Language Studies, Science and Engineering 2

Language-Learner Computer Interactions

Theory, methodology and CALL applications

EDITED BY
Catherine Caws
Marie-Josée Hamel

John Benjamins Publishing Company

Language-Learner Computer Interactions

Language Studies, Science and Engineering (LSSE)

ISSN 2210-7029

The *Language Studies, Science and Engineering* (LSSE) book series seeks submissions of book proposals that address the interface among language studies, science, engineering and education. This book series aims to bring together researchers from the fields of language studies and science/engineering education with the aim of generating new interdisciplinary knowledge. This book series is premised on the concept that science is of central importance in the 21st century and that research informed by linguistic knowledge can contribute to the description, understanding, education and practice of science and engineering. The goal of this series is to enhance educational and professional practices in the sciences and engineering through interdisciplinary interaction between language researchers, science and engineering educators and scientists.

The scope of this book series covers the range of potential contributions that language studies can make to the advancement of science, engineering and educational practices in these fields. Researchers who utilize language based methodologies, such as discourse analysis, computational linguistics, conversational analysis, multimodal analysis, rhetorical analysis, and genre analysis, for the purposes of advancing science/engineering education and professional practice are invited to submit a proposal for this innovative book series.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/lsse

Volume 2

Language-Learner Computer Interactions
Theory, methodology and CALL applications
Edited by Catherine Caws and Marie-Josée Hamel

Language-Learner Computer Interactions

Theory, methodology and CALL applications

Edited by

Catherine Caws
University of Victoria

Marie-Josée Hamel
University of Ottawa

John Benjamins Publishing Company
Amsterdam / Philadelphia

DOI 10.1075/lsse.2

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2016011008 (PRINT) / 2016023443 (E-BOOK)

ISBN 978 90 272 5751 2 (HB)
ISBN 978 90 272 6698 9 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

Part II. Data and elicitation technologies and techniques

List of figures

List of tables

Book series preface

Language Studies, Science and Engineering

I am very pleased to introduce Catherine Caws and Marie-Josée Hamel edited book which is the second publication in the *Language Studies, Science and Engineering* series. The book series was initiated to allow applied linguists and STEM professionals to interact around research methodologies and questions which are of mutual interest. Interdisciplinarity is at the heart of the current book as engineering, science and technology are integrated in an innovative discussion of ways in which language and literacy can be developed. The emphasis is on design and technology, the content is literacy and language and the approach directly drawn from current understanding in engineering. Together this is a powerful combination of disciplines and understandings and functions in the established applied linguistic tradition of utilizing all available resources, approaches and methodologies in solving real world problems and furthering educational issues.

Catherine Caws and Marie-Josée Hamel take this book series in new directions by exploring ways in which technology and the associated conceptual and research methodologies can contribute to issues of language and literacy learning. As such this book represents and impressively exemplifies the ways in which science, technology, engineering and applied linguistics can work in an interdisciplinary sphere and provides value for all involved. The two-way interaction between applied linguistics and STEM is once again shown to be rich ground for exploration and utilization. With a perfect balance of theory, research and practice, this book offers an innovative understanding of what technologically mediated environments can accomplish and the ways in which applied linguistic professionals can work with them.

Hopefully this edition will encourage other professionals to take a careful look at the interdisciplinary zone within which the first two books in this series exist and consider future directions for extending the power of this rich interaction between applied linguistics and STEM.

David Ian Hanauer Editor: *Language Studies, Science and Engineering*

Professor of Applied Linguistics: Graduate Studies in Composition and TESOL, English Department, Indiana University of Pennsylvania
Lead Assessment Coordinator: SEA-PHAGES Program, Hatfull Laboratory, Pittsburgh Bacteriophage Institute, University of Pittsburgh

Acknowledgments

We had a vision for this book for some time and we decided to be bold. After all, with such a talented team, it had to go well – and it did – thanks to the amazing group of scholars who helped us to follow our dream. Naturally our first big thank you goes to each of the authors in this book, each of whom worked so hard in meeting our deadlines and having produced remarkable chapters. Most importantly we are grateful that, in spite of the bumps in the road, they remain good colleagues and friends!

A special note of appreciation goes to the editorial team at the John Benjamins series "Language Studies, Science and Engineering (LSSE)", especially its editor, Professor David Hanauer for his feedback throughout the process, his encouragement and sound advice, Kees Vaes for guiding us during the last phase, and Justin Nicholes for the careful language editing of our manuscript in such a tight timeline.

As a final thank you, we are so grateful for the support of our respective partners, Greg and Graham, and our colleagues at the University of Victoria and the University of Ottawa, Canada.

Contributor biographies

Editors / Authors

Catherine Caws is an Associate Professor of applied linguistics in the Department of French at the University of Victoria (Canada). Her research focuses on Computer Assisted Language Learning, and, more particularly, on learning design, learner-computer interactions and data-driven learning. She has authored many articles related to students' success in higher education and runs a blog on "Teaching French in the 21st century."

Marie-Josée Hamel is a Professor of applied linguistics at the Institute of Official Languages and Bilingualism at the University of Ottawa (Canada), where she holds a Chair in *New Technologies and Computer Assisted Language Learning*. She has been involved in CALL teaching, research and development for the past twenty years. Her interests are in the contribution of linguistic and second language acquisition theories to CALL and in the design of ergonomic CALL environments and resources.

Contributors / Authors

Françoise Blin is a Senior Lecturer in the School of Applied Language and Intercultural Studies at Dublin City University (Ireland). She has been teaching French with the help of technology at Dublin City University for the last thirty years. She is co-editor of *ReCALL* and the current president of the European Association for Computer Assisted Language Learning (EUROCALL). Her more recent works focus on the applications of ecological and activity theoretical approaches to CALL research, design and practice.

Thierry Chanier is a Professor of applied linguistics at Université Blaise Pascal, Clermont-Ferrand, France. His main research interest over the past twenty years has been computer assisted language learning and, since 1999, online learning in telecollaborative situations. He coordinated the Mulce project, which developed an open-access repository of LEarning and TEaching Corpora.

Cathy Cohen is an Associate Professor at the teacher training college at the University of Lyon 1 (France). She teaches courses on language pedagogy and bilingualism, as

well as teaching English for specific purposes. She is a member of the ICAR research laboratory (*Interactions, Corpus, Apprentissages, Représentations*). Her research interests include language pedagogy, teacher education in computer-mediated communication and bilingual acquisition in children and young learners.

Nicolas Guichon is a Professor of linguistics at the University of Lyon 2 (France) and is a member of the ICAR (*Interactions, Corpus, Apprentissages, Représentations*) research team. His research interests include teacher education in computer assisted language learning (CALL), the study of online interaction and materials design.

Trude Heift is a Professor of linguistics at Simon Fraser University, Canada. Her research focuses on the design as well as the evaluation of CALL systems, with a particular interest in learner-computer interactions and learner language. Her work has appeared in leading CALL/SLA journals, and she is co-author, with Mathias Schulze, of *Errors and Intelligence in Computer Assisted Language Learning: Parsers and Pedagogues* (Routledge). She is co-editor of *Language Learning and Technology*.

Mike Levy is an Honorary Professor in the School of Languages and Cultures at the University of Queensland, Brisbane, Australia. His research includes studies on the distinctive role of technology in mediating language learning, including how the technology itself shapes the interaction at both the macro and the micro level. His interests span theory, design and practice, and his work has included studies on digital media, mobile language learning, online cultures, teacher education and learner training. Two recent papers consider research and development of online dictionaries and electronic translation tools. He is on the editorial boards of *ReCALL, CALICO* and *System*. Currently, he is Chair of the Steering Committee for the WorldCALL Conferences held in different parts of the world every four years.

Kyle Scholz is a PhD Candidate in the Germanic and Slavic Studies department and a liaison with the Centre for Teaching Excellence at the University of Waterloo. His research interests include complex adaptive systems and digital game-based language learning, where he explores their applicability in extramural learning contexts. His current research examines the language learning and gameplay trajectories of learners playing the game *World of Warcraft* to support the transfer of language observed and produced in-game to non-gaming contexts.

Mathias Schulze is a Professor of German (Applied Linguistics), the co-editor of the *CALICO Journal*, and the director of the Waterloo Centre for German Studies. His research interests include the application of Complexity Theory in CALL research, the use of natural language processing in CALL, online language learning, and the development of proficiency as a part of second language development.

Jérémie Seror is an Associate Professor at the Official Languages and Bilingualism Institute at the University of Ottawa. His research focuses on advanced literacy development, content-based language learning and the language socialization of multilingual students in educational settings. He has also drawn on his expertise with computer assisted language learning to research the strategies and composition processes of language learners in digital spaces and the application of screen capture technologies for literacy development.

Lijing Shi is currently an assistant coordinator of Mandarin at the London School of Economics and Political Sciences, where she teaches Mandarin Chinese. Her main research areas include online language teaching/learning, and intercultural communicative competence.

Bryan Smith is an Associate Professor of Applied Linguistics in the Department of English at Arizona State University in Tempe, Arizona. His research focuses on the nexus of CALL and SLA. He is the co-editor of the *CALICO Journal* (along with Mat Schulze) and has published widely in the area of computer-mediated communication in language learning.

Ursula Stickler is a Senior Lecturer in the Department of Languages at the Open University in the UK and convenor of the Open Languages Research Group. Her research interests focus on independent and technology-supported language learning and qualitative methods in CALL research. She is involved in European projects, collaborating in the area of teacher training, and she has written a number of articles and book chapters on language learning and teaching. She is a supervisor on the university's PhD and EdD programs.

Steve Thorne holds faculty appointments in the Department of World Languages and Literatures at Portland State University (USA) and the Department of Applied Linguistics at the University of Groningen (The Netherlands). His interests include formative interventions in language education contexts, intercultural communication, indigenous language revitalization, communication across new media and mobile technologies, and research that draws upon contextual traditions of language analysis and usage-based and distributed approaches to language development.

Ciara R. Wigham is a Senior Lecturer in English and applied linguistics at the Language Centre, Université Lumière Lyon 2 (France), where she currently directs a master's program in language education and CALL. Her research interests are multimodal pedagogical communication in online language learning and methodologies for the description of online learning situations. She is a member of the ICAR research laboratory (*Interactions, Corpus, Apprentissages, Représentations*).

CHAPTER 1

Cutting-edge theories and techniques for LCI in the context of CALL

Catherine Caws and Marie-Josée Hamel
University of Victoria, Canada / University of Ottawa, Canada

As an introduction to the field of learner-computer interaction, this chapter argues for a need to generate knowledge about the online language learning process, developing a capacity for doing so by using cutting-edge frameworks and methods grounded in science and engineering. Adopting a posture of CALL engineers, we approach interaction-based research in CALL through the core concept of *design* and discuss LCI investigations in the context of technology-mediated task-based language learning. This chapter also presents the aim of the book; highlights the main features of contributors' chapters; identifies the book's readers and purposes for which it can be used. It summarizes each chapter in order to highlight the variations in theories and methods that this book promotes for the analysis of LCI. As such, this introductory chapter serves to guide readers to better apprehend the book content.

Keywords: learner-computer interaction (LCI), human-computer interaction (HCI), CALL, technology-mediated language learning, design

Introduction

When considering CALL research and practices from a scientific and engineering angle, we recognize that the role of computers and, more generally, of technology in society and especially in education remains far from simplistic, obvious, or unique. Popular media misrepresentations have divided the public between lovers and haters of technology, distinguished by an excessive trust in the power of computers (such as this 2014 article featured in the *New Yorker* <http://www.newyorker.com/> "*Will computers ever replace teachers?*") or by an exaggerated fear of new technologies.

Within the specific context of language learning and teaching, the value, opportunities and challenges brought about by technologies can be examined from

DOI 10.1075/lsse.2.01caw

many angles: the pedagogy, the curriculum, the relation between learner(s) and instructor(s), the evaluation, the learning objectives and tasks, or simply, the tools. Regardless of the approach favoured, *design* remains critical for the success (or failure) of any intervention. And if good design can lead to better learning, we ought to ask ourselves this simple question: How can we design effective, sustainable learning ecosystems mediated by technology?

Our premise in this book is that interaction-based research in CALL can assist us researchers and practitioners in reaching our goal. By providing specific theories and methods centred on the relationship between a human (herewith a learner) and an artefact (herewith a technology), interaction-based research can inform us on specific models and interventions that are common in technology-mediated learning and teaching and that may need further development. More specifically, interaction-based research can guide us in improving the *design* of such learning environments by showing us exactly what learners typically do when interacting with technologies. With such interaction models providing empirical data that are obtained by way of observing and computer tracking, researchers can apply scientific methods to analyse, assess and recycle their findings into further interaction-based interventions in a view to create optimal CALL learning ecosystems. Thus, an iterative process is born for CALL research, partially modelled upon theories and practices from the fields of engineering and sciences.

As an introduction to the rich field of interaction-based CALL research, this chapter presents an overview of the ways in which interconnections between sciences and humanities have led us to rethink, value and reflect upon learner-computer interactions (LCI) and how this re-thinking about LCI from a scientific perspective has allowed us to (re)value the concept of *design*. While introducing key concepts emergent in the field of CALL research centred on LCI in this chapter and subsequently throughout the book, we make an argument for sharpening our understanding of technology-mediated language learning processes using cutting-edge frameworks and methods, several grounded in science and engineering. In order to do so, we adopt the posture of CALL engineers while considering LCI investigations in the context of technology-mediated task-based language learning tasks.

Looking at CALL research and practices through the lenses of scientific theoretical frameworks

As an analogy to HCI (human-computer interaction), learner-computer interaction (coined LCI and woven throughout the book) is the focus of our volume.

Intended to offer a fresh outlook and innovative perspectives, the book looks at CALL research and practices through several lenses of theoretical frameworks inherited from the sciences.

Throughout the chapters, LCI processes are emphasized. While some of these processes are clearly embedded in a second language acquisition (SLA) framework, such as identifying language tasks and their completion patterns or analysing behavioural and metacognitive strategies, other processes may be inherited from engineering practices, such as testing for usability (measuring efficiency, effectiveness and user satisfaction of a system), troubleshooting (a form of re-engineering that is particularly helpful in finding causes of a failed system) or reverse engineering (a process of dissembling or reversing potential malfunction of a design, system or technology). In revisiting and recycling frameworks, approaches, tools and techniques that commonly apply to engineering, HCI, or software design, our primary goal is to sharpen our assessment of design and learning processes, in particular those that relate to language and literacy development in technology-mediated environments. Moreover, our motivation in linking scientific methods and CALL research methods results from the fact that they provide a methodology that can support data elicitation and analysis within a rich theoretical framework. The content of the book is hence unique, rich and varied, going from ergonomics to complex systems, from affordances to personas, from screen-capture to eye-tracking techniques, from specific learning design to recycling empirical data and creating multimodal corpora, in a view to ameliorate language learning ecosystems.

How did we come to consider LCI within the perspectives of scientific and engineering frameworks?

Design is the anchor that binds engineering and LCI, also the link that unites our team of researchers. Indeed, while many other connections with other disciplines can be made, when we reflect upon the true meaning of engineering, clear overlap appears between engineering and applied linguistics research and methods.

On being CALL engineers

The relationship between humans and artefacts (human-made objects) manifests itself clearly through engineering. Adopting an activity theory perspective to engineering enables us to understand the special bond between humans and artefacts. Indeed, one of the goals of activity theory is to analyse the way in which

these artefacts influence interactions, how these interactions evolve and change based on the sociocultural context where mediations occur (Engeström, 1987; Leontiev, 1981). The interaction of humans in and with their environment is mediated by artefacts that humans have *engineered* themselves, exploiting resources or objects that they have at their disposal. At the same time, humans are constantly adapting, shaping or redesigning these artefacts to better suit their needs, purposes and goals. Their observational, analytical and planning skills are core in foreseeing the affordances that resources in their environment have to offer in terms of artefact-building opportunities.

Affordances, understood as intrinsic capacities of objects that reveal themselves through usage, emerge in activities. Human activities and minds are mediated by culturally developed tools (Kaptelinin & Nardi, 2012, p. 972). Artefacts can also be congregated to form complex systems. They evolve through human interventions and are dynamic in essence. Those humans in our society who have acquired knowledge and skills, enabling them to devise such complex systems of artefacts, are referred to as engineers.

Engineers apply scientific theories (in particular, those coming from mathematics and physics) to guide the (re)design and evaluation of complex artefact systems (whether civil, electric, mechanical, environmental or technological). They build models and prototypes based on investigations of needs and analyses of requirements, taking into consideration contextual variables (e.g., physical and social environment). They test these elements using simulations to predict best solutions for design processes and/or outcomes. Engineers work collaboratively, in interdisciplinary teams of thinkers and doers.

CALL researchers and developers do the same. They are engineers in the sense that they approach the design, evaluation and description of complex learning-artefact systems from top-down (theory-driven) and bottom-up (data-driven) perspectives. This dual approach enables them to create abstract models of learning, to build and test concrete prototypes for learning, to simulate learning processes and to anticipate their outcomes. Their motivation for engaging in engineering activities stems from problems that they have identified through empirical investigations, focused on the learners and their learning environments. In that context, analysing learner behaviours and the outcome of such behaviours is critical as a means to inform, and to enrich complex and dynamic learning systems.

Their capacity to resort to their tacit knowledge and experience is enhanced by the fact that CALL researchers, who are CALL developers, are very often also CALL practitioners. This triple hat of thinker, doer and user of CALL systems gives them a privileged insight into the discipline, which engineers might not have the opportunity to acquire.

Digging into theories, borrowing methodologies

CALL, as an applied linguistics discipline, has a relatively long and strong tradition at investigating theoretical research and frameworks, in particular, interactionist second language acquisition (SLA), as well as socio-constructivist and sociocultural perspectives (e.g., Chapelle, 2005; Lantolf & Thorne, 2006). In so doing, CALL has had several goals: for instance, to shed light on LCI and to focus on aspects of language development that have been observed in technology-mediated contexts (e.g., focusing on form, negotiating meaning, producing comprehensible output or identifying ideal conditions for SLA to occur in such contexts (Chapelle, 2005). In contrast to these frameworks, the theoretical perspectives and frameworks (*ergonomics, theory of affordances* and *complex systems*) that are discussed in the first part of this volume have been less explored in the context of CALL. We believe they are innovative and particularly meaningful in the specific context of CALL research and development (R&D) because they unite CALL and engineering, while helping us deepen our understanding of LCI.

