State of Queensland (Department of Education and the Arts) [Accessed online 01/09/07 from http://education.qld.gov.au/curriculum/learning/literate-futures/pdfs/reading-part2.pdf]

Queensland Government (2002c) *Destination 2010: 2002–5* Brisbane, State of Queensland

Queensland Government (2002d) *Productive Pedagogies: Classroom Reflection Manual*. Brisbane, State of Queensland (Department of Education and the Arts) [Accessed online 01/09/07 from http://education.qld.gov.au/public_media/reports/curriculum-framework/productive-pedagogies/pdfs/prodpedpdf]

Rae, J. (2001) Organizing Participation in Interaction: Doing Participation Framework. *Research on Language and Social Interaction*, 34:2, 253-278.

Ramanathan, V; Atkinson, D. (1999) Ethnographic approaches and methods in L2 writing research: a critical guide and review. *Applied Linguistics*, Vol. 20, Number 1, March 1999, 44-70(27)

Rampton, M. B. H. (1990) Displacing the Native Speaker: Expertise, Affiliation and

Inheritance. *ELT Journal*, 44/2, 97–101.

Rassool, N. (1999) *Literacy for Sustainable Development in the Age of Information.* Clevedon, Avon and Philadelphia: Multilingual Matters Ltd.

Rassool, N. (1999) *Literacy for Sustainable Development in the Age Of Information*. US: Multilingual Matters.

Rawles, A. W. (2002) *Ethnomethodology's Program: Working Out Durkheim's Aphorism (Legacies of Social Thought Series) US:* Rowman & Littlefield Publishers.

Reid, J.M. (1993) Historical Perspectives on Writing and Reading in the ESL Classroom. In Carson, J.G. and Leki, I. *Reading in the Composition Classroom*; Boston: Heinle and Heinle, 33-60.

Richardson, L. (1997) *Fields of Play: Constructing an Academic Life*. New Brunswick, NJ: Rutgers University Press.

Roberts, C., M. Cooke, M. Baynham and J. Simpson (2007) Adult ESOL in the United Kingdom: Policy and research.'*Prospect* 22/3, 18-31

Rouhshad, A. Wigglesworth, G & Storch, N. (2015) The nature of face-to-face versus computer-mediated communications in pair interactions. *Language Teaching Research*, 1-21.

Roulston , L. deMarrais, K. & Lewis, J. B. (2003) Learning to Interview in the Social Sciences. *Qualitative Inquiry*, Vol 9, No. 4, 2003, 643-668.

Royce, T. & Bowcher, W. (eds) (2007) *New Directions in the Analysis of Multimodal Discourse*. US: Lawrence Erlbaum.

Royce, T. (2002) Multimodality in the TESOL Classroom: Exploring Visual-Verbal Synergy. In *TESOL Quarterly***,** Vol. 36, No. 2, Summer, 2002.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974) A simplest systematics for the organization of turn-taking for conversation. In *Language*, 50, 696-735.

Satchwell, C. Barton, D. & Hamilton, M. (2013) Crossing boundaries: Digital and non-digital literacy practices in formal and informal contexts in further and higher education. In Goodfellow, R. & Lea, M. (eds) *Literacy in the Digital University: Critical Perspectives on Learning, Scholarship & Technology*. London: Routledge. 42-55.

Sajedi, P. (2014) Collaborative Summary Writing and EFL Students' L2 Development. In *Procedia - Social and Behavioral Sciences*, 98, 1650 – 1657. International Conference on Current Trends in ELT.

Salomon, G. (2000) It's not just the tool, but the educational rationale that counts, (keynote address) Proceedings of ED-MEDIA 2000, Montreal. [Accessed online 01/09/07 from http://www.aace.org/conf/edmedia/00/salomonkeynote.htm]

Samovar, L. A. & Porter, A. S. (2001) *Communication Between Cultures* (4th ed) Belmont, USA: Wadsworth/Thomson Learning.

Sandelowski, M. (1986) The problems of rigor in qualitative research. In *Advances in Nursing Science*. 8(3), 27-37.

Sauro, S. (2011) SCMC for SLA: A research synthesis. *Calico Journal*, 28, 369-391.

Schegloff, E. (2007) *Sequence Organization in Interaction*. Cambridge University Press.

Scheurich, J. (1995) A Postmodernist Critique of Research Interviewing. *Qualitative Studies in Education 8(3): 239–52.*

Schiffrin, D. (1997) *Approaches to Discourse*. Oxford: Blackwell.

Schmidt, R. (1995) Consciousness and foreign language learning: A tutorial on attention and awareness in learning. In R. Schmidt, (ed) *Attention and awareness in foreign language learning* (1-63) Honolulu, HI: University of Hawai`i, National Foreign Language Resource Center.

Schmidt, R. (2010) Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010*, Singapore, December 2-4 (721-737) Singapore: National University of Singapore, Centre for Language Studies.

Schnotz, W. (2002) Towards an Integrated View of Learning From Text and Visual Displays. In *Educational Psychology Review,* Vol. 14, No. 1, March 2002.

Schramm, W. (1954) Procedures and Effects of Mass Communication, reprinted in Schramm (1965) *The Process and Effects of Mass Communication*, Urbana: University of Illinois.

Scollon, R. S. & Scollon, W. S. (2003) *Discourses in Place. Language in the Material World.* London: Routledge.

Seels, B. A. (1994) Visual literacy: the definition problem. In Moore, D & Dwyer, F (eds) *Visual Literacy: A Spectrum of Visual Learning*, (97-112) Englewood Cliffs, NJ: Educational Technology Publications.

SFA (2015) Skills Funding Agency letter to Peter Lauener. [Accessed online, 30/07/2015 from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/406881/Vince_Cable_and_Nick_Boles_to_Peter_Lauener_-__Skills_Funding_Agency.pdf]

Shehadeh, A. (2011) Effects and student perceptions of collaborative writing in L2. In *Journal of Second-language Writing*. Vol. 20, Issue 4, Dec. 2011, Pages 286–305.