Unlike engineering, however, CALL is a younger discipline, anchored traditionally in the humanities. As such, CALL does not have its own dedicated research and development methods (such as usability tests in web design) and tools (such as AutoCAD for engineering design). Consequently, methods for investigating LCI in the context of CALL are not specific to the discipline but rather come from various research traditions, including the following: classroom observation (Good & Brophy, 2000), corpus linguistics (McEnery & Wilson, 2001), conversational analysis (Sidnell, 2010) and discourse analysis (Renkema, 2004). The same can be said about tools and methods that are used to elicit and analyse LCI data in the context of CALL. These vary from traditional (yet powerful) instruments, like questionnaires (Dörnyei, 2010) and interviews (Maurel, 2009), which provide indirect, yet important perspectives on LCI, to methods based, for instance, on the verbalization of actions, decisions and thoughts, such as talk-aloud, stimulated recalls and walk-through (e.g., Gass & Mackey, 2000; Hémard, 2003; Hughes & Parkes, 2003). These insights on learner behaviours allow us to make inferences about strategies that language learners deploy when interacting at the computer.

In the second part of the book, we introduce computer-tracking tools (such as *eye-tracking* and *video screen captures*) and techniques (such as building *personas* and *learner corpora*) that are relatively new and are mainly inherited from cognitive science or software engineering (web design industry). Using these tools and techniques allows us to collect, organize and analyse LCI data in cutting-edge ways. As a result, we obtain new and comprehensive perspectives on LCI, focused on complex and dynamic LCI processes.

LCI investigations in the context of technology-mediated language learning tasks

Interactions constitute a core component of language learning. The basic tenet of interactionist SLA is that language learning has greater chances of occurring through interactions. Chapelle (2005) has explained, "the term interaction [is] the superordinate concept that includes any type of two-way exchanges" (p. 54). She reminds us that learners use linguistic and non-linguistic means and cues in the *exchange process* during language learning activities, where they need to construct meaning in order to reach their individual and/or collective goals (Chapelle, 2005). These goals can be individual or collective and may be expressed in terms of linguistic, cultural, social or communicative competences. The design of goal-oriented activities, formulated as language learning tasks, will also shape the nature of the LCI. Tasks can indeed be considered as more convergent, meaningful and purposeful forms of language learning activities (Ellis, 2003). They consist of a powerful, pedagogical way for language teachers to operationalize concepts conveyed by theoretical approaches and to structure teaching (Guichon, 2012). Language learning tasks should provide learners with rich interaction opportunities so that they direct their attention to the linguistic forms, negotiate meaning and enhance their language output through feedback.

Language learning tasks vary in sizes and scopes. For instance, at the macro level, a language learning task (such as writing a blog) may encompass all interactional aspects while, at the meso level, a task (such as commenting on a blog post) may be centred on negotiation of meaning, and finally, at the micro-level,

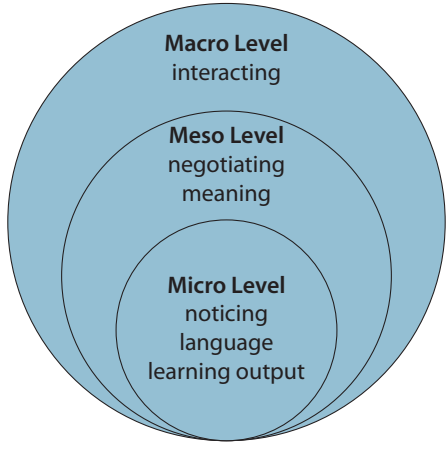


Figure 1.1 Three levels of a language learning task

the same task (such as editing a blog post or a comment) will focus on noticing language learning input.

Technology offers language learners opportunities that Nissen (2011) has referred to as *tooled* opportunities. This concept describes technology-mediated occasions, whereby learners can practice language in authentic situations, individually or collectively, to engage in meaningful projects and initiatives, increase their sociocultural awareness or develop their language autonomy. Language opportunities abound in virtual environments, and learners can easily exercise their autonomy when, for instance, they exchange opinions in an online forum, collaboratively write a wiki, self- or peer-edit a scientific article, produce a YouTube video about themselves, or solve a quest in a gaming situation. These LCI are carefully planned in order to exploit the affordances of technology that, in the context of CALL, are referred to by Mangenot (2013) as the "semio-pragmatic characteristics of technology in relation to communicative practices and pedagogical interventions" (p. 16, our translation).

So what does investigating LCI in the context of technology-mediated language learning tasks really mean? It means observing the exchange process that occurs via, with, and through technology, when learners are attempting to reach personal and common goals. It also means examining the outcome(s) of such an exchange process, analysing whether personal and shared goals have been successfully achieved and looking at the context in which it has occurred. Adopting an ergonomic perspective (i.e., a learner-centred perspective) on the analysis of LCI enables a focus on learner behaviours in technology-mediated tasks. Within such interactions, the observed behaviours (often combined with learners' perspectives on their interactions) can provide hints on the quality of the relationship that exists between the learner and the task, the learner and the tool or the task and the tool. The results of LCI analyses may indicate that the design of a task could be improved, that the usability of a tool could be increased, that learners' strategies could be improved, all to better address learners' needs. In addition, these interactions may, in parallel, reveal how language is involved in the construction of meaning.

Combining other types of empirical data about (and around) this exchange process and its outcome(s), and taking into account individual and contextual variables (e.g., the learners' prior experience and preferences, the task set-up, etc.), will enable a richer understanding of LCI. The basic argument is that we need to look at the learner in a CALL environment that is viewed as a multi-dimensional space. This *complex system* features multiple variables that may have an effect on the learner's language production and the learner's development. Therefore, we need to resort to a multivariate technique to better describe this space and any

phenomena that exist within it. This description has to include quantitative, qualitative and longitudinal elements at once, which is in essence what a complexity theory seeks to achieve.

Empirical data gathered dynamically about learner behaviours in CALL environments can also be transformed into multimodal learner corpora that may be openly accessed for research, training and teaching purposes. Resulting outcomes can be recycled to improve the quality of LCI, to advance SLA theory, or to put into test emergent theories in CALL, such as complexity theory.

About the book

The central aim of the book is two-fold. First, it seeks to explain how these cutting-edge theories and data-elicitation and data-analysis methods enable an in-depth, informed and objective, dynamic and multimodal investigation of language learners' interactions in technology-mediated environments. Secondly, the book describes the purpose of such theories and methods and the contexts (illustrated by case studies) in which they can be applied. Particular attention is given to CALL *design* as we make the case for (multimodal) online language learning tasks and environments that facilitate the language learning process. It also provides recommendations on how language teachers can better scaffold learners online during their language learning process.

In order to reach its objectives, the book proposes an innovative approach to describing CALL research by outlining and highlighting specific connections between research disciplines (such as human-computer interaction, web design and ergonomics, or engineering) originally grounded in the sciences, and computer assisted language learning (CALL) research, a discipline that is traditionally housed in applied linguistics (second language acquisition and second language pedagogy).

Lastly, the book offers fresh perspectives by gathering theoretical reflections and exemplar studies from researchers in applied linguistics who come with rich and varied experience not only in second language acquisition but also in language engineering.

All book contributors bring their background in sciences and language engineering to enrich their research and apply their findings in unusual ways. This enables them to create abstract models of learning, to build and test concrete prototypes for learning, to simulate learning processes and to anticipate their outcomes.

Readership

This book addresses a wide readership: graduate students at the master and PhD levels, scholars involved and/or starting to be involved in CALL research, computer-scientists with a background in the humanities who are looking for new ways to bridge the gap between their discipline and disciplines housed in other faculties at their institution, and any reader, scholar, designer who, like Steve Jobs, believes in the interaction between art and science, i.e., interdisciplinary research and development.

Readers of this book should be able to gain an in-depth understanding of what being a CALL research and development (CALL R&D) engineer entails, by exploring theories and methods, as well as numerous illustrations and examples drawn from LCI research studies that have been conducted in the specific context of CALL research and development.

Book structure

The book is divided in two main parts, allowing the reader to better grasp the connections between the theories and the methods (used for both research and language learning). To enhance this connection, a chapter is used as a pivot between both parts. This division addresses the need to frame CALL research in sound theoretical practices.

Part I of the book (*Frameworks guiding the research*) presents theoretical perspectives that are core in other applied sciences, while only emerging in CALL. It includes three chapters focusing specifically on theoretical concepts (*ergonomics*, Chapter 2), and theories (*affordances*, Chapter 3, and *complex systems*, Chapter 4) that are explained and illustrated in order to present arguments for adopting and adapting them in the context of CALL research and development focusing on LCI analyses. Part I also features a chapter on *design and research* (Chapter 5) which aims at connecting theoretical notions with practical methods.

Part II of the book (*Data and elicitation technologies and techniques*) offers the reader a wide spectrum of possibilities in terms of conducting quantitative and qualitative empirical research on LCI, capturing its complexity, its dynamic process and its purpose(s). It contains five chapters: *learner personas* (Chapter 6), *video screen capture* (Chapter 7), *eye-tracking* (Chapter 8), *desktop videoconferencing* (Chapter 9) and *multimodal corpora* (Chapter 10). They describe technologies and techniques carefully chosen to emphasize the diversity of data-collection and data-analysis methods, and reveal ways in which they could easily be adapted to many other environments in CALL research and language learning research. The

focus on *interaction* as the underpinning characteristic of the volume enables a learner-centred, process-oriented description of what happens in the digital space when learners are engaged with technology.

Chapter summaries

In **Chapter 2**, Caws and Hamel revisit the concept of *ergonomics* in the context of CALL. Viewed as a methodological and theoretical framework that aims to describe interactions between learners and instruments, CALL ergonomics seeks to ameliorate these interactions so that learning can be maximized. Ergonomics is focused on what a learner *does* when interacting with instruments to improve CALL design and enhance interactions. These aspects are discussed in relation to HCI research, where the user plays a central role in influencing the interactions, providing rich data that can be recycled in many ways. The chapter also reflects on CALL ergonomic methods in the context of system evaluation and the analysis of learners' behaviours through direct observations.

Chapter 3 focuses on the theory of affordances, a theory that has been at the forefront of debates within the HCI community since the late 1980s and is also frequently called upon by CALL researchers seeking to adopt an ecological approach to CALL design. In this chapter, Blin explains the concept of affordances as it relates to CALL environments and, more particularly, to those environments that make extensive use of Web 2.0 applications. In doing so, she explores the relationship between technological, educational, and linguistic affordances, drawing on case studies as well as literature.

Chapter 4 introduces the readers to complex adaptive systems in CALL research. Schulze and Scholz argue for and sketch a research paradigm – with its ontological, epistemological, and methodological components – based on the understanding of second language development as a complex adaptive system. This chapter explains that such a complexity-scientific approach to research addresses questions that are central to the use of computers within technology-rich language learning contexts, and for the computational modelling of learning processes to achieve improved individualized instruction in CALL, hence reaching optimal LCI.

In **Chapter 5**, linking theoretical discussions to description of research methods and outcomes, Levy and Caws reflect upon the concept of normalization by exploring two specific areas of CALL work that have proved problematic over time. The first area relates to our understandings of the broader contextual factors that influence CALL activity, and the second relates to our understandings of the nature of interactions when those interactions are mediated via technology

in some way. These two specific areas of exploration offer macro and micro perspectives, and they consider CALL research within a context where technology is ubiquitous, forever changing and evolving, often in disruptive ways.

Chapter 6 forms the first element of Part II of the book. It focuses on case studies detailing individual learner characteristics (profiles) and moment-by-moment interactions. In this chapter, Heift addresses two questions, seeking to devise ways of individualizing instruction suited to a variety of users while, at the same time, addressing the needs of individual users. The case study presented investigates data on learners' help access and clusters learners and their behaviour into different learner personas. It indicates that identifying personas can assist us in better modelling learning processes and individualizing instruction.

Chapter 7 explores the use of video screen capture (VSC) technology as a method to document and analyse online writing task processes in three specific ways: as a tracking tool to collect rich empirical data of interactions produced in real-time, as a retrospection tool to allow users to reflect on their processes and as a scaffolding tool to generate more dynamic and multimodal feedback. To explore these methods, Hamel and Séror report on three specific case studies that are focused on affordances and relevance of VSC for second language (L2) writing pedagogy and the promotion of L2 writer autonomy. The chapter concludes with recommendations for optimal use of VSC as a way to enhance L2 writing tasks design.

Chapter 8, forming a natural continuation to VSC, is focused on using eye-tracking technology to explore the LCI process. Smith, Stickler and Shi examine how CALL researchers are employing eye-tracking technology in explorations of learner interaction in authentic, task-based computer-mediated environments. As they draw upon both cognitive and sociocultural theoretical underpinnings to instructed SLA, current findings from studies employing eye-tracking in CALL are explored, as well as potential areas for growth. The chapter concludes with a discussion on affordances and limitations of eye-tracking technology and recommendations on ways to integrate such technology to other, more established data-collection measures.

In **Chapter 9**, Cohen and Guichon present the methodological issues and challenges related to the analysis of gestural expressions in multimodal, synchronous online exchanges. Making the case for a deeper understanding of semiotic resources to comprehend how they may be better orchestrated in LCI contexts, the chapter analyses the various contributions that have been made to gestural expressions in pedagogical exchanges. The authors address such aspects as ethical issues and technical implications. They also consider determining relevant units of analysis before illustrating these themes by presenting a qualitative study based on synchronous videoconference interactions.

Chapter 10 constitutes the last section of Part II of the book. Taking a more holistic approach, the chapter discusses a staged methodology to build learning and teaching corpora (LeTeC) in a view to better capture the many elements that are at stake in situated learning and LCI. Chanier and Wigham describe the methods used to build the corpora. Most importantly, they argue for a concerted, collaborative research cycle involving a group of researchers in order to facilitate analysis across different online environments, in order to integrate data into larger corpora and in order to contribute further to general linguistics, applied linguistics or Natural Language Processing (NLP).

References

Chapelle, C. A. (2005). Interactionist SLA theory in CALL research. In J. Egbert & G. Petrie (Eds.), *Research perspectives on CALL* (pp. 53–64). Mahwah, NJ: Laurence Erlbaum Associates.

Dörnyei, Z. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York, NY: Routledge.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, United Kingdom: Oxford University Press.

Engeström, Y. (1987). *Learning by expanding: An activity theoretical approach to developmental research*. Helsinki, Finland: Orienta-Konsultit.

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Good, T. L., & Brophy, J. E. (2000). *Looking in classrooms* (8th ed.). New York, NY: Longman.

Guichon, N. (2012). *Vers l'intégration des TIC dans l'enseignement de langues*. Paris, France: Didier.

Hémard, D. (2003). Language learning online: Designing towards user acceptability. In U. Felix (Ed.), *Language learning online: Towards best practice* (pp. 21–42). Lisse, Netherlands: Swets & Zeitlinger.

Hughes, J., & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour & Information Technology*, 22, 127–141. doi:10.1080/0144929031000081341

Kaptelinin, V., & Nardi, B. (2012). *Activity theory in HCI: Fundamentals and reflections*. San Rafael, CA: Morgan & Claypool.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, United Kingdom: Oxford University Press.

Leontiev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in soviet psychology* (pp. 37–71). Armonk, NY: M. E. Sharpe.

Mangenot, F. (2013). Les échanges en ligne comme secteur de pratiques et de recherches en ALAO: Quelles problématiques, quelles évolutions? *OLBI Working Papers*, 5(5), 3–21. doi:10.18192/olbiwp.v5i0.1114

Maurel, M. (2009). The explicitation interview: Examples and applications. *Journal of Consciousness Studies*, 16(10–12), 58–89.

McEnery, T., & Wilson, A. (2001). *Corpus Linguistics*. Edinburgh, Scotland: Edinburgh University Press.

Nissen, E. (2011). Variations autour de la tâche dans l'enseignement/Apprentissage des langues aujourd'hui. *Alsic*, 14. http://dx.doi.org/10.4000/alsic.2344

Renkema, J. (2004). *Introduction to discourse studies*. Amsterdam, Netherlands: John Benjamins. doi: 10.1075/z.124

Sidnell, J. (2010). *Conversation Analysis: An Introduction*. Chichester, United Kingdom: Wiley-Blackwell.

PART I

Frameworks guiding the research

CHAPTER 2

CALL ergonomics revisited

Catherine Caws and Marie-Josée Hamel
University of Victoria, Canada / University of Ottawa, Canada

This chapter revisits the field of educational ergonomics in the light of the current state of learner-computer interactions (LCI) and within the specific context of language learning. The discussion starts by defining the elements that constitute ergonomics in computer assisted language learning (CALL) as a methodological and theoretical framework, reviewing key concepts and principal theories upon which CALL ergonomics is based. The discussion focuses on the motives behind this innovative approach before exploring specific examples of engineering methods that can be applied to CALL research. We argue that methods inherited from human-computer interaction (HCI) or human-centred design (HCD) offer an excellent complement to CALL research and that, vice-versa, CALL ergonomics constitutes a framework that is closely related to HCI research, in that the user plays a central role in influencing the interactions, providing rich data that can be recycled in many ways.

Keywords: ergonomics, CALL research, learner-centred research, design

Introduction

Our journey towards CALL ergonomics started somewhat by accident. As we were developing ourselves into CALL scholars and language educators, we often stumbled upon incidents where either a learner, or a task or a tool used for learning or teaching was failing us. In other cases, the entire environment seemed to be hostile to the type of learning (mediated by technology) that we were trying to construct, and, inadvertently, some of its elements seemed to fluctuate from one day to the next. Discussing our misadventures with colleagues made us realize that such failure was neither accident nor rare occurrence. Like many other language educators, we were working in an environment where technologies were developing at an exponentially fast rate; they were becoming ubiquitous, somewhat invasive, but oh-so insidiously tempting!

DOI 10.1075/lsse.2.02caw

When research in CALL and other computer-supported learning started booming, it became quickly apparent that the ubiquity of the computer, and the accelerated expansion of the Internet, Web 2.0 technologies, and any other technology-mediated language learning tools, had resulted in a somewhat chaotic situation characterized by a clash of behaviours, excessive "awe" or exaggerated "fear" that Bax (2011) summarized rightly in the following:

> These twin features of excessive 'awe' and exaggerated 'fear' when dealing with new or normalizing technologies serve to exemplify the way in which the relationship between technology and society is frequently conceived in popular accounts, namely in absurdly simplistic and polarised terms. Technologies are popularly presented as being either so powerful that they will undoubtedly change every aspect of our practice, or else so evil as to be entirely harmful, with apparently no middle, nuanced or neutral position possible. (p. 2)

In our journey towards effective CALL research, practice and design, it became clear that we would never be able to comfortably understand the full potential of technologies without really pausing and asking ourselves this simple question: *What are students really doing when they are interacting with technologies?* By delving deeper into several CALL research perspectives, we discovered that *ergonomics,* in the context of both *education* and *web design*, offered many promising avenues (Huh & Hu, 2005; Raby, 2005). In the particular case of CALL, we will see that *educational ergonomics* plays an important role in interaction-based research by providing a conceptual framework that looks specifically at the relationship between the user (herewith the *language learner*) and the instrument (herewith the *technology-mediated tool*). *Web ergonomics*, for its part, offers the engineering support, in particular the methods and the technologies enabling CALL researchers to carry observations on learner-computer interactions (LCI), as well as the criteria, guidelines to analyse and measure the quality of such an interaction (Hamel & Caws, 2010). *CALL ergonomics*, a(n) (interdisciplinary) field slowly establishing itself in CALL research and design, can be hence understood as a blend of both educational and web ergonomics.

In this chapter, our objective is to revisit the field of ergonomics in the light of the current state of *LCI* within the specific context of language learning. Our discussion starts with a review of the core concepts grounding these fields of ergonomics from both educational and web-design perspectives (the *what* of ergonomics), taking into account the various theoretical frames and methodological approaches that enrich CALL research. In rethinking the many options that ergonomics offers, as well as the several directions into which this approach can lead our work, we cover and revisit key concepts and studies. We review the principal theories upon which ergonomics (as applied to language learning) is based and

the ways in which these are put into application through cutting-edge tools and techniques borrowed from the web industry. We then focus more specifically on the field of *CALL ergonomics* by looking at the evidences and motives that support its development. *Why* would we want to apply ergonomic principles to CALL research and practices? Before concluding, we comment on several engineering methods that researchers and practitioners can explore to put the principles of CALL ergonomics into practice. In doing do, we focus on the *How* of ergonomics and argue that methods commonly used in human-computer interaction (HCI), software design (SD) and human-centred design (HCD) constitute excellent complements to current practices in CALL ergonomics, and that, in fact, both these disciplines borrow from each other to enrich their respective fields.

The *What*: Understanding ergonomics in the context of CALL

When we think of CALL research, the term *ergonomics* is not the first one that comes to mind. There are many reasons for this. Originally, ergonomics, from the Greek *ergon*, meaning *work*, referred to a scientific area of research that studied the efficiency of human beings in their working environment (Oxford English Dictionary). In the late 1950s, engineering research appropriated the term to refer more generally to "the study of the interaction of men and their environment (now usually defined with special reference to the machine environment)" (*Engineering* 21 Feb 1958 cited by OED). Soon enough, the concept of *design* became a common element within this field of research. Indeed, it seems natural to think that changes in design of a machine will affect its users' behaviours and the ways in which they interact with it. A call for papers recently published in the scientific review *Ergonomics* is quite revealing of the shift that the discipline has seen since its beginning, and on the desire to explore new grounds of applied research in ergonomics. The editors claim that the field has "a long history of innovations" and welcome manuscripts in fields ranging from psychology to social or cognitive fields, including "new ergonomics methodology," "inter-disciplinary insights," or "case studies involving new concepts/new domains/new wicked problems" (p. 1600). A further examination of recent issues of *Ergonomics* reveals that the field is inherently becoming interdisciplinary while focusing primarily on *effects* and *factors* (two words that appear consistently in titles) of various instruments on humans' physical, psychological, or cognitive attributes or performances. While the instruments in focus might have been essentially related to mechanical work when the field started to evolve, we cannot help but notice a shift in recent years in the type of outcomes, environments, or devices that are being tested: video-games, touchscreens, smart phones, cognitive load, dynamic decision-making, 3D display technologies and

user experience, influence of crowd-sourcing on human perception of information, effects of simulated virtual environment as compared to real environment on human behaviour, or learning transfer from virtual to real environments.

When ergonomics is more specifically applied to a learning environment, we find a similar emphasis on making sure that designs fit users' needs, abilities and likes, hence reducing the effort that needs to be produced while maximizing productivity. While computer-mediated language learning may impose new constraints on learners, designing a system that is ergonomically viable is a way to ease adaptability or, in fact, reduce the cognitive load resulting from constantly adapting to new environments or instruments. Related to the idea of adapting to new instruments, or new teaching and learning concepts, ergonomics will also pay special attention to the skills (functional and cognitive) that may be transferred (from one environment to the next), shifted, developed or adapted.

Coming back to CALL contexts, we can extrapolate that when language learners are interacting with a computer (or a mobile device), or with other human beings, through a computer, the *efficiency* of these interactions will have an impact on the overall *languaging* process. In other words, CALL researchers, like engineers, need to analyse these interactions to potentially enhance the design of part or all of their elements (from the instrument itself to its context of use). To illustrate this necessity, Raby et al. (2003) explained, "it is necessary to examine the learners' interactions not just with an instrument (a computer or a textbook), but with the whole learning system devised by the teachers" (p. 7).

Ergonomics in educational contexts has now become more common, and it is recognized as a strong approach to studying learning interactions. Benedyk, Woodcock and Harder (2009) explained that the "original concept of education ergonomics was introduced by Kao" (p. 237) in 1976, and added that the concept was related to a view of educational institutions as work systems where, according to Kao, one objective was the "effective and successful dissemination of knowledge and cultivation of intellectual sophistication" (as cited in Benedyk et al., 2009, p. 237). Drawing from Kao's views, Benedyk et al. (2009) proposed the following:

> From an ergonomic perspective, learning, being the transformation and extension of the learner's knowledge and/or skills, can be viewed as work, and its 'workplace' is the educational environment in which the learning tasks take place, with the 'learning work' consisting of a series of learning tasks. (p. 238)

In describing the general approach to ergonomics, Bertin and Gravé (2010) referred to Laville's (1976) definition that characterises "ergonomics as a combination of science, technology and art" (p. 10). They added, "As a science, its object is the study of man in his work environment. As a technology, it organizes various fields and disciplines in order to design tools and means of production. As an art,

it consists of using available knowledge to transform a given reality or design into a new reality" (Bertin & Gravé, 2010, p. 10). This description illustrates the interdisciplinary nature of ergonomics, a field that originated from industrial production and design, and one that encompasses such fields as psychology, engineering and sociology.

To properly define ergonomics in the specific domain of CALL, we will take the view that CALL ergonomics constitutes both a methodological and theoretical framework that seeks to describe interactions between users and instruments in a view to ameliorate these interactions so that learning or work can be enhanced. An investigation of the common theoretical perspectives associated with CALL ergonomics will guide us in framing more precisely the relevance of the field.

Theoretical perspectives

Two main schools influence ergonomics. The European school is focused on the activity and the analysis of the interaction between the machine and the user. The American school is more focused on the human factors, which refers to design for human use (Sanders & McCormick, 1989) and, in this regard, is interested in designing the best possible machines or programs (e.g., Raby et al., 2003). These two schools find their roots in specific cognitive and sociocultural theoretical currents. Research in CALL ergonomics, in particular interaction-based research, adopts a user-centred approach that is grounded in mediated activity theory or instrumented activity theory (Rabardel, 1995; Raby, 2005; Vérillon & Rabardel, 1995). The basic precept of these theories is that human beings adapt, change, and learn through their interactions with machines, tools, or other human beings. In other words, these interactions are socially and culturally constructed (e.g., Leontiev, 1981; Rabardel, 1995; Vygotsky, 1978). While Piaget believed that adaptation to new environments was predominantly the result of biological transformations of human beings, Vygotsky (1978), then Leontiev and other sociocultural theorists, considered that most human development was, in fact, the result of an *artificial* process in which the "acquisition of instruments plays a leading role" (p. 82).

At first, the instrumented activity theory could be seen as going against the possibility of reaching a state of *normalization*, that is, a situation in which technology has become so invisible that humans interact with it seamlessly and naturally (see Bax, 2011; Chapter 5, this volume). However, Vérillon and Rabardel (1995) made an important distinction between the tool and the instrument by explaining that the tool (considered here as the initial agent) becomes an instrument once "the subject has been able to appropriate it for himself – has been able to subordinate it as a means to his ends – and in this respect, has integrated it with

his activity" (p. 85). Moreover, when considering learner-computer interactions, one feature that needs to be emphasized is that technologies must *not* be studied as forming a single agent capable of change or transformation (Bax, 2011). Instead, as Bijker (1997) explained, "one must study how technologies are shaped and acquire their meanings in the heterogeneity of social interactions" (as cited in Bax, 2011, p. 6).

Rabardel (1995) considered that the ergonomic method sits within an *anthropocentric* approach, in which humans already possess skills that have the potential to be further developed. Within this approach, artefacts are understood (and analysed) as mediators of activities. Moreover, they are viewed as being the result of human transformations within social practices. Rabardel (1995) described tools that are used for the development or acquisition of knowledge as cognitive instruments and considers that an *instrument* is made up of two elements: an artefact and a scheme of use. This distinction is particularly suitable to the analysis of technologies. Indeed, if we consider that a CALL instrument is characterized by a specific method to use it (defined as the *efficient* use of the instrument), the analysis of the interaction, between this particular artefact and the learner, will potentially reveal the gap between the way in which we think the interaction should occur and what the user of the artefact is actually doing. Rabardel (1995) believed that the *gap*, between the predicted usage and the real usage of the artefacts, is a "sign that users contribute to the conception of how artefacts should be used" (our translation) (p. 124). An example of this gap within CALL is the use of video screen capture technology by language learners when writing (see Chapter 7, this volume). Using the technology allowed learners to *see* their writing process and better reflect on it. Its *affordances* as a documentation and retrospection tool emerged in activity (Baerentsen & Trettvik, 2002; Chapter 3, this volume) when the learners interacted with the technology. Their creative and collaborative usage of this particular *instrument*, coupled with their teacher guidance through careful *task design* and *scaffolding*, enable new affordances to reveal themselves in meaningful action. This must be a goal, as Baerentsen and Trettvik (2002) stated: "Successfully conveying the possibilities for meaningful action offered by a technology to the user should be top priority in the design of interactive systems" (p. 971).

If we consider that learner-computer interactions happen within highly dynamic and complex systems, it is easy to imagine the variations in interactions from a user to the next, hence, the need to observe users' behaviours and re-evaluate the design of systems, software, contexts of learning and/or language learning tasks. Learner persona (see Chapter 6, this volume) can be drawn to model user behaviours with the goal to personalise (interface and content) design options in response to the core trend and idiosyncratic characteristics, e.g., learning styles

and preferences, they identify. This multi-faceted view of interactions with instruments is also evident in Engeström's Activity Model (1987). Considered to represent a third (and on-going) phase of Activity Theory (AT), Engeström's model clearly situates the interactions within a social practice (e.g., Lantolf & Thorne, 2006; see also Figure 2.1 below).

Within that social practice, individuals, or groups of individuals, will typically share an *object* that becomes an outcome through the mediation by the tool/instrument (in our case a technology). That mediation through the technology also occurs within an environment that is regulated by implicit or explicit rules, regulations, norms or conventions (e.g., Lantolf & Thorne, 2006). Let us take the use of micro-blogging (namely Twitter) as an example of technology that can mediate communication between language learners and their peers, and other users (such as native speakers). The *community* is an important facet in the use and success of Twitter. If Twitter is used within a language course, this community will be made up of each user (symbolized by a Twitter identity @name) and their shared interest or sub-group (symbolized by the hashtag #subgroup). Micro-blogging in Twitter is regulated by specific conventions and constraints, such as the 140 characters maximum per message. Using Engeström's AT model, Figure 2.1 illustrates a language learning activity mediated through micro-blogging.

By applying an ergonomic approach to the analysis of interactions within such micro-blogging environments, one could focus on the overall design of the learning tasks to ensure that they are conducive to learning and communicating in the other language. CALL ergonomics presumes that computer-mediated language learning environments constitute complex dynamic systems (see Chapter 4, this volume). These differ from linear systems because they exhibit many elements, agents or processes. Within systems, "produced by a set of components

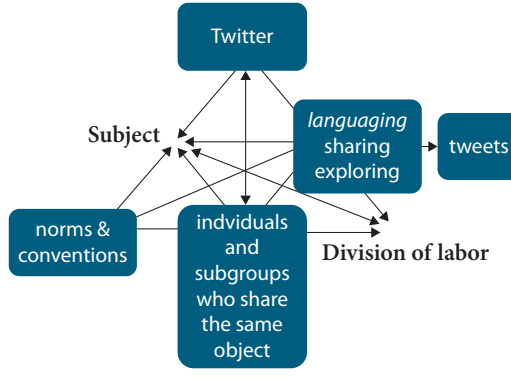


Figure 2.1 Language learning activity mediated through micro-blogging (Twitter)

that interact in particular ways to produce some overall state or form at a particular point in time" (Larsen-Freeman & Cameron 2008, p. 26), *change* is an important feature of the dynamic environment, and the dynamic component is a direct result of the many external and internal elements that may affect or influence it (Larsen-Freeman & Cameron, 2008). It is not difficult to imagine that learning systems and, more particularly, language learning systems that are mediated by technologies constitute highly dynamic systems. Technologies (understood herewith as any tools with which language learners interact) are continuously changing, either because they require updates to better meet their users' demands or (perceived) needs, or because they have been surpassed by other technologies that offer more affordances, i.e., possibilities for *meaningful* action (Baerentsen & Trettvik, 2002). These and other elements of these complex CALL environments constitute *change*; for instance, users come with different cultural, linguistic or social skills, as well as equipped with different technological devices and developing their own personal learning environments (PLE) (Guth, 2009). Computer labs are designed in multiple ways, so are virtual learning platforms (such as Moodle, Blackboard, Canvas, Coursera, Second Life, etc.), and even institutions' policies and practices will affect learning environments due to their unstable nature. Such complexity within CALL defines, in itself, the rationale for further exploring the benefits of educational ergonomics as learners interact in many ways with many artefacts, embracing a global, holistic perspective, focusing on reaching goals rather than acquiring detailed bits of knowledge in a linear fashion (e.g., Bertin & Gravé, 2010). Due to their ubiquitous nature, systems have become embedded cultural artefacts with which individuals interact regularly to perform common and routine tasks (e.g., Selber, 2004; Vérillon & Rabardel, 1995). Consequently, the multiplications of interactions that take place, either within the learning environment or outside of it, have created a situation with no set limits: Learners move back and forth, often unconsciously, between the local and global sphere, sometimes hanging precariously between the personal/private and/or the educational/semi-public spheres.

In conclusion, the field of ergonomics studies individuals at their work place to "describe and interpret these men/machines interactions" (p. 3), in order to "find better ways of adapting machines or technical environments" (p. 3) to the users' characteristics (Raby et al., 2003). Because the user plays a central role in influencing the interactions, ergonomics values a *human factor* (i.e., the usage) while at the same time paying special attention to the tool (i.e., the design) (Rabardel, 1995). A good *fit* between the user, the tool and the context of use, i.e., the environment, is what ergonomics is all about. Let us now look at the aspects that can motivate an ergonomic approach to LCI.

The *Why*: CALL ergonomics as a scientific process

CALL ergonomics presents several advantages to the research, practice and design of activities and learning contexts that are mediated by technology (e.g., Bertin & Gravé, 2010; Raby et al., 2003). Raby et al. (2003) suggested that one of the prime reasons to adopt an educational ergonomic approach is that "the preoccupation of the majority of language students, teachers and researchers is to improve work situations" (p. 4). Another reason, indirectly mentioned by Benedyk et al. (2009), is that educational settings are extremely varied, hence requiring an analysis approach that allows for the identification of constraints to learning. Within such environments, "the task of the ergonomist is to identify design problems for the effective completion of the learning tasks, and to structure solutions" (Benedyk et al., 2009, p. 238). While focusing particularly on distance language learners, Bertin and Gravé (2010) advocated for *didactic ergonomics* because it offers a more accurate representation of a learning situation, based on a dual perspective, "drawing on systemics as well as interactionist theories" (p. 6), that can enhance our comprehension of interactions. Like Bax (2011), they warned against an "unreasoned integration of Information and Communication Technology (ICT) in the classroom (the 'gadget' trend)" (p. 6) and, consequently, feel that "didactic ergonomics has sprung from [the] need to examine how artefacts can be used to instrument the language situation" (p. 6). They explain researchers' trajectory towards an ergonomic approach to language learning as follows:

> If one accepts that the pedagogic relation focuses on the learner, there remains to understand how the other components of the situation can be organized coherently so that the learner-centred process will be facilitated. Another question is raised because the absence in any one of the former models of a technological pole: how should the instrumental (process-oriented) nature of technology be defined in relation to the human actors (the users)? (Bertin & Gravé, 2010, p. 11)

As a type of experimental/field research, CALL ergonomics provides an avenue to validate the use of technologies for language learning and teaching in realistic, authentic environments. By considering educational ergonomics as a field research, Raby et al. (2003) went as far as insisting that observations occur in the authentic physical settings where learners or teachers are working, and not in a laboratory, because "the finest details of a subject's activity are influenced by sociological, cultural, organizational factors which will disappear in the traditional laboratory condition" (p. 4). The question of authentic settings versus laboratory settings will be revisited later.

CALL ergonomics is also a scientific process. As such, it will seek to collect data using various tools (see Part II, this volume) during experiments, with a view toward developing scientific knowledge on mental or behaviouristic models, and taking into account all the possible factors or elements that affect interactions between humans and machines. As a scientific process, educational ergonomics also "combines human expertise with technological potential" (Bertin & Gravé, 2010, p. 6) and seeks to integrate the social, political and cultural relationships with the *technological dimension* (Bertin & Gravé, 2010).

On a more practical level, tracking students can help researchers investigate specific principles of second language acquisition (SLA). Fischer (2007) explained, "researchers can investigate specific principles in relatively carefully controlled conditions in which tracking students' interactions with variables representing those principles gives us a perhaps limited, but relatively unobscured, view of the operation of SLA processes" (p. 429). Aspects such as (socio)linguistic quality of learners' input, methods of negotiating meanings, or signs of metacognitive skills through corrective feedback are examples of data that can be recycled to redesign activities and/or systems. Moreover, as expressed by Fischer (2007), "analyses of tracking data can be used to address practical questions […] provide evidence to make decisions about instructional design […] shed light on the critical question of learner autonomy and the need for learner training […] and place views of students' self-reports in a more realistic context" (p. 429).