Simpson, J. Cooke, M. Callaghan, J. Hepworth, M. Homer, M. Baynham, M. Allen, T. Grant, R. Sisimayi, S. *(2011b) ESOL Neighbourhood Audit Pilot (Harehills, Leeds*) HENNA project. [Accessed online 10/08/15 from http://www.education.leeds.ac.uk/research/projects/henna-project.]

Simpson, J. (2003) *The discourse of computer-mediated communication: A study of an online community*. University of Reading. School of Linguistics and Applied Language Studies. Unpublished PhD thesis.

Simpson, J. (2007) Adult ESOL in England: Policy, Practice and Research. In Faux, N (ed) *Low-educated Second-language and Literacy Acquisition: Research, policy and practice.* Richmond, VA: Commonwealth University of Virginia.

Simpson, J. (2011) Telling tales: Discursive space and narratives in ESOL classrooms. *Linguistics and Education*, 22, 10-22.

Simpson, J. (ed) (2011) *The Routledge Handbook in Applied Linguistics*. New York and London: Routledge.

Simpson, J. Cooke, M. & Baynham, M. (2008) *The right course? An exploratory study of placement practices in ESOL and Literacy.* London: NRDC.

Simpson, J. (forthcoming) English for Speakers of Other Languages (ESOL): Language education and migration. In Hall, G. (ed) *The Routledge Handbook of English Language Teaching*. London: Routledge.

Smeyers, P. (2001) Qualitative Versus Quantitative Research Design: A Plea for Paradigmatic Tolerance in Educational Research. *Journal Of Philosophy of Education*, Vol. 35, No. 3, 2001.

Smith, K. (2005) *Handbook of Visual Communication*. New Jersey: Lawrence Erlbaum Associates.

Snyder, L. (ed) (2002) *Silicon Literacies: Communication, Innovation and Education in the Electronic Age.* London: Routledge.

Steffensen, M.S. & Goetz, E.T. & Cheng, X. (1999) The images and emotions of bilingual Chinese readers: A dual coding analysis. *Reading Psychology, 20,* 301-324.

Stenbacka, C. (2001) Qualitative research requires quality concepts of its own. *Management Decision, 39*(7), 551-555.

Stevick, E. W. (1986) *Images and Options in the Language Classroom*. USA: Cambridge University Press.

Stokes, S. (2006) Visual Literacy in Teaching and Learning: A Literature Perspective, in *Electronic Journal for the Integration of Technology in Education*, vol. 1, no. 1. [Accessed online 22/07/07 from http://ejite.isu.edu/Volume1No1/pdfs/stokes.pdf]

Storch, N. & Aldosari, A. (2010) Learners' use of first-language (Arabic) in pairwork in an EFL class. *Language Teaching Research,* 14, 355–375.

Storch, N. & Wigglesworth, G. (2007) Writing tasks: The effects of Collaboration. In M. García Mayo (ed), *Investigating tasks in formal language learning* (157–177) Clevedon, UK: Multilingual Matters.

Storch, N. & Wigglesworth, G. (2010) Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second-language Acquisition,* 32, 303–334.

Storch, N. (1998) A classroom-based study: Insights from a collaborative text reconstruction task. *ELT Journal*, 52, 291–300.

Storch, N. (1999) Are two heads better than one? Pair work and grammatical accuracy. *System*, 27, 363–374.

Storch, N. (2001) How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research*, 5, 29–53.

Storch, N. (2002) Patterns of interaction in ESL pair work. *Language Learning*, 52, 119–158.

Storch, N. (2004) Using activity theory to explain differences in patterns of dyadic interactions in an ESL class. *The Canadian Modern Language Review*, 60, 457–480.

Storch, N. (2005) Collaborative writing: Product, process, and students' reflections. Journal of Second-language Writing, 14(3), 153–173.

Storch, N. (2007) Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research*, 11, 143–159.

Storch, N. (2008) Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness*, 17, 95–114.

Streeck, J, Goodwin, C & LeBaron, C. (eds) (2011) *Embodied Interaction: Language and Body in the Material World.* New York: Cambridge University Press.

Street, B. (1998) New Literacies in Theory and Practice. What are the Implications for Language in Education. *Linguistics and Education*, 10(1): 1-24.

Street, B. (ed) (1993) *Cross-cultural approaches to literacy.* Cambridge: Cambridge University Press.

Street, B. V. (1995) *Social Literacies, Critical Approaches to literacy in Development, Ethnography and education;* Harlow: Pearson Education Limited

Street, B. V. (ed) (1993) *Cross-cultural approaches to literacy*. Cambridge: Cambridge University Press.

Street. B. (1995) *Social Literacies, Critical Approaches to Literacy in Development, Ethnography and Education.* Essex: Pearson Education Limited

Swain, M. & Watanbe, Y. (2013) Languaging: Collaborative Dialogue as a Source of Second-language Learning. In *The Encyclopaedia of Applied Linguistics*, Carol A. Chapelle (ed) Blackwell Publishing Ltd.

Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. and Madden, C. (eds) *Input in Second-language Acquisition,* 235-256. New York: Newbury House.

Swain, M. (1995) Three functions of output in second-language learning. In Cook, G. and Seidelhofer, B. (eds) *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson,* 125-144. Oxford: Oxford University Press.

Swain, M. (1998) Focus on form through conscious reflection.' In C. Doughty & J. Williams (eds) *Focus on form in classroom second-language acquisition* (64–81) Cambridge: Cambridge University Press.

Swain, M. (2000) The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (Ed.), *Sociocultural theory and second-language learning* (97–114) Oxford: Oxford University Press.

Swain, M. (2006) Languaging, agency and collaboration in advanced second-language learning. In H. Byrnes (ed), *Advanced language learning: The contributions of Halliday and Vygotsky* (95–108) London, England: Continuum.