Ergonomics is thus profoundly entangled with human factors (Sanders & McCormick, 1989). In CALL, as well as learning in general, the cognitive domain of learning plays a key role. Chalmers (2003) explains that when designing software for learning, some learning theories can help with understanding behaviours. For instance, schema theory that is based on the idea that human understanding is nurtured by schema can explain that users/learners will often have specific cognitive expectations, which explain certain behaviours. Ergonomics based research will help capture these models and assess the level of training that may help adapt new systems to learners. Moreover, Ellis and Goodyear (2010) have considered learning to be situated. As we have seen with complexity theory, activity theory, and the concept of instrumental genesis, the idea of *situatedness* refers to the fact that learning and cognition are social and physical environments and that these environments contribute to the shaping of processes and outcomes. They added, "cognition can be distributed across individuals and artefacts, such that what a single [student] can do on their own may be different from what they can do when working with other people and/or with tools and other physical or digital resources" (Ellis & Goodyear, 2010, p. 26).

In sum, CALL ergonomics provides a promising framework in which the user/learner plays a central role in influencing the interactions, providing rich

data that can be recycled in many ways. Yet the reality is that these notions (related to the computer environment becoming a primary space for knowledge building and creating) have not been fully integrated/digested by the academic community at large except for a few individuals involved in CALL. Research in the learning sciences (e.g., Ellis & Goodyear, 2010), while contributing to the design and iterative enhancement of tools, resources, techniques or processes, value the idea that learning is occurring increasingly through networked systems in which roles and tasks of actors (learners as well as instructors) are constantly shifting. To that effect and realizing the sharp shifts in learning today, Ellis and Goodyear (2010) have made the case that what is often missing from the equation is "good design". Indeed, while human-computer interaction (HCI) has influenced CALL research for some time, other scientific models (such as engineering) could also be influential because they could help specify the structural relations between all entities involved in productive network learning. Having had the chance to better understand the motives behind CALL ergonomics, let us now reflect on existing and promising methods of applying ergonomics based research.

The *How*: Reflecting on CALL ergonomic (evaluation) methods

There are many CALL contexts, in which applying an ergonomic approach to practice, research and design makes sense. In particular, as new CALL systems are being designed and developed at a fairly rapid pace, it has become even more urgent to understand their role and effectiveness in order to (re)assess their requirements, improve their design (e.g., Chalmers, 2003; Colpaert, 2006; Felix, 2005; Hémard, 2006) and enhance the quality of the LCI.

Depending on the goal of the study, ergonomic (evaluation) methods will vary. Part II of this volume, notably Chapters 7–9, gives specific examples of applying sound ergonomic principles to the analysis of computer-mediated interactions in a second language. Some of the techniques and tools used to collect data (e.g., video screen capture, or eye-tracking) allow researchers to take a close and in-depth look at learners' behaviours, whilst collecting data within naturally occurring language learning contexts. The fact that the users (e.g., [pre-service] teachers and learners) are working in authentic language learning contexts while being observed is particularly in line with the settings recommended by Raby et al. (2003). However, some settings are more controlled for reasons that will be discussed later.

CALL system design should be an iterative process (Colpaert, 2006), involving the users early on and all along its various phases of development. An iterative process allows researchers to identify interaction problems as they occur, in

relation to the many variables that are characteristic of a context of learning that is in a constant state of change. While focusing specifically on language courseware (i.e., tutoring systems), Colpaert (2006) opted for the "ADDIE approach (analysis, design, development, implementation, and evaluation), in which each stage delivers output which serves as input for the subsequent stage" (p. 115). Although this approach is described in Colpaert's study as an effective courseware development, it can also be used as a basis for evaluating LCI by inserting an ergonomic methodology at the evaluation stage of the cycle. Like ergonomics, ADDIE is a methodology originally used in engineering, more particularly in computer engineering and software design in an aim to produce systems that have been tested and (re)designed to optimize their effectiveness.

Several ergonomic analyses (also referred to as: measurements, assessments, evaluations) involving (relatively small) groups of participants will facilitate this iterative design process. These analyses are particularly tailored to LCI and can be contrasted to various methods used in HCI and SD where testing with users (either expert evaluators or *real* users) provides data that are constantly reinvested into the (re)design of (further) systems. Assessing through direct manipulations or pure heuristic methods will involve several stages from training and evaluation, to rating, debriefing, and retesting. In HCI, as in LCI, we seek to identify potential user interface errors and successes to further prevent these errors and enhance efficiency of the system (both in terms of content and interface). In a more general manner, the goal is to design a useful and enjoyable experience for the user/learner, reaching what is referred to as *usability* (see Chapter 7, this volume) or *quality in use* (Bevan, 1999). Measuring *quality in use* implies methods that have been carefully devised and embedded in the design cycle of CALL systems.

Ultimately, such an iterative design process should be included in action-research initiatives (Bax, 2011), hence empowering the users, and potentially leading to changes in practices, i.e., to innovations. Discussing the fine line that exists between design and development research and action-research in language didactics, Guichon (2007) has argued that it is not rare that the outcome of a research in this applied discipline leads to an innovation, or the conception of a system (p. 42). Most CALL systems are typically designed by language educators (as part of a team of developers), as an answer to an identified problem or need, with the double purpose of testing a given theory (e.g., SLA) and introducing the system to an already targeted clientele.

General views on methods

Within the field of CALL, while research and development have already led to a better understanding of tools, learning strategies, didactics, or personas (see Part II, this volume), many questions still require empirical investigation, such as the following:

- To what extent do artefacts (i.e., CALL systems) enhance or transform our abilities to communicate, interact, and work with others?
- What types of interactions occur when a learner is connected to a mobile or static device?
- How does the design of a tool, and/or a language-task, affect the learning experience?

These and other questions related to LCI can be explored within an ergonomic research paradigm. To that end, CALL ergonomists will use specific tools and measures to understand and analyse what learners actually *do* when they are working with technology "for the finest details of a subject's activity are influenced by sociological, cultural, organisational factors" (Raby et al., 2003, p. 4). They will perform process-oriented analyses of LCI by means of *learner-task-tool* observations at the computer (e.g., Hamel & Caws, 2010). In order to fully grasp and understand these observations and the behaviours they reveal in a more comprehensive and holistic manner, other types of ergonomic analyses/measurements should be performed, such as needs analyses.

A user needs analysis is an essential first step in setting up research and/or designing new tools, systems or environments. Former experience with the learned language and with technologies can highly influence the success or failure of interactions with new systems being developed. Moreover, learners' metacognitive knowledge and skills have been shown to help learners reinforce their autonomy in such new systems (Hauck, 2005).

Ergonomics also values *behaviours* (verbal and physical) and the mental activity of the user/learner. As noted by Raby et al. (2003):

> Unlike many CALL studies that limit themselves to account for learners' representations or productions, ergonomics also takes into account their behaviours. In order to analyse a work situation, ergonomists or work analysts point out the relationships that unite behaviours and mental processes into a task model.(p. 4)

Hence, CALL ergonomics looks at mental activity (schemas) and behaviours through the task process, as illustrated in Figure 2.2.

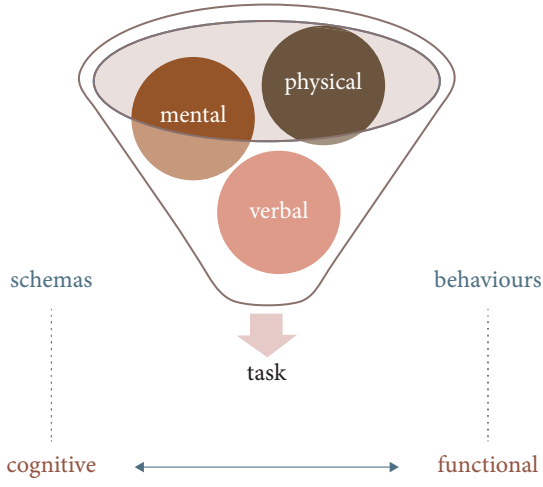


Figure 2.2 Ergonomics' view on schemas and behaviours through the task process

Ergonomics is *context* dependent. By understanding the context and analysing specific learning situations, CALL ergonomics digs deeper within the cognitive and functional effects of new systems, and the effects that new e-learning tasks may have on human cognition and behaviours. It does so not only by placing the user at the centre of the investigation but also by focusing on the processes of learning rather than relying solely on outcomes. A focus on processes is crucial to better assess the learners' abilities to cope with the needs of systems that are becoming more and more dynamic and varied (e.g., online, at home, in collaboration, on a mobile device, etc.). For example, some ergonomics based research on CALL systems that appeared at first glance to be *usable and useful* showed that (some) learners were not always performing well, meaning that (aspects of) the systems were not entirely adapted to their (various) needs (e.g., Caws, 2013; Hamel, 2012). As such, by analysing a "work situation (or the association of a subject and a task in set conditions)" (Raby, 2005, p. 184), empirical data that are collected (physical and verbal behaviours, performances, and processes) can further be recycled into systems design, as well as new learning environments.

Conditions for observations

When discussing the methods used by CALL ergonomists to collect valuable empirical data on LCI for design and/or learning purposes, it is essential to discuss the conditions under which observations occur. First, we need to recall some of the attributes of the theories that frame the research. As said in our *What* section

above, LCI occur within dynamic, complex systems, and the activities that are mediated by technologies involve many components. These components (notably, the space, the actors, the community, the rules and regulations under which the activity takes place, or the specific instrument that mediates the interactions) need to be present and/or considered as potential variables when researchers undergo their ergonomic experiments.

Earlier on in this chapter, we also explained that for most ergonomists, observations of work conditions should occur in the environment where the human being is actually and physically working. Raby et al. (2003) also insisted on this condition being applied to CALL ergonomics research. While we agree in principle, in that the social, cultural or institutional factors do influence learning, we argue that, in fact, if we consider that the physical settings might also negatively affect the LCI (as it is often the case when classrooms and CALL labs are designed without taking into account the needs and requirements of their future users), there is also room for observations in extended (at home) and semi-experimental (in an ergonomic lab) settings. Conditions under which interactions occur can be accommodated, (re)arranged, while keeping the measurements procedures intact so to come closer to finding the optimal contextual settings which will enhance the *quality* of the LCI.

Conditions should match the aim of the experiment. For instance, running usability tests on a system being developed might initially be conducted with a small set of learners only (e.g., Hamel, 2012). Nielsen (1993) explained that after five users, saturation in terms of problems with a system would be reached. Usability tests as per the software design industry are typically run in ergonomic labs, where user behaviour is being monitored individually (Hamel, 2012). When a system prototype has reached maturity, i.e., a functional level robust enough to allow for a wider deployment, then ergonomic evaluations could be conducted in more naturalistic conditions. In Hamel and Séror's study (see Chapter 7, this volume), authentic learning conditions were kept intact so that the LCI processes and behaviours observed were not induced, but rather mirrored the reality.

Other challenges concern the overall settings of experiments with CALL, and, more generally, educational ergonomics. For instance, Benedyk et al. (2009) addressed one of these concerns, as presented by previous studies (such as the one by Kao). Although their study does not concern CALL environments, the challenges that the authors addressed are similar to several issues in CALL contexts. For instance, they explained that one real challenge to the application of ergonomic principles to education contexts is that instead of presenting one "worker" (as is the case in traditional ergonomics), these environments typically feature two main actors, namely the teacher and the learner, who are co-dependant, in that "the measure of effective teaching is successful learning" (Benedyk et al.,

2009, p. 238). Moreover, this co-dependence occurs in somewhat dislocated time and space: Both actors may share the same space or not, and may interact with artefacts in synchrony or asynchrony. They explained,

> Within such a variety in settings exposing learners to a wide variety of influencing factors, some of which are subtle and intangible, the task of the ergonomist is to identify design problems for the effective completion of the learning tasks, and to structure solutions. (p. 238)

To address these challenges, the authors propose a holistic model set in two stages: one stage that focuses on the learner, separating him/her in a single learning context, and another stage where the ergonomic approach extends to include all the external factors that affect the learner interactions (such as the instructor, the physical setting, the artefacts, or the peers) (p. 238).

Ergonomic criteria

The software design industry, like several other industries, relies on the International Organization for Standardization (ISO) standards to ensure that HCI systems being developed comply with sets of internationally approved requirements, specifications and/or guidelines. The prime objective of an ISO standard is indeed that "products and services are safe, reliable and of good quality" (International, n.d.). Ergonomic analyses performed on HCI systems should enable the evaluation of its *usability*. To this end, the ISO 9241 norm, which concerns *Ergonomics of human-system interaction*, stipulates that *usability* is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-11: 1998, definition 3.1).

As an industry evolves, so will the ISO standards that monitor this industry. The notion of *usability*, for instance, was extended to that of *quality in use* to better reflect its user-centredness, as Bevan (2009) recalled:

> This wider interpretation of usability was incorporated in the revision of ISO 9126-1 (2001), renamed "quality in use" as it is the user's perspective of the quality when using a product [3]. The software quality characteristics: functionality, reliability, efficiency, usability, maintainability and portability contribute to this quality. (p. 2)

In the context of CALL, ergonomic analyses performed on LCI systems should equally enable the evaluation of their *usability* or *quality in use*, ensuring that their design follows the same ISO standards, and be guided by the same main three user-centred, goal-specific and context-dependent criteria: *effectiveness*, *efficiency* and

satisfaction. With reference to the ISO 9241 standard, *effectiveness* will be measured against parameters related to the learner's success in achieving the specified goals set by the language task to be accomplished. It focuses on the task outcome: its accuracy and completeness. *Efficiency* will be measured against parameters related to the learner performance in achieving goals set by the language task during its accomplishment. It focuses on the task process: the efforts, the (physical, cognitive) resources deployed by the learner and the time spent on task. *Satisfaction* will be measured against parameters related to the learner perception of task goal achievement, and of software qualities, as stated in the ISO 2196 standard above. It focuses on the learner's experiences, attitudes, beliefs, and feelings. In Hamel (2012, 2013), specific parameters were devised to account for these ergonomic criteria in measuring the usability of an online dictionary for advanced learners of French (see Chapter 7, this volume).

These standards offer a comprehensive set of ergonomic evaluations/analyses that can be used to assess the quality-in-use of HCI/LCI systems.

Ergonomic analyses and evaluations

In our connected Web 2.0 world, the UX (User eXperience) industry is flourishing. It has led to a strong community of UX experts believing that knowing your users, taking their views into account, having them participate in a system design are core in achieving an optimal fit between their goals and the system being developed. Many websites provide useful descriptions of the types of methods and tools that can be used to conduct user-centred (i.e., ergonomics) evaluations, and, in particular, how to assess *usability*. A most-known website is that of the two "fathers" of *usability*: Jacob Nielsen and Don Norman. Called NN/g Nielsen Norman group <http://www.nngroup.com>, this website contains UX research reports (e.g., on *User testing* and, in particular, on *How to conduct usability studies*) and articles (e.g., *Usability 1010, User testing, Web usability*). Other websites, such as that of the Usability Professionals' Association <www.upassoc.org>, also gather usability resources (e.g., *Guidelines* and *Methods*), as well as publications (e.g., *Journal of Usability Studies, UX – User experience magazine*).

Dedicated to the instructional context is the IAR: Instructional Assessment Resources <www.utexas.edu/academic/ctl/assessment/iar/>, which proposes a series of comprehensive modules on how to assess students, teaching, *technology* and programs (even how to conduct research), in an approach very much in line with educational ergonomics (see also Scapin & Bastien, 1997). This assessment approach considers the following stage: *Planning*, *Gathering data* and *Reporting results* in a cyclic and iterative manner. If, for instance, our focus is on assessing

instructional technology, the *Planning* phase comprises five steps: (a) Describe the instructional technology and the learning context; (b) Identify the stakeholders and their needs; (c) Determine the assessment purpose using central questions; (d) Identify how you will use the assessment results; and, (e) Choose the appropriate assessment method(s) and plan implementation.

During the initial phase of an ergonomic evaluation, we should recall the importance of understanding the user context that will eventually help understand the user behaviour. By means of a *task analysis*, a task model can be built, describing the (planned) task components and structure, as well as the user intentions and goals. According to Preece et al., this type of ergonomic analysis is "used to ensure that the conceptual model being developed is working in the way it is intended and that it is supporting the users' tasks" (as cited in Hémard, 2006, p. 266). This will help construct task scenarios for user tests.

Taking into account ergonomic criteria, usability tests put users in real task scenarios and monitor the *efficiency* of the task process, the *effectiveness* of the task outcome and the user *satisfaction*. They can be conducted early (on a paper/wireframe prototype, for instance). However, often at that initial stage of system development, a *walkthrough* method (Hémard, 2003, 2006) will be applied, which consists of providing users with a task script and asking them to verbalise (in a *talk-aloud protocol* or in conversation with the experimenter) the steps taken during the scripted task process. Usability tests can be conducted midway, as formative assessments to identify strengths and weaknesses of versions of functional prototypes. They can even be run comparatively (against similar systems). Comparative measurements are often performed by domain experts with checklists of heuristics (sets of ergonomic criteria) against which systems are compared. That method is called *benchmarking*. An example of benchmarking in a CALL context can be found in Handley and Hamel (2005), where we describe a study aiming at benchmarking speech synthesis for language teaching and learning purposes.

These methods used to elicit empirical user/LCI data can be further classified in direct and indirect methods. Direct methods are often referred to as objective whereas indirect methods are often referred to as subjective. User data are considered objective if prompted directly, in a non-obstructed, non-intrusive manner, with little or no inferences on what is being observed (behaviours, task outcome). On the other hand, user data are considered subjective if they solicit opinions, judgements, or interpretations (e.g., user background, experience, satisfaction). *Observation* and *Usability testing* can be considered direct, objective methods, while *Survey, Interview*, *Focus group* can be considered indirect, subjective methods. *Walkthrough* and *benchmarking* fall somewhere in the middle, since the LCI data elicited is a mixture of observations and interpretations.

Gathering user/LCI data, especially in an educational research context, needs to take into account ethical issues as well. Informed consent must be obtained from users, explaining which types of data will be collected, how and for what purpose. When classroom studies involving learners are considered, the notion of a captive population should be taken into careful consideration. The IAR website discusses ethical questions related to: willingness to participate, anonymity and confidentiality, data security (storage, destruction), and compensation.

Finally, *Reporting results* will imply analysing the data, in a qualitative, quantitative or mixed-manner (we will not expand on this here). Ideally, data should be cross-analysed and attempts should be made to correlate results. Hamel (2013), for instance, has provided an example of how questionnaires can inform usability tests conducted on a dictionary prototype. In that study, learner background and experience (pre-test questionnaire), efficiency and effectiveness scores (usability test) and learner satisfaction (post-test questionnaire) were correlated and showed relationships between experience and performance, satisfaction and success.