Swain, M. (2010) Talking-it through: Languaging as a source of learning. In R. Batstone (ed) *Sociocognitive perspectives on language use/learning* (112-130) Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1998) Interaction and second-language learning: Two adolescent French immersion students working together. The Modern Language Journal, 82(3), 320–337.

Swain, M., & Lapkin, S. (2001) Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (eds) *Researching pedagogic tasks: Second-language learning, teaching and testing* (99–118) London: Longman.

Swain, M., & Lapkin, S. (2002) Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285–304.

Swain, M., Lapkin, S., Knouzi, I., Suzuki, W., & Brooks, L. (2009) Languaging: University students learn the grammatical concept of voice in French. *The Modern Language Journal*, 93, 5–29.

Synder, L. (2008) *The Literacy Wars,* Australia: Allen & Unwin.

Tannen, D. (1989) *Talking Voices. Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Tannen, D. (2005) *Conversational Style, Analyzing Talk Among Friends*. USA: Oxford University Press.

Taura, H. (1998) Bilingual Dual Coding in Japanese Returnee Students. In *Language, Culture and Curriculum.* Vol. 11, No. 1, 1998.

TLang project (2014-2018) Translation and Translanguaging. Investigating Linguistic and Cultural Transformations in Superdiverse Wards in Four UK Cities. For further information see: http://www.birmingham.ac.uk/generic/tlang/about/index.aspx

Tomlinson, P. (2005) *Can They Really Mean That? Mixing up Psychology, Epistemology and Ontology…?* University of Leeds: School of Education.

Trager, G. L. (1958) Paralanguage: A first approximation. *Studies in Linguistics*, 13, 1-12.

Tunbridge, M. (1995) The cyberspace cowboy. In *Australian Personal Computer*, December: 2-4.

Tusting, K. Barton, D. McCulloch, S. Hamilton, M. & Bhatt, I. (2015) The dynamics of knowledge creation. Academics' writing practices in the contemporary university workplace. For further information see http://wp.lancs.ac.uk/acadswriting/.

Ulhirová, L. (1994) On the role of the PC as a relevant object in face-to-face communication. In *Journal of Pragmatics*, 22, 511-527.

Unsworth, L. (2001) *Teaching Multiliteracies Across the Curriculum: Changing Contexts of Text and Image in Classroom Practice.* Buckingham: Open University Press.

Uzawa, K. (1996) Second-language Learners' Processes of L1 Writing, L2 Writing, and Translation from L1 into L2, *Journal of Second-language Writing* 5 (1996) 271294.

Valsiner, J. (ed) (1986) *The Individual Subject and Scientific Psychology*. New York: Plenum.

van Leeuwen, T. (2011) Multimodality. In Simpson, J. (ed) (2011) *The Routledge Handbook in Applied Linguistics*. New York and London: Routledge.

Villamil, O. S., & Guerrero, M. C. M. de. (1998) Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, 19, 491–514.

Vygotsky, L. S. (1978) *Mind in Society: The development of higher psychological processes*. Harvard University Press: London.

Walker, J. & Chaplin, S. (1997) *Visual culture, An introduction*. Manchester: Manchester University Press.

Walther, J. B. & Tidwell, L. C. (1995) Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. In *Journal of Organizational Computing*, 5:4, 355-378.

Wang, Y. & Guo, M. (2014) A Short Analysis of Discourse Coherence. *Journal of Language Teaching and Research*, Vol. 5, No. 2, 460-465, March 2014.

Warschauer, M. (1999) *Electronic literacies: Language, culture, and power in online education.* Lawrence Erlbaum Associates, Publishers: Mahwah, NJ.

Warschauer, M. (2006) Literacy and technology: Bridging the divide. In D. Gibbs & K.-L. Krause (eds) *Cyberlines 2: Languages and cultures of the Internet* (163-174) Albert Park, Australia: James Nicholas.

Watanabe, Y. & Swain, M. (2007) Effects of proficiency differences and patterns of pair interaction on second-language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11, 121–142.

Watanabe, Y. (2008) Peer-peer interaction between L2 learners of different proficiency levels: Their interactions and reflections. *Canadian Modern Language Review*, 64, 605-635.

Weinstein-Shr, G. (1993) Literacy and Social Process: A Community in Transition. In Street, B. V. (1993) *Cross-cultural approaches to literacy*. Cambridge: Cambridge University Press.

Wenger, E. (1998) *Communities of Practice: Learning, Meaning, and Identity*, Cambridge: Cambridge University Press.

Wertsch, J.V. (1998) *Mind as Action.* New York: Oxford University Press.

Widdowson, H. G. (2003) *Defining issues in English language teaching*. Oxford: Oxford University Press

Wigglesworth, G., & Storch, N. (2009) Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26, 445–466.

Wilden, A. (1987) *The Rules Are No Game: The Strategy of Communication*. London: Routledge & Kegan Paul

Williams, J. (2012) The potential role(s) of writing in second-language development. *Journal of Second-language Writing*, 21, 321-331.

Winter, G. (2000) A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research. *The Qualitative Report*, Vol 4, Nos 3 & 4, March, 2000. [Accessed on 21/06/08: http://www.nova.edu/ssss/QR/QR4-3/winter.html]

Wolcott, H. (1990) On Seeking– and Rejecting–Validity in Qualitative Research. In E. Eisner and A. Peshkin (eds) *Qualitative Inquiry in Education: The Continuing Debate*, 121–52. New York: Teachers College Press.

Wood, D. J., Bruner, J. S., & Ross, G. (1976) The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology*, 17(2), 89-100.

Woodal, B. (2002) Language switching: Using the first-language while writing in a second language. In *Journal of Second-language Writing*, 11, 7-28.

Woodrow, L. (2006) Anxiety and Speaking English as a Second Language. *RELC Journal.* Vol. 37(3), 308-328, 2006.

Woulds, S. & Simpson, J. (2010) Dead on the page no more! The case for authentic, locally appropriate ESOL materials. *Language Issues*, 21.1, 4-20.

Wright, A. (1976) *Visual materials for the Language Teacher.* London: Longman.

Yacci, M. (2000) Interactivity Demystified: a structural definition for distance education and intelligent computer-based instruction. *Education Technology*, 5-16.

Appendix A Transcription Notation Conventions

Simplified Jefferson symbols for speech	
(.)	noticeable pause
(.4) (2.4)	timed pauses
↑hello ↓hello	noticeable rise or fall of intonation
[hello [hello	brackets used to indicate overlapping talk
.hh	in-breath
hh,	out-breath
he(h)llo	(h) humour or laughter as word is spoken
hell-	sharp cut-off of word
hell:o	speaker has stretched the preceding syllable
(hello to you)	unclear so transcriber's best guess
( )	completely unclear talk
<u>hello</u> HELLO	underlined indicates loud voice and capitals LOUDER
°hello°	quietly spoken words
>hello<	spoken quickly
<hello>	spoken slowly
→	line of significant importance
((*sneeze*)) ((*cough*))	sounds not possible to write phonetically
cześc {hello}	curly brackets used to provide translation and comments
Symbols used for interaction	
~~St~~	strikethrough for when letters are deleted on screen
⌨	actions with keyboard
⌨ h a p	letters line up with when talk and typing are simultaneous
🖱	actions with mouse
💻	keyboard and screen, e.g. typing and looking at the monitor
❑	actions with screen, such as looking at the monitor
❑+✎	interaction with objects, e.g. writing and looking at the monitor
❑----------❑	hyphens used to indicate a continued action
🗔	actions with electronic translator
✋☜☞☝☟	gestures (at screen and between each other)
🗣	a learner turns to face the other
😐	learner's gaze is not focussed on any specific person or object
☺ ☹	learner smiles or shows displeasure
✎	writing on paper
🕮	actions with paper instructions
⌛	waiting for the computer to respond
🗁	opens a drop down menu
⇦⇨⇧⇩⇖⇗⇙⇘	direction of mouse across the screen
⇘----------⇘	hyphens used to indicate continuous movement across screen
⇳	movement up or down of the vertical scroll bar
⇔	movement left or right of the horizontal scroll bar

Appendix B Sakia & Gamda higher-level actions

timing	item	higher-level action
0:00:00.0	1.0	HLA begins: create a booklet about Greenpeace
0:01:23.3	1.1	HLA begins: create a booklet template
0:02:40.9	1.1	HLA ends: create a booklet template
0:02:56.6	1.2	HLA begins: set-up the booklet layout
0:04:37.0	1.2	HLA ends: set-up the booklet layout
0:04:51.6	1.3	HLA begins: insert text and images into booklet
0:08:12.0	1.3	HLA ends: insert text and images into booklet
0:08:21.1	1.4	HLA begins: match text and images
0:18:18.1	1.4	HLA ends: match text and images
0:18:24.3	1.5	HLA begins: balance the columns
0:24:47.7	1.5	HLA ends: balance the columns
0:24:52.2	2.0	HLA begins: create a protest image for page 1 of booklet
0:25:07.5	2.1	HLA begins: decide on what their protest image will be about
0:26:34.6	2.2	HLA begins: find pictures on the internet for protest image
0:29:23.9	2.1	HLA ends: decide on what their protest image will be about
0:39:03.9	2.2	HLA ends: find pictures on the internet for protest image
0:44:31.9	2.2.1	HLA begins: find pictures on the internet for protest image
1:11:43.5	2.2.1	HLA ends: find pictures on the internet for protest image
1:18:19.6	2.3	HLA begins: write suitable title for protest image
1:29:31.2	2.3	HLA ends: write suitable title for protest image
1:30:13.5	2.0	HLA ends: create a protest image for page 1 of booklet
1:32:10.1	1.6	HLA begins: identify a heading for first for page 1 of booklet
1:33:03.4	1.6	HLA not completed: identify a heading for page 1 of booklet
1:33:17.8	1.7	HLA begins: write some text for the fourth page of the booklet
1:48:55.7	1.7	HLA ends: write some text for the fourth page of the booklet
1:48:55.7	1.0	HLA ends: create a booklet about Greenpeace
2:02:36.2	3.0	HLA begins: design a website about recycling
2:09:12.0	3.1	HLA begins: create a folder structure for website
2:10:55.8	3.1	HLA ends: create a folder structure for website
2:33:11.4	3.2	HLA begins: create a master/template webpage
2:47:36.6	3.2	HLA ends: create a master/template webpage
2:47:42.0	3.3	HLA begins: create an index page for website
2:50:06.9	3.3.1	HLA begins: find a YouTube video and embed in index
3:04:24.1	3.3.1	HLA ends: find a YouTube video and embed in index
3:10:21.8	3.3	HLA ends: create an index page for website
3:10:39.1	3.4	HLA begins: create information page for website
3:14:14.8	3.4	HLA ends: create information page for website
3:16:06.3	3.5	HLA begins: create thoughts page for website
3:20:30.5	3.5	HLA ends: create thoughts page for website
3:21:17.4	3.6	HLA begins: create email page for website
3:28:07.1	3.6	HLA ends: create email page for website
3:28:07.1	3.0	HLA ends: design a website about recycling