The objective of conducting ergonomic analyses is to develop a comprehensive understanding of the user experience, in our case the learner experience, with a focus on their behaviours and mental activity when interacting with/using technology with specific purposes in mind to attain specific language learning goals. Figure 2.3 below summarizes how LCI (the learner experience of technology usage) can be investigated through a comprehensive set of ergonomic analyses.

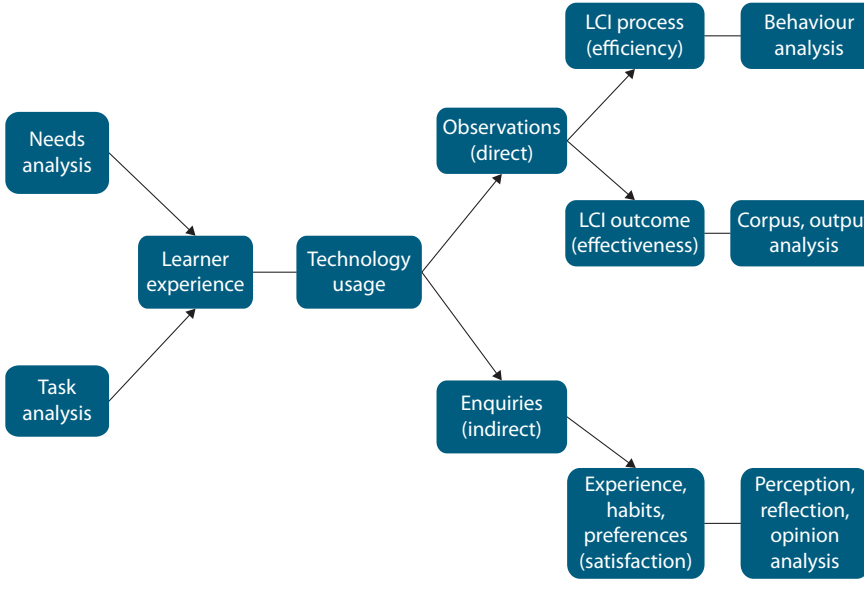


Figure 2.3 A comprehensive set of ergonomic analyses to investigate LCI

Conclusions

Understanding what learners do

CALL ergonomics assumes that learners grow continuously. Under this assumption and as imposed by the complexity of LCI, changes occur constantly, over time and space. Cameron and Larsen-Freeman (2008) compared this non-linearity to mathematics, suggesting that it referred to "a change that is not proportional to input" (p. 31). This non-linearity causes challenges, and the authors proposed that one alternative (amongst others) to face such challenges is to "construct simulated models of the models that explore behaviour over time" (Cameron & Larsen-Freeman, 2008, p. 31). Within such dynamic environments (may they be simulated or authentic), CALL ergonomists will focus one aspect of their work on observing learners to see how behaviours adapt to changes, and how learners' mental models develop through interactions with the systems.

Just as methods and research in HCI, UI and UX found their roots and inspiration in anthropology and ethnography, Fischer (2007) stated, "computer-based tracking can be characterized as a form of ethnography research. As ethnographers enter a community of practice and interview informants to collect data on a sociocultural phenomenon, so, too, can the computer collect data on how students use software" (p. 411). These studies focus more directly on the learners' interactions with specific software, or even specific components of these software, that are considered by Fischer (2007) as tutor (i.e., allowing students to complete language learning exercises) as opposed to tools that permit communication in the L2 via the computer. Interestingly enough, the fast development of computer-mediated communication tools, in particular those mediated by Web 2.0 technologies, has helped tremendously in providing ample data on learners' output, probably less on learners' process, hence the requirement to track learners' processes in a more objective, scientific way.

Observing and understanding learners' behaviours will often lead to surprising evidence, hence the need for heuristic evaluations similar to those used in HCI. In the case of language learning software, for instance, it is common to see users take the fastest route to the targeted item, omit steps in the process, or ignore some of the components of the software. Fischer (2007) added, "the evidence is consistent and compelling; many students make only minimal use of some software components, which raises questions about what constitutes effective instructional design and also has self-evident consequences for software development" (p. 414). While such behaviours may be troubling, it often results from the fact that development of software and tools has been largely influenced by what the designer (who is not necessarily a language learner or instructor)

believes is needed. However, a proper design (and redesign) should be based on learners' observations, needs and goal assessment, interviews, analysis of activities, namely what is described as need-finding in HCI and SD. These observations will often be repeated in a cyclic process.

Broadening the scope of ergonomic measurements

In the CALL research literature about design, Hémard (2003) criticised the scope of usability studies, their lack of longitudinal approach and the fact that CALL design should aim at *acceptability* (the adoption stage for CALL system), even to what Bax (2011) referred to as *normalization* (the integration stage for CALL system), and ultimately to achieve what Levy (2013) referred to as *sustainability* (the green, i.e., maintainable stage for a CALL system). Stakes are high in measuring against these ideal, yet desirable, ergonomic criteria and will involve widening the further concept of *quality in use* to make room for parameters defining efficiency, namely, that take into account learning from errors, and from efforts and time spent during the task process which, in learning situations, can and should be beneficial to language learning (Hamel, 2013).

Hornbaeck (2006), looking at current practice in measuring usability, has also held a similar discourse. Based on a review of 180 usability studies published in HCI journals, the author identified problems related to *how* usability is being measured. Namely, Hornbaeck (2006) stated that (a) domain experts are rarely used in such evaluations; (b) the HCI outcome (effectiveness criteria) is not systematically evaluated; (c) learning and retention factors are not taken into account; (d) there is an unclear relationship made between usage patterns and quality-in-use; (e) satisfaction questionnaires used are not valid instruments; and (f) some studies unknowingly mix objective (observation) and subjective (perception) measures (p. 97). He formulated recommendations (in terms of challenges), namely for "focusing on macro measures, such as those related to cognitively and socially complex tasks, and long-term use" (Hornbaeck, 2006, p. 97).

A recycling metaphor

One aspect of CALL ergonomics that merits particular attention is the recycling metaphor proposed earlier by Caws and Hamel (2013) and inspired by other research and discussions on the learning cycle (e.g., Bertin & Gravé, 2010; for action-research within a neo-Vygotskyan approach, see Bax, 2011; for the ontological iterative process, see Colpaert, 2006). The concept is based on a requirement to run a series of ergonomic measurements and to recycle *everything*: the user/LCI data

collected, the user/LCI analysis results, and the user/LCI data elicitation methods, and to reinvest it into further design and development of systems as well as into pedagogical practice to enrich it (its pedagogical tasks and scenarios) with models of task processes, portraits learners, and personas (see Chapter 6, this volume). Methods can be recycled into teaching (see Chapter 7, this volume); LCI data can be reused for teacher-training purposes (see Chapter 9, this volume); such complex empirical outcomes should be stored as open-access LCI corpora (see Chapter 10, this volume).

In summary, CALL ergonomic research can and must pursue a dual agenda: that of investigating the *learner experience* to shed light on both behaviours and performances in order to optimize CALL design and pedagogical interventions so that both of these reach quality or good learner fit (Chapelle, 2001). Ultimately, ergonomic evaluations, which are powerful and comprehensive methods to elicit user/LCI data, should help inform, perhaps challenge and advance, SLA theory.

At a time when institutions seem to put a lot of emphasis on the acquisition of abilities that students can apply to the work place, it seems quite fitting to understand exactly how they interact with instruments no matter what the learning environment. As such, CALL can rightly claim a right and a role to play in ergonomics based research.

References

Bærentsen, K. B., & Trettvik, J. (2002). An activity theory approach to affordance. *Proceedings of NordiCHI 2002* (pp. 51–60). New York, NY: Association for Computing Machinery.

Bax, S. (2011). Normalisation revisited: The effective use of technology in language education. *International Journal of Computer-Assisted Language Learning and Teaching*, 1(2), 1–15. doi: 10.4018/ijcallt.2011040101

Benedyk, R., Woodcock, A., & Harder, A. (2009). The hexagon-spindle model for educational ergonomics. *Work,* 32(3), 237–248. doi: 10.3233/WOR-2009-0822

Bertin, J. C., & Gravé, P. (2010). In favor of a model of didactic ergonomics. In J. C. Bertin, P. Gravé, & J.-P. Narcy-Combes (Eds.), *Second language distance learning and teaching: Theoretical perspectives and didactic ergonomics* (pp. 1 –36). Hershey, PA: Information Science Reference. doi: 10.4018/978-1-61520-707-7.ch001

Bevan, N. (1999). Quality in use: Meeting user needs for quality. *Journal of Systems and Software*, 49(1), 9–96. doi: 10.1016/S0164-1212(99)00070-9

Bevan, N. (2009). Extending quality in use to provide a framework for usability measurement. In M. Kurosu (Ed.), *Human Centered Design* (pp. 13–22). Proceedings of HCI International 2009, San Diego, California, USA. Berlin: Springer-Verlag.

Bijker, W. (1997). *Of bicycles, bakelites, and bulbs: Towards a theory of sociotechnical change*. Cambridge, MA: The MIT Press.

Chalmers, P. (2003). The role of cognitive theory in human–computer interface. *Computers in Human Behaviour*, 19, 593–607. doi: 10.1016/S0747-5632(02)00086-9

Chapelle, C. (2001). *Computer applications in second language acquisition*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524681

Colpaert, J. (2006). Toward an ontological approach in goal-oriented language courseware design and its implications for technology-independent content structuring. *Computer Assisted Language Learning*, 19(2), 109–127. doi:10.1080/09588220600821461

Caws, C. (2013). Evaluating a web-based video corpus through an analysis of user interactions. *ReCALL*, 25(1), 85–104. doi:10.1017/S0958344012000262

Caws, C., & Hamel, M.-J. (2013). From analysis to training: Recycling interaction data into learning processes. *OLBI Working Papers*, 5, 25–36. doi:10.18192/olbiwp.v5i0.1116

Ellis, R., & Goodyear, P. (2010). *Students' experiences of e-learning in higher education*. New York, NY: Routledge.

Felix, U. (2005). E-learning pedagogy in the third millennium: The need for combining social and cognitive constructivist approaches. *ReCALL*, 17(1), 85–100.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behaviour in CALL. *Computer Assisted Language Learning*, 20(5), 409–442. doi:10.1080/09588220701746013

Guichon, N. (2007). Recherche-développement et didactique des langues. *Les Cahiers de l'ACEDLE*, 4(1), 37–54.

Guth, S. (2009). Personal learning environments for language learning. In M. Thomas (Ed.), *Handbook of research on Web 2.0 and second language learning* (pp. 451–471). Hershey, PA: Information Science Reference. doi:10.4018/978-1-60566-190-2.ch024

Hamel, M.-J. (2012). Testing aspects of the usability of an online learner dictionary prototype: A product- and process-oriented study. *Computer Assisted Language Learning*, 25(4), 339–365. doi:10.1080/09588221.2011.591805

Hamel, M.-J. (2013). Questionnaires to inform user tests in CALL. *International Journal of Computer-Assisted Language Learning and Teaching*, 3(3), 56–76. doi:10.4018/ijcallt.2013070104

Hamel, M.-J., & Caws, C. (2010). Usability tests in CALL development: Pilot studies in the context of the Dire autrement and FrancoToile projects. *CALICO*, 27(3), 491–504. doi:10.1558/cj.v27i3.491-504

Handley, Z., & Hamel, M.-J. (2005). Establishing a methodology for benchmarking speech synthesis for computer assisted language learning. *Language Learning and Technology*, 9(3), 99–120. Retrieved from <http://llt.msu.edu/vol9num3/pdf/handley.pdf>

Hauck, M. (2005). Metacognitive knowledge, metacognitive strategies, and CALL. In J. Egbert & G. Petrie (Eds.), *Research perspectives on CALL* (pp. 65–86). Mahwah, NJ: Laurence Erlbaum Associates.

Hémard, D. (2003). Language learning online: Designing towards user acceptability. In U. Felix (Ed.) *Language learning online: Towards best practice* (pp. 21–42). Lisse, Netherlands: Swets & Zeitlinger.

Hémard, D. (2006). Design issues related to the evaluation of learner-computer interaction in a web-based environment: Activities v. *tasks. Computer Assisted Language Learning*, 19(2–3), 261–276. doi:10.1080/09588220600821586

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64, 9–102. doi:10.1016/j.ijhcs.2005.06.002

Huh, K., & Hu, W. (2005). Criteria for effective CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 9–24). New York, NY: Lawrence Erlbaum Associates.

International Organization for Standardization. (n.d.). Standards. Retrieved from <http://www.iso.org/iso/home/standards.htm>

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, United Kingdom: Oxford University Press.

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford, United Kingdom: Oxford University Press.

Laville, A. (1976). *L'ergonomie*. Paris: PUF.

Leontiev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in soviet psychology* (pp. 37–71). Armonk, NY: Sharpe.

Levy, M. (2013). Design-based research and the quest for normalization in CALL. In J. Rodriguez & C. Pardo-Ballester (Eds.), *Design-based research in CALL* (Vol. 11, pp. 31–40). CALICO Monograph Series. San Marcos, TX: Computer Assisted Language Instruction Consortium.

Nielsen, J. (1993). *Usability engineering*. Cambrige, MA: Academic Press.

Rabardel, P. (1995). *Les hommes et les technologies: Approche cognitive des instruments contemporains*. Paris, France: A. Colin.

Raby, F., Baillé, F., & Bressoux, P., & Chapelle, C. (2003). Ergonomic theory and practice: What language learners do in a self-access room. *Asp enligne*, 41–42. Retrieved from <http://asp.revues.org/1175> doi:10.4000/asp.1175

Raby, F. (2005). A user-centered ergonomic approach to CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL Research Perspectives* (pp. 179–190). New York, NY: Lawrence Erlbaum Associates.

Sanders, M., & McCormick, E. (1987). *Human factors in engineering and design*. NewYork, NY: McGraw-Hill.

Scapin, D., & Bastien, J. (1997). Ergonomic criteria for evaluating ergonomic quality of interface systems. *Behaviour and Information Technology*, 16, 220–231. doi:10.1080/014492997119806

Selber, S. (2004). *Multiliteracies for a digital age*. Carbondale, IL: Southern Illinois University Press.

Vérillon, P., & Rabardel, P. (1995). Cognition and artefacts: A contribution to the study of thought in relation to instrumented activity. *European Journal of Psychology of Education*, 10(1), 7–101. doi:10.1007/BF03172796

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

CHAPTER 3

The theory of affordances

Françoise Blin
Dublin City University, Republic of Ireland

In the last decade, the term "affordance", coined by the ecological psychologist James Gibson (1986), has become a buzzword in CALL research. Often used to denote *possibilities* offered by technologies, the concept has been imported into CALL from cognate domains, such as human-computer Interaction (HCI). However, the CALL community has yet to engage in in-depth discussions on its meaning and usefulness for CALL research and design. The concept remains confusing, often misunderstood, and, at times, misused. This chapter provides an introduction to the concept of affordances, with a view to clarify its meaning and potential applications within CALL. Following a brief overview of Gibson's theory of affordance, it presents and discusses leading HCI interpretations and conceptualizations of affordance that are particularly relevant to CALL researchers and designers. More specifically, it explicates HCI cognitivist and post-cognitivist views of affordances before exploring their relation to CALL affordances and their possible place within a CALL research agenda focusing more particularly on learner-computer interactions.

Keywords: affordances, human-computer interaction (HCI), design, activity theory, phenomenology

Introduction

According to the sociologist Hutchby (2001), "different technologies possess different affordances, and these affordances constrain the ways that they can possibly be 'written' or 'read'" (p. 447). Since the beginning of this millennium, and although it is yet to appear in dictionaries, the term *affordance* has become a buzzword in the human-computer interaction (HCI), Educational Technology and CALL literature, as well as in the public discourse on the integration of digital technologies in education. The term, which was originally coined by ecological psychologist James J. Gibson (1986) "to denote action possibilities provided to the actor by the environment" (Kaptelinin, 2014), was first introduced to the HCI

DOI 10.1075/lsse.2.03bli

community through Norman's (1988) seminal book, *The psychology of everyday things (POET)*, re-edited in 2002 under the title *The design of everyday things* (Norman, 2002). In his original definition, Norman (1988) defined affordances as "the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used" (p. 9), thus implying some relationship between the affordances of a physical product and its *usefulness and usability*.

As Norman's original definition quickly spread throughout the HCI community, "some inherent ambiguities [led] to widely varying usage in the HCI literature" (McGrenere & Ho, 2000, p. 179), perhaps casting some doubt on the usefulness of the concept in design studies and, more particularly, in the domain of Interaction Design. It generated a debate within and between various communities, which is still ongoing today. Different theoretical perspectives have inspired different design models, with varying degrees of success in terms of usability and user experience. As evidenced by the numerous attempts to clarify the meaning of the concept and the ensuing debates, the concept of affordance nevertheless remains a core HCI concept to researchers and practitioners seeking to improve the usability and usefulness of systems and applications.

The lack of clarity and consensus on the meaning of affordances is not the prerogative of HCI researchers and practitioners. Many examples of divergent understandings of affordances can also be found in the educational technology and CALL literature. To my knowledge, however, the CALL community is yet to engage in a theoretical discussion on the meaning of the concept. Numerous educational technology and CALL articles or book chapters have the word *affordances* in their title or in the body of their text – often as part of a collocation, such as educational affordances, learning affordances, pedagogical affordances, cognitive affordances, social affordances, or linguistic affordances. Some propose or investigate lists of affordances offered by various technologies, e.g., Web 2.0 technologies or virtual worlds (Conole & Dyke, 2004; Dalgarno & Lee, 2010; de Haan, Reed, & Kuwada, 2010). Others explore a certain kind of affordances within a given technological context, e.g., "linguistic affordances in telecollaborative chat" (Darhower, 2008). While some authors make their understanding of affordances explicit (e.g., Berglund, 2009; Darhower, 2008; Hoven & Palalas, 2011; Levy & Steel, 2015; Newgarden, Zheng, & Liu, 2015; Zheng & Newgarden, 2012), many do not.

A significant consequence of the lack of clarity about the meaning of affordances is a plethora of taxonomies, design models, and empirical studies that, at best, cannot be compared and, at worst, contain some intrinsic incoherence, due to "divergent ontological and epistemological understandings of the concept" (Bonderup Dohn, 2009, p. 153). Such divergent understandings may lead to tensions and misunderstandings, both at the theoretical and practical levels, in the

design processes and empirical studies that are supposedly informed by a theory of affordances (Bonderup Dohn, 2009). Kaptelinin and Nardi (2012a) make a similar point when they warn us that "unruly theoretical mixing and matching risks illogic and inconsistency" (p. 8). The concept of affordances is probably most useful to CALL researchers and designers seeking to improve the usability, usefulness, and user experience of CALL systems, and to support language learners in their interactions with computers and others speakers of the target language. By mixing and matching incommensurable approaches to affordances, or by combining a view of affordances with an incommensurable theory of second language acquisition or development, our attempts to make our designs more usable, more useful, and enjoyable may be severely constrained. In addition, the validity, reliability, or trustworthiness of our empirical studies is not guaranteed.

This chapter provides an introduction to the concept of affordances with a view to clarify its meaning, so that it can be useful and relevant to CALL researchers and designers. Following a brief overview of Gibson's theory of affordance, it outlines leading HCI interpretations and conceptualizations of affordance that are particularly relevant to the CALL community. More specifically, it explicates cognitivist and post-cognitivist views of affordances before exploring educational and linguistic affordances and their place within a CALL research agenda, focusing more particularly on learner-computer interactions.

Gibson's theory of affordances

Gibson's theory of affordances is an integral part of his *Ecological approach to visual perception* (Gibson, 1986), which marked a departure from "the information-processing paradigm that previously dominated research in the psychology of perception" (Albrechtsen, Andersen, Bødker, & Pejtersen, 2001, p. 7), as well as from Cartesian dualism, which sees mind and body as separate yet interacting entities – i.e., the mind controls the body, and the body can influence the mind (Descartes, 1647). According to Gibson (1986), animals (and humans) pick up information about the environment in which they live directly from the "ambient optic array" which he defines as "a structured arrangement of light with respect to a point of observation" (Gibson, 1970). For Gibson, "action and perception are linked through real-world objects that afford certain forms of action possibilities for particular species or individuals" (Albrechtsen et al., 2001, p. 6). The actor perceives these action possibilities as affordances (Kaptelinin & Nardi, 2012b), which are "what [the environment] offers the animal, what it provides or furnishes, either for good or ill" (Gibson 1986, p. 127). Gibson coined the word *affordance* to mean "something that refers to both the environment and the animal in a way

that no existing term does" and that "implies the complementarity of the animal and the environment" (Gibson, 1986, p. 127). Gibson explains this complementarity (or mutuality) in the following terms:

> [An affordance] is equally a fact of the environment and a fact of behaviour. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer. (Gibson, 1986, p. 129)

From a Gibsonian perspective, affordances are thus action possibilities that are offered by the environment to the animal and that are determined by both the objective properties of the environment and the action capabilities of the animal (Kaptelinin, 2014). For example, "[w]ater affords breathing for a fish, but not for a human. A chair affords sitting for an adult, but not for an infant" (Linderoth, 2012, p. 49). Affordances can be positive or negative, as illustrated by Gibson's own examples:

> [A knife] affords cutting if manipulated in one manner, but it affords being cut if manipulated in another manner. Similarly, but at a different level of complexity, a middle-sized metallic object affords grasping, but if charged with current it affords electric shock. (Gibson, 1986, p. 137)

Another characteristic of affordances is that they "exist irrespective of whether or not they are perceived by the observer" (Kaptelinin & Nardi, 2012b, p. 968):

> The affordance of something does not change as the need of the observer changes. The observer may or may not perceive or attend to the affordance, according to his needs, but the affordance, being invariant, is always there to be perceived. (Gibson, 1986, pp. 138–139)

Finally, as noted by Kaptelinin (2014), Gibson (1986) does not distinguish between animals and humans, nor between natural and cultural environments. Affordances can be provided both by natural objects and by objects created by humans, such as tools, in their attempt to alter the natural environment.

The above overview is but a brief and simplified account of Gibson's theory of affordances. It nevertheless provides an entry point to the exploration of key issues that have been the focus of much debate since Norman's (1988) introduction of the concept to the design and HCI communities. Such issues include the relationship between affordances and perception, the role of culture in the creation and perception of action possibilities for humans, the specificity of tool affordances compared to affordances offered by other *natural* objects, or the role of learning in the perception of affordances (Kaptelinin, 2014).

Affordances in HCI

As noted by Kaptelinin (2014), "the sheer volume of HCI literature that uses the concept of affordances makes it impossible to cover all relevant work" (Sec. 44.3). Attempts at classifying the wide range of HCI perspectives on affordances can be found in the works of Vyas, Chisalita, and Dix (2008), Kaptelinin (2014), and Pozzi, Pigni, and Vitari (2014), to mention but a few.

Different conceptualizations and interpretations of affordance can be loosely attributed to *cognitivist* and *post-cognitivist* HCI. According to Vyas et al. (2006), "[a] cognitivist would describe affordance as a set of observable technology attributes provided by a designer" (p. 93). By contrast, these authors have labelled the post-cognitivist activity theoretical and phenomenological accounts *interaction-centred*, meaning that "affordances of a system emerge during users' actual interaction with it" (Vyas et al., 2008, p. 4). Assigning an interpretation of affordance to one or another of cognitivist or post-cognitivist views is challenging, however. As remarked by Vyas and his colleagues (2006), the paradigm shift observed in HCI between the 1980s and the 1990s did not always translate into a fundamental re-framing of affordances. Baerentsen and Trettvik (2002) argue that this is largely due to the fact that Cartesian dualism still pervades our theories of the mind and of "our environment and our place in it" (p. 51), as well as HCI. They suggest that

> the problem with affordances stems from the attempt to adapt it to the dualistic Procrustes bed of cognitivism, with the result that it is reduced into something fundamentally foreign to Gibson's use of the concept. (Baerentsen & Trettvik, 2002, p. 52)

While being cognisant of possible ontological and epistemological inconsistencies, and perhaps even conflicts, the following sections give an overview of the main *cognitivist* and *post-cognitivist* contributions to the HCI debate on affordances.

The cognitivist view

Early debates within the HCI community have attempted to clarify the meaning of affordances in HCI and have primarily focused on the relationship between affordances and perception. This section briefly examines the contributions of three authors who continue to influence the field, as evidenced by the number of citations they have received to date: Norman (1988, 1999, 2013), Gaver (1991), and McGrenere and Ho (2000).

Having first defined affordances as "the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used" (Norman, 1988, p. 9) – which was a marked departure from Gibson's (1986) view that affordances were independent of perception – Norman (1999) later made a distinction between *perceived* and *real* affordances, before eventually separating affordances, i.e., *real affordances*, from information about them (Norman, 2013).

Gaver (1991), expanding on Norman's (1988) earlier definition, explored "the notion of affordances as a way of focussing on the strengths and weaknesses of technologies with respect to the possibilities they offer the people that might use them" (Gaver, 1991, p. 79). He provided a framework for separating affordances from the perceptual information about them, thus keeping with Gibson's view. This allowed him to distinguish between *correct rejections* and *perceptible, hidden*, and *false* affordances (see Table 3.1 below). According to Gaver (1991), "the actual perception of affordances will […] be determined in part by the observer's culture, social setting, experience and intentions" (p. 81).

Gaver (1991) also introduced the concepts of *sequential* and *nested* affordances, which he saw as required to understand affordances for complex actions. Sequential affordances refer "to situations in which acting on a perceptible affordance leads to information indicating new affordances" (Gaver, 1991, p. 82). Nested affordances refer to grouping of affordances in space, with one affordance serving "as context for another one" (Kaptelinin, 2014, Sec. 44.3.2.1). Finally, Gaver (1991) called for an exploration of "other modes for communicating affordances for action" (p. 83), such as tactile information and sound, which can also give information about affordances.

McGrenere and Ho (2000) discussed the ambiguities in Norman's (1988) original definition and further explored Gibson's (1986) concept of affordance. In line with Gaver (1991), they called for a clear distinction between the existence of affordances and the information that specifies it, while claiming that the former was "independent of the actor's experiences and culture, whereas the ability to perceive the affordance may be dependent on these" (McGrenere & Ho, 2000, p. 180). Stemming from this distinction, they argued for differentiating between

Table 3.1 Separating affordances from the information available about them (adapted from Gaver 1991, p. 80)

Perceptible affordances:	perceptual information is available for an existing affordance
Hidden affordances:	perceptual information is not available for an existing affordance
False affordances:	information suggests a non-existing affordance
Correct rejections:	there is no affordance for a given action, nor information suggesting it

the *usefulness* and *usability* of designs, the former having previously been somewhat neglected by the HCI community. According to them,

> The usefulness of a design is determined by what the design affords (that is, the possibilities for action in the design) and whether these affordances match the goals of the user and allow the necessary work to be accomplished. The usability of a design can be enhanced by clearly designing the perceptual information that specifies these affordances. Usable designs have information specifying affordances that accounts for various attributes of the end-users, including their cultural conventions and level of expertise. (McGrenere & Ho, 2000, p. 184)

McGrenere and Ho (2000) further differentiated between the affordances offered by the "physical" system (e.g., the physical interactions with devices) and by an application, which also offers possibilities for action at different hierarchical levels. Building on Gaver's (1991) nested and sequential affordances, they argued that affordances offered by the software or application "exist (or are nested) in a hierarchy and that the levels of the hierarchy may or may not map to system functions" (McGrenere & Ho, 2000, p. 185). For example, a word processing application affords, at the highest level, writing and editing, and at a lower level clicking, scrolling, dragging and dropping (McGrenere & Ho, 2000, p. 184). However, when a user clicks on a button, his/her goal is not to click on the button per se, but rather to invoke the associated function: "button clickability is nested within the affordance of function invokability" (McGrenere & Ho, 2000, p. 185). In line with Gaver's (1991) concept of sequential affordances, clicking a button may also result in the display of a drop-down menu, giving the user the possibility to then select an option.

Finally, McGrenere and Ho (2000) rejected the binary view of affordances (i.e., an affordance exists or does not exist) and introduced *the degree of an affordance*, which can be used to describe "the ease with which an affordance can be undertaken" (p. 185). In addition, they proposed a second dimension, the *degree of perceptual information*, which "describes the clarity of information that describes the existing affordance" (McGrenere & Ho, 2000, p. 185), in other words, the usability of the design. These two dimensions are incorporated into a framework for design, whose goal is, firstly, to determine the necessary affordances (usefulness) and, secondly, to "maximise each of these dimensions" (McGrenere & Ho, 2000, p. 185), which relate to usability.

The *cognitivist view* of the concept of affordances and its associated models continue to influence designs and empirical studies, not only in HCI but also in a variety of domains, including CALL (see for example Levy & Steel, 2015).

The post-cognitive view

This section will present two approaches that have been very influential in post-cognitivist HCI (Kaptelinin et al., 2003; Kaptelinin, 2014): Leontiev's (1978) activity theory and phenomenology (Heidegger, 1962). Phenomenology and activity theory have some similarities, while being radically different in other aspects. Kaptelinin and Nardi (2012a) noted that both approaches have different points of departure. From an activity theoretical perspective, social (or collective) activities are the interface between subjects and the world: Subjects are constituted by practical activities that transform both themselves and the environment. By contrast, phenomenology is not so much concerned with "how subjects come to exist" but rather how they make sense of their existence and how the world reveals itself to them (Kaptelinin & Nardi, 2012a, p. 51). Another key difference between the two approaches relates to the phenomenological notion of embodiment, which has inspired theoretical and empirical work in HCI and, more particularly, the development of the concept of *embodied interaction*, i.e., "interaction with computer systems that occupy our world, a world of physical and social reality" (Dourish, 2001/2004, p. 3). Although this would be "theoretically plausible" (Baumer & Tomlinson, 2011; Kaptelinin & Nardi, 2012a), activity theoretical HCI has not explicitly explored the role of the body in interactions, except perhaps for Kaptelinin's (1996) work on functional organs (Leontiev, 1981), which "combine natural human capabilities with artefacts to allow the individual to attain goals that could not be attained otherwise" (Kaptelinin & Nardi, 2012a, p. 28).

Kaptelinin (2014) outlines some similarities between activity theory and phenomenology on the one hand, and Gibson's ecological psychology on the other: Despite their different philosophical underpinnings, and despite the fact that neither activity theory nor phenomenology has a theory of affordances as such, the notion of mutuality (or complementarity) of the environment and the actor, as well as a tight relationship between perception and action, can be found in both, albeit in different ways (Kaptelinin 2014, Sec. 44.3.3). Activity theoretical and phenomenological approaches to affordances are said to account for complex affordances, a concept that has emerged in the context of rapid development of complex technologies and which is not fully addressed by cognitivist approaches to affordances.

Simple and complex affordances

Phenomenological and activity theoretical accounts of affordance have become particularly attractive to the Computer Supported Collaborative Work (CSCW)

and Computer Supported Collaborative Learning (CSCL) communities, who have to deal with complex collaborative activities and thus seek to design and evaluate increasingly complex systems and platforms. As noted by Vyas et al. (2008), designers have traditionally decided what affordances should be offered to users of a system. However, as many users actively participate in interactions, they also transform them in unexpected ways. In addition, and according to Vyas et al. (2008),

> the perception and acting out of affordances may lead to reflection on the artefacts, their uses (potential actions) and people's roles (constraints upon actions). Once users are aware of this their perceived affordances change also. (Vyas et al., 2008, p. 8)

Through cycles of change, artefacts and affordances are thus modified, and both embody the practices, norms and values of the community that created and used them (Vyas et al., 2008, p. 8). This dynamic view of affordances is not adequately addressed by Gibson's theory or by its cognitivist interpretations.

The complexity and the dynamical nature of affordances is also the focus of Turner's (2005) work. Turner distinguished between *simple* and *complex* affordances. Simple affordances are those "operating in a classic Gibsonian 'perception-action loop'" (Turner, 2005, p. 788), such as turning a knob to increase the volume of the sound on a device or dialling a number on a phone. While he recognized that simple affordances remain essential to the design and "creation of tangible, ubiquitous and pervasive devices" (Turner, 2005, p. 790), Turner argued that many systems are likely to offer more complex affordances. For example, in the context of a collaborative system, the affordance "highlighting some aspect of an object" is an action that "embodies not only one's perception, but serves to direct the attention of others" (Turner, 2005, p. 792). Turner further observed that, in the case of CSCW, "artefacts mediating cooperation are frequently socially constructed and their affordances can be seen to differ from one workplace to another" (Turner, 2005, p. 793). These artefacts constitute "boundary objects", initially defined by Star (1989) as "common objects [that] form the boundaries between groups through flexibility and shared structure" and whose "materiality derives from action" (Star, 2010, p. 603). Boundary objects develop within and between groups of people, and their affordances embody the culture, history, and practice of these various communities of practice (Wenger, 1998).

Turner (2005) proposed two distinct philosophical approaches that could illuminate how complex affordances may operate: Ilyenkov's (2012) concepts of ideal and significances, and Heidegger's (1962) phenomenology. Turner argued that both Ilyenkov and Heidegger pointed to a similar approach to affordance: "a

thing is identified by its use and that use, in turn, is revealed by way of its affordances/significances" and thus both, directly or indirectly, "equate context and use" (Turner, 2005, p. 787). Turner (2005) concluded that "affordance and context are one and the same" (p. 787).

Turner rooted this conclusion in Ilyenkov's (1977) concepts of ideality and significance, in his work on "the relationship between the material and the ideal in human life" as well as "his formulation of the concept of the artefact" (Cole, 2012, p. 9). An artefact is "an aspect of the material world that has been modified over the history of its incorporation into goal directed human action" (Cole, 2012, p. 9). Artefacts, including technologies, are both material and ideal, as explained by Cole (2012):

> By virtue of the changes wrought in the process of their creation and use, artefacts are simultaneously ideal and material. They are manufactured in the process of goal directed human actions. They are ideal in that their material form has been shaped by their participation in the interactions of which they were previously a part and which they mediate in the present. (pp. 9–10)

Significances are then ideal properties, such as values and meanings, which are acquired by an artefact as the result of purposive activity (Turner, 2005; Turner & Turner, 2002). For Turner and Turner (2002), significance was a *cultural affordance*, i.e., a set of features that arose from the making, using or modifying of the artefact, and which encompassed the values of the culture that created it. Ilyenkov's concept of *ideal-material* artefacts, along with his work on dialectics and contradictions, are foundational concepts of activity theory, which will be explored in later sections.

Relying on Heidegger's phenomenology, Turner (2005) argued that the Heideggerian notions of familiarity, breakdown, and more particularly, equipment, could enhance our understanding and use of complex affordances. According to Heidegger, a world is made of "everyday practices, equipment and common skills shared by specific communities" (Turner, 2005, p. 796). It comprises the totality of interrelated pieces of equipment that are being used for a specific task. It also comprises the set of purposes to which these tasks are put, as well as the identities that are assumed while performing these tasks. We demonstrate our everyday familiarity (i.e., our involvement or *being-in-the-world* and our understanding/know-how of activities) by coping (i.e., dealing "with little or no conscious effort" [Turner, 2013, back cover]) with situations, tools and objects as they present themselves to us, and by our understanding of the referential whole, which is embedded in and manifesting itself in our activities. It is therefore impossible to separate the context, i.e., the world, in which we are active from the action possibilities that present themselves to us. The level and nature of coping are

likely to reveal themselves in response to breakdowns or disturbances that may occur when using a tool (for an extended discussion of coping, including skilful coping, and breakdowns, see Dreyfus [1991, 2014] and Turner [2013]).

Before coming to the specific discussion of these complex affordances in learner-computer interaction in CALL, we will contextualize them in our prevalent approach in Activity Theory.

Affordances in activity-theoretical HCI

Among the different variants of activity theory, the closely related versions proposed by Leontiev (1978) and Engeström (1987/2014) appear to be dominating activity theoretical HCI (for a detailed overview of both versions as they are applied to HCI, see Kaptelinin and Nardi 2012a). For Leontiev (1978), "activity is the basis for psychic phenomena and the fundamental unit of psychological analysis" (Baerentsen & Trettvik, 2002, p. 53). Whereas Leontiev was primarily concerned with activities of individuals, Engeström (1987/2014) extended Leontiev's original model and developed a model of collective activity. In both models, however, human or *life* activity is understood as a systemic, dynamic, and hierarchical formation organized around three layers or constituents – activity, actions, and operations – which relate to needs, intentions, and conditions, respectively.