Appendix C Darras and Shourok higher-level actions

timing	item	higher-level action
0:00:00.0	1.0	HLA begins: create a booklet about Greenpeace
00:00:35	1.1	HLA begins: create a booklet template
00:04:27	1.1	HLA ends: create a booklet template
00:04:43	1.2	HLA begins: set-up the booklet layout
00:04:59	1.1.1	HLA begins: create a booklet template
00:06:59	1.1.1	HLA ends: create a booklet template
00:07:06	1.2	HLA ends: set-up the booklet layout
00:08:16	1.3	HLA begins: insert text and images into booklet
00:10:29	1.3	HLA ends: insert text and images into booklet
00:10:30	1.4	HLA begins: match text and images
0:15:35.4	1.4	HLA ends: match text and images
0:17:16.3	1.5	HLA begins: balance the columns
0:24:01.8	1.5	HLA ends: balance the columns
0:25:09.9	2.0	HLA begins: create a protest image for page 1 of booklet
0:26:14.2	2.1	HLA begins and ends: no discussion on their protest image
0:26:24.5	2.2	HLA begins: find pictures on the internet for protest image
0:36:19.5	2.2	HLA ends: find pictures on the internet for protest image
0:41:32.0	2.3	HLA begins: write suitable title for protest image
0:46:27.2	2.3	HLA ends: write suitable title for protest image
0:47:15.4	2.2.1	HLA begins: find pictures on the internet for protest image
0:50:44.5	2.2.1	HLA ends: find pictures on the internet for protest image
0:56:35.4	2.0	HLA ends: create a protest image for page 1 of booklet
0:58:08.1	1.6	HLA begins: identify a heading for first page of the booklet
1:01:20.8	1.6	HLA ends: identify a heading for first page of the booklet
1:01:34.3	1.7	HLA begins: write some text for the fourth page of the booklet
1:12:36.1	1.7	HLA ends: write some text for the fourth page of the booklet
1:17:51.5	1.0	HLA ends: create a booklet about Greenpeace
1:27:29.2	3.0	HLA begins: design a website about recycling
1:30:24.3	3.1	HLA begins: create a folder structure for website
1:32:26.0	3.1	HLA ends: create a folder structure for website
1:33:00.7	3.2	HLA begins: create a master/template webpage
1:46:42.1	3.2	HLA ends: create a master/template webpage
1:46:53.1	3.3	HLA begins: create an index page for website
1:55:46.6	3.3	HLA ends: create an index page for website
1:55:50.2	3.4	HLA begins: create information page for website
1:59:02.8	3.4	HLA ends: create information page for website
1:59:27.0	3.5	HLA begins: create thoughts page for website
2:06:43.6	3.5	HLA ends: create thoughts page for website
2:07:29.7	3.6	HLA begins: re-create index page using YouTube video
2:15:58.6	3.6	HLA ends: re-create index page using YouTube video
2:18:46.7	3.0	HLA ends: create a website about Greenpeace

Appendix D pre-task interview with Sakia

1	SW: So the job you have now is what?
2	It's being a receptionist, answering the phone calls, dealing with customers, photocopying, faxing, something like that.
3	SW: But you said your profession is actually teaching?
4	Yes, I'm a teacher of Polish language. I was studying in Poland.
5	SW: Have you recently qualified? Have you done much teaching?
6	In Poland I was studying five years and on the fourth year I had some lessons with students in the primary school. In my final year I was teaching in a high school for two months.
7	SW: Your present job, why are you leaving that?
8	Why? Because I would like to work in my profession.
9	SW: Do you know the Polish Centre?
10	Yes, but I've never been in.
11	SW: They do a lot of teaching there.
12	But it's only on Saturday and it's free, voluntary work.
13	SW: Will you teach when you go back to Poland?
14	I'm not sure because teachers in Poland have really low money.
15	SW: What about this country?
16	I'm not sure; perhaps work as a teaching assistant at first. In a primary school I could help the children.
17	SW: So where are you living now?
18	In the city centre, near Park Lane college.
19	SW: Is it nice?
20	Yes. I share with friends. A very small garden.
21	SW: What is your typical day?
22	I go to work. Tuesdays and Wednesdays I go to school. Then I go back to work. I finish about six. I do the shopping and when I go back home I do some learning, watch TV, read books, talk with my friends, speak with my family in Poland.
23	SW: And how do you communicate with your family?
24	Through Skype or I just ring them.
25	SW: Do you use things like Messenger.
26	I have messenger but my mum uses Skype so I use Skype.
27	SW: Do you have any sense of a community?
28	I know we have a Polish Centre but I've never been there. My only Polish friends are Gamda and her husband? Then I have Italian and English friends.
29	SW: In the class, we've been designing images and websites as forms of communication, do you see any benefits to this?
30	I didn't know how to use these types of programs we use now.
31	SW: Some of the students for example are developing their own websites. Diane for example through her website is trying to connect ladies from the ivory coast. If you created a website, what would it be about?
32	I'm not sure. I didn't have time to think about this.
33	SW: Do you use English quite regularly?
34	I'm watching English TV. I'm reading English books. I speak English with all of my friends, except Gamda and her husband. So I think it's good for me.
35	SW: You don't think you have problems using English.
36	Sometimes I do because of my vocabulary.
37	SW: When do you think you sometimes might have problems?
38	When I read I still use a dictionary. It is the way I learn more words.
39	SW: What types of reading do you do?

40	Last week I want to the pictures and I saw a film so I went and bought the book. I like criminal books.
41	SW: That helps to explain some of the images you created in class. When did you first use computers?
42	When I was small, at home.
43	SW: What about at school?
44	At first I learned on my own and then with friends and my father. Then I had lessons in school.
45	SW: What is the main reason why you use computers now?
46	To communicate with my family, friends, check emails, Skype. At work I have to use the computer at work for typing.
47	SW: Do you ever have any problems using computers?
48	No.
49	SW: Thinking about your English and the future, what is important to you?
50	I was talking about this with Gamda. I was saying to her that I wanted to learn English to be able to speak well, to write essays, to have normal English.
51	SW: What about computers and the future? What is important for you?
52	You mean about the website?
53	SW: Well, any type of communication. You communicate now with your family. Are there other important reasons why you will use computers in the future?
54	It depends on what type of website I create.
55	SW: I agree. Thank you. Any questions for me?
56	No.