According to Bødker and Klokmose (2011), this tripartite structure of activities "provides three sets of analytical glasses, each of which focuses on an important aspect of human activity: motivation (by asking why?), goal-orientation (by asking what?) and function (by asking how?)" (p. 320). Activities are collective, oriented toward one or more objects, and motivated by a need, which can be biological, psychological, or social. This motive gives sense and direction to intentional, tool-mediated, and goal-oriented actions, which are carried out through a series of automated operations that are contingent on material conditions. Activities are dynamic in so far that the relationships between these three constituents are flexible, as explained by Lektorsky (2009): "an action can become an activity, a goal can transform into a motive, a task can become an operation, and so on" (p. 77). As noted by Bødker and Andersen (2005, p. 360), human activity is constantly developing as a result of systemic contradictions (Ilyenkov, 1977; Engeström, 1987/2014), and because of the construction of new needs and mediating tools.

Of particular interest to us in the context of this chapter is the activity theoretical re-framing of affordances. One of the key arguments put forward for this re-framing is the limited scope of Gibson's (1986) original theory with regards to the current needs of HCI (Kaptelinin & Nardi, 2012b). As discussed earlier,

cognitivist views of affordances have been criticised for their dualistic underpinnings, which are contrary to Gibson's monist stance (Baerentsen & Trettvik, 2002). Baerentsen and Trettvik (2002) further argued that cognitivist views did not capture activity as a core foundation of the theory of affordances: "objective features of the environment only become affordances when some organizms relate to them in their activity" (p. 54). However, they also suggested that the concept of activity in Gibson's theory was itself underdeveloped, which constituted an obstacle to further applications of affordances in HCI. According to these authors, Gibson's concept of affordances was limited to low-level interactions, i.e., at the level of operations, between the organizm and the environment. Another limitation of Gibson's theory was identified by Kaptelinin and Nardi (2012b), who have argued that the theory has not provided adequate conceptual tools for understanding human actions mediated by historically and culturally constructed tools.

Baerentsen and Trettvik (2002) extended Gibson's theory by matching Leontiev's levels of activity, actions, and operations to three types of affordances: need-related, instrumental, and operational. Need-related affordances relate to motives and needs (activity level), and instrumental affordances – to the action possibilities that are shaped by the socially constructed artefacts available to us (actions level). Operational affordances, i.e., Gibson's original affordances, relate to the level of operations and are further divided into two types: adaptive operational affordances and consciousness operational affordances. Whereas adaptive affordances are the product of human adaptation to the environment as the result of phylogenetical development, consciousness affordances have been learned through active participation in cultural-historical forms of praxis (Baerentsen & Trettvik, 2002, pp. 55–58).

While retaining the notion that affordances are action possibilities offered by the environment to the actor as well as relational properties between the two, and building on Baerentsen's and Trettvik's structure of affordances, Kaptelinin and Nardi (2012b) have proposed a mediated action approach to affordances underpinned by Vygotsky's (1978) concept of tool mediation and by Leontiev's (1978) activity theory. According to them, affordances emerge "in a three-way interaction between actors, their mediational means, and the environments" (Kaptelinin & Nardi, 2012b, p. 974).

Kaptelinin and Nardi (2012b) have identified two levels of direct instrumental affordances offered by a technology: (a) handling affordances, i.e., possibilities for interacting with the technology, and (b) effecter affordances, i.e., possibilities for employing the technology to make an effect on an object (Kaptelinin & Nardi, 2012b, p. 972). For example, "a computer mouse affords moving it on a horizontal surface (handling affordance), which causes changing the pointer's position on the computer screen (effecter affordance)" (Kaptelinin, 2014, Sec. 44.3.3.1.3).

Taken together, handling and effecter affordances "define instrumental technology affordances as possibilities for acting *through* the technology in question on a certain object" (Kaptelinin & Nardi, 2012a, p. 6, emphasis in original).

In real life, however, mediation is heterogeneous, dynamic, and "consists of *webs of mediators*" (Bødker & Andersen, 2005, p. 354). In addition to instrumental affordances, Kaptelinin and Nardi (2012b) have identified auxiliary affordances (e.g., maintenance and aggregation affordances), which emerge "in the complex relations within webs of mediation" (Kaptelinin & Nardi, 2012b, p. 972). They also have remarked that some form of instruction is often needed to enable users access to a tool's instrumental and auxiliary affordances, and thus emphasize the central role of learning affordances. They note that, in the case of digital technologies, learning affordances are often embedded within the technologies themselves (e.g., through tips, help screens, icons and other signs). Finally, Kaptelinin and Nardi (2012b) argued that the action capabilities of the actor are dynamic and "can quickly change as a result of tool switching" (Kaptelinin & Nardi, 2012b, p. 974).

Summary

The previous sections have outlined selected cognitivist and post-cognitivist conceptualizations and interpretations of the concept of affordance within the domains of HCI and interaction design. All share Gibson's original definition as a point of departure. Early Gibsonian and cognitivist HCI views of affordance have been criticised for their limitations in capturing the dynamics and complexity of technological environments and associated human activities, their overemphasis on direct perception, and their focus on the lower end of interactions (i.e., at the operational level). On the other hand, post-cognitivist HCI views of affordance understand them as possibilities for human actions in cultural environments. Affordances are embedded in cultural contexts and emerge in the interactions between active persons, artefacts, and cultural environments. Affordances and actors' capabilities are also dynamic. They can change across time and space, not only as a result of ontogenetic development and learning, but also as a result of breakdowns and new needs, that is, as a result of a re-orientation of the activity in which actors participate.

The nature and the role of artefacts are core to a post-cognitivist view of affordances. From a phenomenological viewpoint, Turner (2005), recalling Wenger (1998), remarked that "all designed artefacts are boundary objects both between and within the communities of practice of designers and users" (Turner, 2005, p. 799). From an activity theoretical perspective, not only do designed artefacts possess the dual characteristics of being simultaneously ideal and material, they

also present two interrelated facets of artefact use: the possible uses and the intended use (Baerentsen & Trettvik, 2002, p. 59). The intended use of a designed artefact can be conceptualized as its ideal form, which encompasses the designer's intentions, cultural-historical meanings and values, as well as his/her vision of *what it is the user should do and why with the artefact*. The possible uses are what users actually do with a given artefact. An unintended use of an artefact may unleash a chain or web of new action possibilities, i.e., new affordances, which, when enacted, will contribute to the transformation of the activities and the environment.

Whether from a cognitivist or post-cognitivist perspective, the concept of affordance provides HCI researchers and interaction designers with conceptual and analytical tools that can help them make interactive technologies more intuitive, more usable, and more useful. Different authors have proposed conceptualizations of affordance that have led to the construction and use of a variety of models supporting design methods and processes, as well as empirical investigations of human-computer interactions. Cognitivist views of affordances are commonly associated with *user-centred designs*. Within this tradition, empirical studies may seek to investigate whether designed affordances are perceived by users with a view to enhance their discovery and their usability or usefulness in relation to pre-determined tasks. Post-cognitivist perspectives are often associated with *activity-centred designs* (Gay & Hembrooke, 2004) and promote a much wider research agenda, including a focus on technology use in dynamic and complex human activities.

The activities at the centre of attention in this book are learner-computer interactions in CALL, which we will now discuss specifically.

Affordances in CALL

Language learning is a dynamic and complex human activity, even more so in technology-rich learning environments (see Chapter 4, this volume). As remarked by Garrett (2009), "'CALL' is not shorthand for 'the use of technology' but designates a dynamic complex in which technology, theory, and pedagogy are inseparably interwoven" (Garrett, 2009, pp. 719–720). Therefore, a theory of affordances of potential use to CALL researchers and designers cannot be reduced to the technological and interaction-design dimensions. Rather, it needs to relate the latter to educational and language affordances, which will be discussed in the following section.

Educational and linguistic affordances

From an ecological perspective on language education, van Lier (2008) emphasizes the relationship between affordances and learning: "[w]hile being active in the learning environment the learner detects properties in the environment that provide opportunities for further action and hence for learning" (van Lier, 2008, p. 598). Learning environments are very diverse in terms of the opportunities they provide. According to Kirschner, Strijbos, Kreijns, and Beers (2004), "education is always a unique combination of technological, social, and educational contexts and affordances" (p. 50). Affordances for learning are thus the combination of technological, social, and educational affordances. Kirschner et al. (2004) proposed a design framework based on two principles, "(a) the systemic and emergent properties of educational, social, and technological affordances and (b) the implementation of interaction design to assure both usability and utility" (p. 63).

Drawing on Gibson's (1986) theory of affordances and on Kirschner (2002), Kirschner et al. (2004) have reminded us that technological, social, or educational affordances are characterized by two relationships. First, there must be a reciprocal relationship between the learner and the learning environment, and, second, there must be a perception-action coupling: Once a need to do something becomes salient, an affordance will be perceived by the learner and will invite and guide him/her to act on it. However, the realisation of the affordance "may depend on factors such as expectations, prior experiences, and/or focus of attention" (Kirschner et al., 2004, p. 50). Furthermore, the technology-mediated learning environment must fulfil the learner's intentions, which must be supported or anticipated by meaningful affordances (Kirschner, 2002; Kirschner et al., 2004).

Focusing more specifically on CSCL environments, Kirschner et al. (2004) related technological affordances to the notion of usability, which is concerned with "whether a system allows for the accomplishment of a set of tasks in an efficient and effective way that satisfies the user" (p. 50), and adopted Kreijns, Kirschner, and Jochems' (2002) definition of social affordances:

> Social affordances are properties of CSCL environment that act as social-contextual facilitators relevant for the learner's social interactions. When they are perceptible, they invite the learner to act in accordance with the perceived affordances, i.e., start a task or non-task related interaction or communication. (p. 13)

Finally, educational affordances are "the characteristics of an artefact that determine if and how a particular learning behaviour could possibly be enacted within a given context" (Kirschner et al., 2004, p. 51). They can be defined as "the relationships between the properties of an educational intervention and the characteristics of the learners that enable particular kinds of learning by them" (Kirschner

et al., 2004, p. 51). Educational affordances can be operationalized through tasks, which offer possibilities for action, interaction between students, and, from a distributed cognition perspective, coaction (Zheng & Newgarden, 2012).

According to Kirschner et al. (2004), "task ownership, task character, and task control are defining factors in the educational affording of environments" (p. 54). Each of these dimensions can be described along a continuum, depending on the level of student vs. teacher engagement and agency. At one end of the continuum, the teacher defines the problem space (task ownership), constructs different elements of the task (task character), and determines who does what (task control). At the other end of the continuum, students take ownership of the task, which can be a real-life problem relevant to them, and exercise control over who does what. A given task will offer different affordances for learning to different learners in different technological, social, and educational contexts (Kirschner et al., 2004).

Tasks are also a crucial element of contemporary language pedagogy and CALL (see Thomas & Reinders, 2010). Drawing on Vygotsky (1986) and Leontiev (1981), the sociocultural perspective on tasks differentiates between "task" and "activity." For Coughlan and Duff (1994), a *task* is "a kind of 'behavioural blueprint' provided to subjects in order to elicit linguistic data" (p. 174). By comparison, an *activity* is what learners actually do when performing the task: "It is the process, as well as the outcome, of a task, examined in its sociocultural context" (Coughlan & Duff, 1994, p. 174).

Zheng and Newgarden (2012) have argued that we need to move from a focus on task to a focus on the design of learning environments "where learners can participate, interact, select, and evaluate the effect of language action" (p. 27). I would further argue that we need to reconceptualize tasks and technology-mediated learning environments as boundary objects between designers, teachers, and learners. The tasks and technologies that are embedded in language learning environments are both ideal and material (physical or digital). For example, tasks are material in so far as they have been reified in the form of descriptors that are likely to include instructions, guidelines, and resources. A language task (or a technology) is also ideal, as it embodies the values, beliefs, and views of language and language learning that have been culturally and historically constructed by the task designer and the learners who engage with it. As learners perform a given task, or use a particular technology, they transform it, giving it a significance, i.e., a cultural affordance (Turner & Turner, 2002).

"An ecological perspective on language learning sees language as part of larger meaning-making resources that include … all the affordances that the physical, social, and symbolic worlds have to offer" (van Lier, 2008, p. 599). For language learning, such an ecology should possess a rich "semiotic budget" (van Lier, 2000, p. 252) that will provide "opportunities for learning to the active, participating

learner" (van Lier, 2000, p. 253). The concept of affordance is thus core to an ecological perspective on language learning, and van Lier (2000) proposed to replace the cognitivist notion of "input" by that of "affordance." He later defined linguistic affordances as "relations of possibilities between language users [that] can be acted upon to make further linguistic action possible" (van Lier, 2004, p. 95).

According to Zheng's (2012) eco-dialogical model, L2 learners need to learn to take skilled linguistic action in order to realise the values of affordances in complex environments, such as massively multiplayer online role-playing games (Newgarden et al., 2015). Drawing on Cowley (2012), they have defined skilled linguistic action as "managing activity under material and cultural constraints" (Newgarden et al., 2015, p. 23). Cowley (2012) said that learners, in taking linguistic action, "link linguistic patterns with affect, artifacts and social skills" (as cited in Newgarden, Zheng, & Liu, 2015, p. 23). Conversely, Newgarden et al. (2015) have argued that evidence of skilled linguistic actions indicates students' linguistic capabilities in terms of accuracy, fluency, and pragmatic competency.

CALL affordances

Drawing on Kirschner et al. (2004), I define CALL affordances as a unique combination of technological, social, educational, and linguistic affordances. Understanding this combination and operationalizing it for design or research can be challenging for two reasons. First, CALL systems come in various forms. Some are specifically designed for language learning, and others are integrated within an institutional virtual learning environment or make use of general applications that have not been designed with language teaching and learning in mind. In either case, the challenge may be to ensure that the affordances – be they technological, social, or educational – that have been embedded in the system can support the emergence, perception, and realisation of linguistic affordances. Second, a theory of affordances may not be directly relevant to some approaches to SLA and to language pedagogy. For example, the concept of linguistics affordances is absent from SLA cognitive interactionist theories, which have been extensively drawn upon in the development of CALL applications, more particularly in the development of tutorial CALL.

Tutorial CALL systems (Levy, 2009) have traditionally provided learners with opportunities to receive help with comprehension or feedback on their production (Chapelle, 2009, p. 745). Opportunities for learner-computer interactions that have been specifically designed for language learning can be construed as affordances under certain conditions. Requesting help when reading a text (e.g., in the form of glosses [Türk & Erçetin, 2014]) or when watching a video clip (e.g., in the form of video captions [Hsu 2015]) are action possibilities offered by a tutorial

CALL system to the active language learner. The provision of feedback also offers possibilities for further linguistic actions. According to Darhower (2008), feedback is an SLA construct that "most closely approximates linguistic affordances" (p. 49). Receiving feedback from the computer gives learners the possibility to notice gaps and to correct their errors (Chapelle, 2009) by enacting technological affordances that have been engineered by CALL designers. However, to become affordances, these action possibilities need to relate to the needs and capabilities of active users.

Intelligent CALL (ICALL) systems (Schulze & Heift, 2013) are particularly promising with regards to the possible engineering and realisation of CALL affordances. For example, the integration of expert and learner models in Intelligent Language Tutorial Systems (ILTS) can provide the basis for a sophisticated user-centred design, whose usability and utility can be enhanced through the implementation of an affordance-based approach to interaction design. Non-tutorial ICALL tools (e.g., grammar checkers, online dictionaries and corpora) that can be accessed when needed in the context of a language learning activity (e.g., writing a text individually or collaboratively, interacting orally or in writing with others in the context of telecollaborative projects, etc.) can also contribute to an increased language awareness (Schulze & Heift, 2013), which, in turn, can help learners pick up linguistic affordances that can be acted upon in the context of a given language learning activity.

While cognitivist HCI interpretations of affordances may have their place in a user-centred design of the basic functionalities of a CALL application, they do not offer conceptual and analytic tools for understanding the perception and realisation of linguistic affordances within the CALL environment. Post-cognitivist HCI interpretations of affordance, however, appear to be a natural fit with ecological perspectives on language learning and CALL. Still, to be a source of affordances, a CALL system must be designed with the concept of affordances in mind (Hoven & Palalas, 2011). Activity-centred design models that strive to integrate technological, social, educational, and linguistic affordances into an overarching design framework remain few and far between. A notable exception is the ecological constructivist framework proposed by Hoven and Palalas (2011) and operationalized in the context of mobile learning.

Empirical studies that are underpinned by post-cognitivist theories of HCI, learning, and SLA not only can assist us in our design endeavours but also can provide new insight into the interaction between the different components of a CALL system and into learner trajectories. Blin, Nocchi and Fowley (2013) have examined the emergence and realization of technological, educational and linguistic affordances of a simulation in Second Life® performed by a group of students of Italian. Some educational affordances had been engineered in the interactional

and action-oriented approach that underpinned the design of the simulation and of its component tasks, as well as their integration in the broader language curriculum. The environment contained many artefacts, such as buildings, objects, scripted objects, notecards, native-speaker avatars, etc., and thus presented a rich semiotic budget to students. Linguistic affordances were expected to emerge as learners began to actively respond to the task and to interact with objects and other avatars. It was indeed observed that linguistic affordances emerged in synchronous and asynchronous multimodal interactions between avatars, and between avatars and some scripted objects carrying semiotic resources that had been placed in the environment. A detailed activity-theoretical and affordance-based analysis of these interactions and, more specifically, of breakdowns, enabled the authors to identify the emergence of learning chronotopes (Bakhtin, 1981) that revealed learners' trajectories across multiple spaces and timescales.

Spatiotemporal features of affordances have also been explored by Zheng and Newgarden (2012) from a dialogical and distributed view of language (Cowley, 2009; Linell, 2009). Exploring language learning activities in virtual worlds (see also Newgarden et al., 2015), they have argued that "the pedagogies that grew from the input-output model do not account for the multiple timescales across which learning occurs or the dynamic, distributed, multimodal nature of meaningmaking" (Zheng & Newgarden, 2012, p. 16), and call for a reconceptualization of language and language learning from language as a code to languaging, i.e., "language as action" (Linell, 2009, p. 273).

Educational and linguistic affordances interact across multiple spaces and on different timescales (e.g., macro and micro levels, respectively), and are connected by social and technological affordances (e.g., at the meso level) offered by learning environments. In the case of virtual worlds, technological affordances include displaying information attached to scripted objects, moving within and across different places, zooming on objects, entering text in the local chat, or activating vocal communication (Blin et al., 2013). Failure to perceive and enact these technological affordances may constrain the realisation of longer-term educational affordances, which, in turn, may impact the emergence of linguistic affordances in unpredictable ways.