Appendix E pre-task interview with Gamda

1	SW: Where do you live and how do you get here?
2	I'm from Poland and I live in Leeds now.
3	SW: Whereabouts in Leeds?
4	About one year and three months.
5	SW: Where in Leeds?
6	I live near city-centre.
7	SW: What do you do on most days?
8	Working.
9	SW: Where do you work?
10	Actually, I'm working in a restaurant as a waitress.
11	SW: How do you find that?
12	I like it.
13	SW: The money?
14	The money is better because you have tips.
15	SW: What have you enjoyed doing in the class?
16	Everything.
17	SW: What stands out for you?
18	I've never done Fireworks and Dreamweaver before. I found these interesting.
19	SW: Have you had to design images before?
20	No. Never.
21	SW: How do you find trying to communicate through images rather than words?
22	Now, in the 21st century, this is very helpful. Most people are communicating through pictures through the internet.
23	SW: What about websites as a form of communication?
24	Most people are using websites to find something. They are asking questions and they can find the answers. It's good to know how websites are made.
25	SW: We have created lots of websites in class. Some of the students have created their own websites. What do you think your website might be?
26	The first thing is about me but I don't know yet.
27	SW: Many of the students have created websites for their community. Do you have a sense of Polish community, here in Leeds?
28	I don't meet many Polish people.
29	SW: Any reason?
30	I am working a lot. Sometimes 50 hours a week for me. In my free time I am learning at home.
31	SW: What English do you use on most days?
32	I'm watching television but usually with subtitles because this is helping me a lot. When I'm reading I can understand what the people are talking about.
33	SW: Do you ever have any problems using English?
34	Yes. I don't know many words, like smart words. I'm using words which I know and I try to learn new words but it is going slowly.
35	SW: When you have most difficulty with English?
36	Writing. Now it is better than when I first came to England. I was scared to write something because I wasn't sure if it was correct or not. Sometimes poems and short sentences.
37	SW: Do you have to write much at work?
38	No.
39	SW: Why do you write now then? We do a bit in class. What about at home?
40	I have a book with English exercises.
41	SW: When did you first use computers?
42	In my country. At school. I don't remember how old.
43	SW: Did you use computers a lot?

44	We started computer lessons in school but not in primary school.
45	SW: How are those lessons different to what you are doing now?
46	I have a computer at home which I use every day. It's different because before I was just reading internet sites. Now I know how they are created.
47	SW: Why do you use the internet at home?
48	I look for information that is interesting to me at this moment.
49	SW: You mean like search engines, google?
50	Yes.
51	SW: What about communicating with family and friends?
52	We use Skype mostly.
53	SW: Do you English or Polish?
54	Mostly Polish as that is where my parents are. My sister is in France. That is why we communicate with each other using Skype.
55	SW: Do you ever have any problems using technology?
56	Maybe when learning something new but when I practice I don't have many problems. I do see a difference though. My sister is seven years younger than me. She started using computers younger than I did and she is much better.
57	SW: Think about your English and computers, what do you hope will happen in the future?
58	I hope to be able to communicate without any problems. I try to do what I can to learn English. That will help me here and in my country.
59	SW: Do you spend much time with English, native speakers?
60	I would like but ...
61	SW: Your friend speaks Polish. You work in an Italian restaurant.
62	That is why I am going to change my job. Perhaps to an English restaurant.
63	SW: How do you think you will use technology in the future?
64	I don't know.
65	SW: Some of the students for example want to go to university and learn more about web design. What would be your ideal job?
66	I think to do English websites you need a high level. My occupation is an accountant. Now I have two ways. Be an accountant or a website designer.
67	SW: What are the obstacles for you working as an accountant in this country?
68	I haven't thought about this. I think I need to complete more courses.
69	SW: Any questions for me?
70	No.

Appendix F pre-task interview with Darras

1	SW: Where do you live?
2	I live in Lincoln Green.
3	SW: Do you like it?
4	Yes I really like it.
5	SW: But you've been asked to leave.
6	Yes. I have to home. The Support Team are helping me?
7	SW: Who is the Support Team?
8	They help asylum seekers and refugees to find a home.
9	SW: Do you have leave to remain?
10	I'm still waiting.
11	SW: What would you say is your nationality, your culture?
12	I'm Kurdish. I'm from Kurdistan.
13	SW: Do you belong to a Kurdish community?
14	No.
15	SW: Who are your friends then? Who do you socialise with?
16	I have a social worker.
17	SW: A social worker is paid to support you. Who are your friends?
18	I have some Kurdish friends. Most of my friends are from different countries, not Kurdistan.
19	SW: Is there a Kurdish community in Leeds? Or is there not enough Kurdish people in Leeds?
20	Not enough Kurdish people.
21	SW: What have you need doing in your lessons?
22	The website is not too bad now. Before it was hard but not that hard.
23	SW: What have you enjoyed doing?
24	I can't remember.
25	SW: Do you enjoy designing images to communicate?
26	Yes, it was really helpful and sometimes I do it at home.
27	SW: Can you give me an example of an image you've created?
28	In class or at home?
29	SW: Any.
30	A good citizen. I really like that image. It's important for refugee people to get citizenship.
31	SW: Thinking of English and literacy, do you use English with your friends?
32	Yes.
33	SW: What about at home?
34	On the computer using chat and sometimes with my Kurdish friends when they come to my home.
35	SW: Do you use Messenger?
36	PalTalk and chat rooms. Actually, I've started using Skype.
37	SW: Do you use audio?
38	Yes and I talk to people from Poland, England. And when I write to Kurdish people I have to write in English letters.
39	SW: So you use phonetic spellings, write in Kurdish but use English letters?
40	Yes.
41	SW: Does that ever create any problems?
42	Sometimes. My friend writes to me and sometimes I can't read.
43	SW: Do you ever have any problems using English?
44	Yes, sometimes.
45	SW: Can you give me an example?
46	Understanding new words. Sometimes I hear new words but I don't understand.
47	SW: Is that in class or out on the street?

48	Especially outside. Some people talk very quickly.
49	SW: When did you first use computers?
50	When I first came to England and joined this class. In 2005 I think.
51	SW: And before that.
52	I was a farmer-boy.
53	SW: Why do you use computers now?
54	I like to improve my typing, get information and get a good job.
55	SW: When do you use the internet?
56	When I want to find information. I search for things.
57	SW: But you always have to use English?
58	Yes.
59	SW: Do you ever have any problems using a computer, things you don't understand, things break down?
60	Yes I always have that problem. My computer. I don't understand it. I am lazy with that.
61	SW: Technical problems?
62	Yes.
63	SW: What do you when things go wrong?
64	Sometimes I ask my friends. I use restore, one week or two weeks ago.
65	SW: Thinking of your English, what do you want in the future?
66	Actually, I want to be a translator.
67	SW: Okay.
68	That is my dream. I must improve my English.
69	SW: How will computers be useful to you in the future?
70	It's very useful now because if you work in an office you need to know.
71	SW: So things like Word and Excel?
72	Yes.
73	SW: What about image design and web design?
74	Actually I have no idea.
75	SW: Some students in the class are creating websites about themselves or about their community. What do you think?
76	Actually I have never thought about this. Believe me, I have many problems. I have no chance to think about this.
77	SW: Housing, welfare?
78	Yes. And I still don't know my future. What I should do.
79	SW: Any questions for me.
80	No. Thank you very much.

Appendix G pre-task interview with Shourok

1	SW: Where are you living now?
2	In Leeds, Lincoln Green.
3	SW: What is it like?
4	It's small, a flat, a place to live.
5	SW: Is there any sense of community?
6	Most people are separate, on their own.
7	SW: What about neighbours?
8	The problem there is many multicultural people, for example, your neighbours are from other countries. It's too hard to talk to them.
9	SW: So you think there are many differences because there is a lot of multiculturalism?
10	Yes.
11	SW: In class we've talked about multiculturalism as a good thing.
12	It's so personal when it's your home. It's easy in class to talk with other people but not when it's your own home.
13	SW: What do you do on most days?
14	Most of the time using the internet and coming to college.
15	SW: What do you use the internet for?
16	I look for things.
17	SW: What type of things? What does the internet give you? If I took away your computer what would you lose?
18	I would lose many things. I'm in this country but I can see what is happening in my country. I can see movies and people in my country.
19	SW: So the internet for you is a window through which you can look into your country?
20	So easy, to watch the news and movies.
21	SW: Thinking of the lessons, what have you been learning on the course?
22	I really like this course because I like using the computer. Sometimes I don't go to bed until six in the morning, working on my website when I get problems.
23	SW: Images and websites provide people in the class with a different way of communicating. What do you think?
24	If you look at my website its name is iran4all. It gives me a different way to 'talk' to others. If you can't say it in one way you can say in a different way. It's like Rimi said, if you can't speak in good English you can say it in another way, for example, designing a protest image and saying it that way. If we don't like something and we can't say it we can communicate through images or file; something like that.
25	SW: I notice on your website that there is no writing. Why?
26	I need to get Persian language for keyboard. I could write in Persian/Latin.
27	SW: Have you been avoiding writing in English on your website?
28	No. I can write English and speak English. But for others who come to my website.
29	SW: What English have you used today? Do you use English all the time?
30	Not really. The problem is I have many friends who speak my language. The way you get to speak good English is to go to college; continue studying to get to university. It's too difficult to find an English friend or people from another country. You can't always trust others.
31	SW: Do you think then that your English language prevents you from saying what you want to say?
32	Yes, yes. For example, I have some problems with my home. Telephone bills and things like that. I know what I want to say but I can't say it in English sometimes. I get angry. It makes me nervous. And people they will not try to understand you.

33	SW: So using English with ‘native’ speakers, on the telephone, is where you have problems.
34	Using pronouns for someone behind the telephone. Using good pronouns and they will do your work.
35	SW: (laughs)
36	Believe me Stephen.
37	SW: I do. It is something I have no experience of. When did you first use a computer? And why do you use them now?
38	It was a long time ago in my country. I just used it for pictures and music and movies. Saving files on my hard drive. After that, things like Messenger and community and communicating with others.
39	SW: When you said Messenger you used the word ‘community.’ Earlier, when you talked about where you live, you said it was hard to have a community. Do you have a sense of community through Messenger and the internet?
40	Yes. I’ve got nearly 50 friends. Some of them are Iranian. And others in different countries that are a long way from here. Australia. Canada.
41	SW: Do you use the webcam as well?
42	You can see them with the webcam.
43	SW: Do you use a microphone?
44	Yes.
45	SW: What language do you use when typing?
46	Sometimes English. Well, I always type in English but it’s Farsi in English.
47	SW: So you use phonetic spellings?
48	Yes, using English letters for the sounds.
49	SW: Do you ever have any problems using technology?
50	Sometimes you can’t fix things but sometimes you can.
51	SW: What’s important to you, thinking about your English, computers and the future?
52	I really want to go to university and I need English.
53	SW: What would you like to study?
54	In my country I did 2 years of business management but then I had to leave the country. But I like computers as well.
55	SW: Computers are just for your own use?
56	I like to continue at university as well but I don’t know. I know something about my subject that I studied in the past.
57	SW: Any questions for me?
58	No, thank you very much.

Appendix H post-task interview with Polish learners

1	SW:	What was it like having to share the computer?
2	Gamda:	It was a good idea.
	Sakia:	For us.
3	SW:	Yes?
4	Gamda:	We get different ideas about many things.
	Sakia:	We couldn't finish yesterday.
5	SW:	How did you decide whose ideas to use if you wanted to do one thing
6	Gamda:	We choose something.
	Sakia:	We talk. If I do this by myself without Gamda maybe I do a little bit different.
7	SW:	And it would have been more quicker?
8	Gamda:	Yes.
	Sakia:	Probably.
9	SW:	So that's some of the bad things perhaps, about sharing? What were some of the good things? Were there any good things?
10	Gamda:	(they laugh)
	Sakia:	
11	SW:	Did you learn anything new from each other?
12	Gamda:	I think so.
	Sakia:	We usually support ourselves. Sometimes I look at Gamda's screen and sometimes she looks at mine.
13	SW:	What about the mouse then? Who had control of the mouse?
14	Sakia:	We were trying to share? (they laugh)
	Gamda:	Sometimes I'm thinking about what I'm doing so I just say 'take the mouse.'
15	SW:	Is that because you felt Sakia sometimes knew better than you how to do something?
16	Sakia:	Sometimes yes so she [says
	Gamda:	[okay do your way. (they laugh)
17	SW:	So you pushed your way in to take control of the mouse? Okay. What did you prefer doing the most? Was it Publisher, the leaflet, the website?
18	Gamda:	Dreamweaver. The website.
	Sakia:	When we were sharing the computer? Or?
19	SW:	Yes, when you were sharing?
20	Sakia:	I think the booklet. We did very quick.
	Gamda:	Mhmm.
21	SW:	Could you tell me about some of your image designs, colour choices?
22	Gamda:	Like to show mother earth.
	Sakia:	We were trying to pick together. To be more creative together.
23	SW:	So it was shared?
24	Gamda.	Yes. Shared.
25	SW:	Okay. That's it. Thank you.

Appendix I post-task interview with Kurdish learners

1	SW:	I want to ask for your thoughts, your honesty. How did it feel having to share the computer?
2	Darras:	Actually it's not easy.
3	SW:	Why?
4	Darras:	Because you have different opinion and your partner will not agree.
5	SW:	What do you think Shourok?
6	Shourok:	How do you mean?
7	SW:	Did you feel at times you wanted to do something that you couldn't?
8	Darras:	Yeah. I liked that picture and he choose another picture. (Darras and Shourok laugh)
9	Shourok:	It was better. That was my idea.
10	Darras:	Yeah. Maybe he was right but my opinion was different.
11	SW:	So how did you decide?
12	Darras:	(laughs) He was right.
	Shourok:	We just did.
13	SW:	So were there any good things about sharing?
14	Darras:	Yes, yes, actually, I didn't know something but when he did it I learn it.
15	SW:	So you actually learned something from …
16	Darras:	Yes I learned from him yes.
17	SW:	The mouse? Who controlled the mouse? (they both laugh) (to Darras) Did you feel you wanted to control the mouse but you couldn't?
18	Shourok:	Yes, yes, yes I could see.
	Darras:	I wanted but couldn't (laughs) I tried to but he didn't let me
19	SW:	You could sense he sometimes wanted to use the mouse?
20	Shourok:	Yes.
	Darras:	Sometimes you work with other people and I think it's enjoyable sometimes. Yes, sometimes.
21	SW:	And did you speak in English all the time?
22	Shourok:	Yeah. All the time. Sometimes I answer in Kurdish.
	Darras:	Yes. Kurdish. But I wanted to speak English all the time.
23	SW:	Did you think about colours, pictures to use?
24	Darras:	Mostly pictures.
	Shourok:	Some colours. Mostly green.
25	SW:	You were trying to create texts for a Kurdish audience?
26	Darras:	Yes.
	Shourok:	We decided to find pictures about Kurdistan, and Iran.
27	SW:	Which did you enjoy most, designing Publisher, the image, the website?
28	Darras:	For me I think image, or Publisher.
29	SW:	Why not the web design?
30	Darras:	What about you Shourok?
31	Shourok:	I think maybe all of them.
34	SW:	Anything you found difficult to do, other than the web design?
35	Darras:	Everything was easy because of him. When you work with someone else it is easier because you have something else and he has something else. When you forget something maybe your partner knows.
36	SW:	We have completed lots of designs, do you think we should share more?
37	Shourok:	Myself, I find it difficult.
38	SW:	You think sometimes it can be helpful?
39	Darras:	Yes. Sometimes you forget something and your partner can help.

Appendix J classroom task instructions

From www.greenpeace.org

1. Greenpeace exists because this fragile earth deserves a voice. It needs solutions. It needs change. It needs action.

2. To help us get our message across to people around the world we design leaflets, booklets, images, videos and websites.

3. The modern global communicator of today isn't just bilingual. He or she can communicate in multiple ways, using sound, words, images and colours.

4. He or she understands how to design communication to suit the audience and purpose.

DESIGN BRIEF:

- Create a booklet, a protest image and a website about Greenpeace
- **Purpose:** to persuade people to stop polluting the earth and look after it
- **Audience:** people from your country. How would you persuade them?

1. Design an image protesting against the destruction of the earth's natural resources.

   Open **Fireworks** to create a design for page 1 of your booklet. You will need to create a new canvas, size 400 pixels wide X 400 pixels high.

   Use the internet to find images. Design your protest image and save your image in the **Greenpeace** folder.

5. For **page 4** of your booklet create a header called 'My Thoughts' and add a text box. Write a few sentences about Greenpeace or environmentalism. What do you think?

6. Save and print your leaflet so it is back-to-back.

Appendix K consent form

Title: Language, literacy and technology: embodied peer-interaction and collaborative writing in an ESOL classroom

Contact: Stephen Woulds, Leeds City College

		Please Initial Box
1.	I confirm that I understand the purpose of the study and have had the chance to ask questions.	
2.	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason. If I do not wish to answer any questions I can refuse.	
3.	I agree to the interview being audio recorded.	
4.	I agree to the classroom session being video and audio recorded.	
5.	I agree to the use of anonymised quotes in publications.	
6.	I agree for the data collected from me to be used in future research.	
7.	I understand that some of the data collected during the study may be looked at by individuals from the University of Leeds. I give permission for these individuals to see this data.	
8.	I agree to take part in the above study.	

_____ _____ _____

Name of Participant Date Signature

_____ _____ _____

Name of Researcher Date Signature