Conclusion

Designing affordances for language learning in technology-mediated environments presents a formidable challenge to CALL designers and researchers. This is mainly due to the complexity of the concept itself and to its various interpretations across distinct yet interconnected disciplines. According to Kaptelinin (2014),

> the main challenges for employing new conceptualizations of affordances (or related concepts) in HCI include clarifying the meaning of the concept, as well as its place within a certain research agenda, and making it useful and relevant to designers and other HCI practitioners. Whether or not it can be achieved appears to be critical for determining the future of affordances as an HCI concept. (Section 44.5.3)

As evident throughout this chapter, this is also true for CALL. The future of affordances as a concept in CALL requires an interdisciplinary approach that integrates the notions of educational, social, technological, and linguistic affordances in an ontologically and epistemologically coherent manner: I believe that HCI post-cognitivist views of affordances offer overarching frameworks that are compatible with ecological, activity theoretical, CAS, and distributed views of language and language learning. The future of affordances in CALL will depend on the way such frameworks can be operationalized to be of use to designers and researchers.

References

Albrechtsen, H., Andersen, H. H., Bødker, S., & Pejtersen, A. M. (2001). *Affordances in activity theory and cognitive systems engineering*. Roskilde, Denmark: Risø National Laboratory. Retrieved from <http://orbit.dtu.dk/fedora/objects/orbit:88142/datastreams/file_7726876/content>

Baerentsen, K. B., & Trettvik, J. (2002). An activity theory approach to affordance. In *Proceedings of the second Nordic conference on human-computer interaction* (pp. 51–60). New York, NY: Association for Computing Machinery. doi:10.1145/572020.572028

Bakhtin, M. M. (1981). *Dialogic imagination: Four essays*. Austin, TX: University of Texas Press.

Baumer, E. P. S., & Tomlinson, B. (2011). Comparing activity theory with distributed cognition for video analysis: Beyond "Kicking the Tires." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 133–142). New York, NY: Association for Computing Machinery. doi:10.1145/1978942.1978962

Berglund, T. Ö. (2009). Multimodal student interaction online: An ecological perspective. *ReCALL*, 21(02), 186–205. doi:10.1017/S0958344009000184

Blin, F., Nocchi, S., & Fowley, C. (2013). Mondes virtuels et apprentissage des langues: Vers un cadre théorique émergent. *Recherches et Applications*, 54, 94–107.

Bødker, S., & Andersen: B. (2005). Complex mediation. *Human-Computer Interaction*, 20(4), 353–402. doi:10.1207/s15327051hci2004_1

Bødker, S., & Klokmose, C. N. (2011). The human-artifact model: An activity theoretical approach to artifact ecologies. *Human-Computer Interaction*, 26(4), 315–371. doi:10.1080/07370024.2011.626709

Bonderup Dohn, N. (2009). Affordances revisited: Articulating a Merleau-Pontian view. *International Journal of Computer-Supported Collaborative Learning*, 4(2), 151–170. doi:10.1007/s11412-009-9062-z

Chapelle, C. A. (2009). The relationship between second language acquisition theory and computer assisted language learning. *The Modern Language Journal*, 93, 741–753. doi:10.1111/j.1540-4781.2009.00970.x

Cole, M. (2012). Preface. *The ideal in human activity: A collection of the writings of Evald Vasilyevich Ilyenkov*. Marxists Internet Archive. Retrieved from <https://www.marxists.org/>

Conole, G., & Dyke, M. (2004). What are the affordances of information and communication technologies? *ALT-J*, 12(2), 113–124. doi:10.1080/0968776042000216183

Coughlan, P., & Duff, P. (1994). Same task, different activities: Analysis of a SLA task from an activity theory perspective. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 173–193). Norwood, NJ: Ablex.

Cowley, S. J. (2009). Distributed language and dynamics. *Pragmatics & Cognition*, 17(3), 495–508. doi:10.1075/p&c.17.3.01cow

Cowley, S. J. (2012). Cognitive dynamics: Language as values realizing activity. In A. Kravchenko (Ed.), *Cognitive dynamics in linguistic interactions* (pp. 1–32). Newcastle upon Tyne, United Kingdom: Cambridge Scholars.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

Darhower, M. A. (2008). The role of linguistic affordances in telecollaborative chat. *CALICO Journal*, 26(1), 48–69. doi:10.1558/cj.v26i1.48-69

De Haan, J., Reed, W. M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, 74(2), 74–94. Retrieved from <http://llt.msu.edu/vol14num2/dehaanreedkuwada.pdf>

Descartes, R. (1647). La description du corps humain. In V. Cousin (Ed.), *Oeuvres de Descartes* (Vol. 11). Paris, France: Levrault.

Dourish, P. (2001/2004). *Where the action is: The foundations of embodied Interaction* (Kindle edition). Cambridge, MA: The MIT Press.

Dreyfus, H. L. (1991). *Being in the world: Commentary on Heidegger's "Being and Time", Division 1*. Cambridge, MA: The MIT Press.

Dreyfus, H. L. (2014). *Skillful coping: Essays on the phenomenology of everyday perception and action*. M. A. Wrathall (Ed.). Oxford, United Kingdom: Oxford University Press. doi:10.1093/acprof:oso/9780199654703.001.0001

Engeström, Y. (1987/2014). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge: Cambridge University Press.

Garrett, N. (2009). Computer assisted language learning trends and issues revisited: Integrating innovation. *The Modern Language Journal*, 93, 719–740. doi:10.1111/j.1540-4781.2009.00969.x

Gaver, W. W. (1991). Technology affordances. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 79–84). New York, NY: Association for Computing Machinery.

Gay, G., & Hembrooke, H. (2004). *Activity-centered design: An ecological approach to designing smart tools and usable systems*. Cambridge, MA: The MIT Press.

Gibson, J. J. (1970). Terms used in ecological optics. In J. Pittenger, E. S. Reed, & M. Kim (Eds.), *James Gibson's purple perils: A selection of James J. Gibson's unpublished essays on the psychology of perception*. Retrieved from <http://caribou.cc.trincoll.edu>

Gibson, J. J. (1986). *The ecological approach to visual perception* (Kindle edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heidegger, M. (1962). *Being and time* (J. MacQuarrie & E. Robinson, Trans.). New York, NY: Harper.

Hoven, D., & Palalas, A. (2011). (Re)conceptualizing design approaches for mobile language learning. *CALICO Journal*, 28(3), 699–720. doi:10.11139/cj.28.3.699-720

Hsu, C.-K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL*, 27(01), 84–103. doi:10.1017/S0958344014000214

Hutchby, I. (2001). Technologies, texts and affordances. *Sociology*, 35(2), 441–456. doi:10.1177/S0038038501000219

Ilyenkov, E. V. (1977). *Problems of dialectical materialism*. Moscow, Russia: Progress.

Ilyenkov, E. V. (2012). *The ideal in human activity*. Marxists Internet Archive Publications. Retrieved from <https://www.marxists.org/admin/books/activity-theory/ilyenkov/ideal-activity.pdf>

Kaptelinin, V. (1996). Distribution of cognition between minds and artifacts: Augmentation of mediation? *AI & SOCIETY*, 10(1), 15–25. doi:10.1007/BF02716751

Kaptelinin, V. (2014). Affordances. *The Encyclopedia of Human-Computer Interaction* (2nd ed.) Retrieved from <https://www.interaction-design.org/books/hci.html>

Kaptelinin, V., & Nardi, B. (2012a). *Activity theory in HCI: Fundamentals and reflections*. San Rafael, CA: Morgan & Claypool.

Kaptelinin, V., & Nardi, B. (2012b). Affordances in HCI: Toward a mediated action perspective. *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems* (pp. 967–976). New York, NY: Association for Computing Machinery. doi:10.1145/2207676.2208541

Kaptelinin, V., Nardi, B., Bødker, S., Carroll, J., Hollan, J., Hutchins, E., & Winograd, T. (2003). Post-cognitivist HCI: Second-wave theories. *CHI'03 extended abstracts on Human factors in computing systems* (pp. 692–693). New York, NY: Association for Computing Machinery.

Kirschner, P. (2002). Can we support CCSL? Educational, social and technological affordances for learning. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 7–47). Heerlen, Netherlands: Open University of the Netherlands. Retrieved from <https://www.ou.nl/documents/Promoties-en-oraties/Oraties/Oraties2002/oratieboek_PKI_DEF_Klein_ZO.pdf>

Kirschner, P., Strijbos, J.-W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development*, 52(3), 47–66. doi:10.1007/BF02504675

Kreijns, K., Kirschner: A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society*, 5(1), 8–22.

Lektorsky, V. A. (2009). Mediation as a means of collective activity. In A. Sannino, H. Daniels, & K. D. Gutiérrez (Eds.), *Learning and expanding with activity theory* (pp. 75–87). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511809989.006

Leontiev, A. N. (1978). *Activity, consciousness and personality*. Englewood Cliffs, NJ: Prentice Hall.

Leontiev, A. N. (1981). *Problems of the development of mind*. Moscow, Russia: Progress.

Levy, M. (2009). Technologies in use for second language learning. *The Modern Language Journal*, 93, 769–782. doi:10.1111/j.1540-4781.2009.00972.x

Levy, M., & Steel, C. (2015). Language learner perspectives on the functionality and use of electronic language dictionaries. *ReCALL, FirstView*, 1–20. doi:10.1017/S095834401400038X

Linderoth, J. (2012). Why gamers don't learn more: An ecological approach to games as learning environments. *Journal of Gaming & Virtual Worlds*, 4(1), 45–62. doi:10.1386/jgvw.4.1.45_1

Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept. *Proceedings of Graphics Interface 2000* (pp. 179–186). Montreal, Canada. Retrieved from <http://graphicsinterface.org/wp-content/uploads/gi2000-24.pdf>

Newgarden, K., Zheng, D., & Liu, M. (2015). An eco-dialogical study of second language learners' World of Warcraft (WoW) gameplay. *Language Sciences*, 48, 22–41. doi:10.1016/j.langsci.2014.10.004

Norman, D. A. (1988). *The psychology of everyday things* (Vol. 11). New York, NY: Basic Books.

Norman, D. A. (1999). Affordance, conventions, and design. *Interactions*, 6(3), 38–43. doi:10.1145/301153.301168

Norman, D. A. (2002). *The design of everyday things* (Reprint ed.). New York, NY: Basic Books.

Norman, D. A. (2013). *The design of everyday things: Revised and expanded edition* (Kindle ed.). New York, NY: Basic Books.

Pozzi, G., Pigni, F., & Vitari, C. (2014). Affordance theory in the IS discipline: A review and synthesis of the literature. Retrieved from <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1228&context=amcis2014>

Schulze, M., & Heift, T. (2013). Intelligent CALL. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer assisted language learning* (Kindle edition, pp. 249–266). London, United Kingdom: Bloomsbury.

Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In L. Gasser & M. Huhns (Eds.), *Distributed artificial intelligence* (Vol. 2, pp. 37–54). San Mateo, CA: Morgan Kaufmann.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values*, 35(5), 601–617. doi:10.1177/0162243910377624

Thomas, M., & Reinders, H. (Eds.). (2010). *Task-based language learning and teaching with technology* (Kindle). London, United Kingdom: Bloomsbury.

Türk, E., & Erçetin, G. (2014). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning*, 27(1), 1–25. doi:10.1080/09588221.2012.692384

Turner, P. (2005). Affordance as context. *Interacting with Computers*, 17(6), 787–800. doi:10.1016/j.intcom.2005.04.003

Turner, P. (2013). *How we cope with digital technology*. San Rafael, CA: Morgan & Claypool.

Turner, P., & Turner, S. (2002). An affordance-based framework for CVE evaluation. In X. Faulkner, J. Finlay, & F. Detienne (Eds.), *People and computers XVI – Memorable yet invisible* (pp. 89–103). New York, NY: Springer. doi:10.1007/978-1-4471-0105-5_6

van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford, United Kingdom: Oxford University Press.

van Lier, L. (2004). *The ecology and semiotics of language learning*. Boston, MA: Kluwer. doi:10.1007/1-4020-7912-5

van Lier, L. (2008). Ecological-semiotic perspectives on educational linguistics. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 596–604). Malden, MA: Blackwell. doi:10.1002/9780470694138.ch42

Vyas, D., Chisalita, C. M., & Dix, A. (2008). *Dynamics of affordances and implications for design* (Report). The University of Twente, Netherlands. Retrieved from <http://doc.utwente.nl/64769/1/Affordance_VyasChisalitaDix.pdf>

Vyas, D., Chisalita, C. M., & Van Der Veer, G. C. (2006). Affordance in interaction. *Proceedings of the 13th Eurpoean conference on cognitive ergonomics: Trust and control in complex socio-technical systems* (pp. 92–99). New York, NY: Association for Computing Machinery.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511803932

Zheng, D. (2012). Caring in the dynamics of design and languaging: Exploring second language learning in 3D virtual spaces. *Language Sciences*, 34(5), 543–558. doi:10.1016/j.langsci.2012.03.010

Zheng, D., & Newgarden, K. (2012). Rethinking language learning: Virtual worlds as a catalyst for change. *International Journal of Learning and Media*, 3(2), 13–36. doi:10.1162/ijlm_a_00067

CHAPTER 4

CALL theory

Complex adaptive systems

Mathias Schulze and Kyle Scholz
University of Waterloo, Canada

We conceptualize learner-computer interactions in CALL as a complex adaptive system. Writing from a complexity-theoretical understanding of second language development, we sketch a cognate research paradigm, discuss the characteristics of these interactions as complex adaptive systems. From these characteristics, recent literature on second language development and CALL, we discuss mixed-method methodologies that have the potential to capture the complexity of the non-linear processes of learner-computer interactions in CALL.

Keywords: complexity theory, complex adaptive systems, second language development, digital gaming

Introduction

Processes of learner-computer interaction are complex because a number of actors – learners, instructors, and L1 speakers – and components – computational hardware and software – participate in them and interact with one another. There are also a community and multiple components in their environment, which influence these processes, e.g., other learning materials, linguistic artefacts, and educational institutions. Language learning processes are complex because they involve many internal and environmental variables and components, such as proficiency, aptitude, motivation, and (online) learning environments and materials. These variables are not stable; they interact with one another and are therefore subject to change. To capture the interaction and interdependency of actors and components, and their variables better, we describe language learning processes as dynamic systems. Dynamic systems are, essentially, processes; we prefer the term system since it denotes an integrated whole formed through the interdependence and interaction of its components and variables. In the dynamic change

DOI 10.1075/lsse.2.04sch

of such a system, its variables co-adapt continuously. Because there are so many variables and components in the system, which change and co-adapt, we call such systems, as language learning and learner-computer interaction, complex adaptive systems (CAS). It is very important at this stage to reiterate that CAS are, essentially, complex *processes*. In other words, when we say CAS or just system, we mean the learning, not the learner; we mean the second language development (SLD), not a structure of acquired and applied knowledge; we mean the learner-computer interaction, not the software or the computer; and, we mean the online gaming, not the digital game.

Why use CAS in research, and how did CAS come into applied linguistics and CALL? Since the late 1980s, we witnessed a proliferation of research approaches, concepts, and metaphors of complexity, well beyond mathematics and the natural sciences from where they originated. Books like Gleick's (1987) *Chaos: Making a new science* popularized research on complex and (ostensibly) chaotic systems and made it accessible also for scholars in the social sciences and humanities. Over the three decades since, complexity theory, dynamic systems theory, and chaos theory – related theories that discuss complex processes with a slightly different emphasis – have been applied widely to social phenomena and in areas such as developmental psychology (van Geert, 1994; van Geert & Steenbeck, 2005), bilingualism (Herdina & Jessner, 2002), and pedagogy (Davis & Sumara, 2008). Larsen-Freeman (1997), in her seminal article "Chaos/complexity science and second language acquisition," introduced complex adaptive systems to researchers in applied linguistics and provided the impetus for the evolution of a new research paradigm. In this chapter, we will argue that research on CAS in CALL can provide an integrative and contextualized perspective on learner-computer interactions and language learning processes. We will first sketch the main tenets of a CAS research paradigm, in which (second) language use, second language development (SLD), and learner-computer interaction can be investigated. In the main part, we outline the characteristics of CAS, outline selected previous CAS research in CALL, and suggest methods for analysing learner-computer interactions in CALL from a CAS perspective.

A research paradigm for CAS in CALL

We hope it will become apparent in this chapter that a CAS perspective on learner-computer interaction necessitates a change in our research paradigm. In 1997, Chapelle argued that "CALL would benefit from addressing questions similar to those posed about other L2 classroom learning and from applying the methods used to study L2 learning in other types of classroom activities" (p. 19). As she

asserts, the underlying challenge is the lack of a well-founded and robust research paradigm in CALL. Such a scientific paradigm of "universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of practitioners" (Kuhn, 1996, p. 10) can provide the cornerstones for research in CALL. We need to ask questions about the relevant ontology (what is it we want to know and observe, how can it be categorized?), epistemology (what can we know of it, how can this knowledge be developed?), and methodology (how can we find out about it?). The answers to these questions need to be commensurable so that the scientific paradigm is coherent and the practical research based on it is effective.

For research on CAS in CALL, we pre-suppose – to answer the questions on *ontology* – that language is emergent (Bybee, 1998; Langacker, 2008; MacWhinney, 2006) and consists of fixed, item-based, and abstract linguistic constructions (Tomasello, 2003, 2007). The emergence of language on an individual plane, or that of a speech community, can be observed after recording written and oral language use over periods of time, for example in text corpora. Language use and language development – both in the L1 and the L2 – are in a dialectical relationship. On the one hand, an individual's SLD is a complex process, which is embedded in, and determined and influenced by, social, historical, and cultural processes, and, on the other, each individual participates in the co-construction of social, historical, and cultural processes through his or her second language use (Lantolf, 2006; Lantolf & Thorne, 2006; Swain, Kinnear, & Steinman, 2011). Language learning processes are complex and multivariate (Larsen-Freeman, 1997), and, therefore, SLD is nonlinear.

In our *epistemology*, the language's grammar – essentially a taxonomy of linguistic constructions – is a phenomenon that can be described and explained a posteriori. In other words, it is through language use and subsequent reflection and analysis that linguists and non-linguists alike develop (and can formulate) a grammar of a language or its parts. Fundamentally, usage-based grammar is "epiphenomenal, a by-product of a communication process. It is not a collection of rules and target forms to be acquired by language learners" (Larsen-Freeman, 2002, p. 42). Similarly, we can observe the behaviour of individual language learners over time and infer information about individual cognitive variables. However, when reasoning about observed learner-computer interactions, we need to be aware of the limitations. CAS are deterministic, but cause-effect relationships are complex and often disproportionate and, therefore, frequently unpredictable. This is so, in large part, because of the nonlinear development of the CAS. Thus, moving away from metaphors of (complete) acquisition, we prefer the term second language *development* (SLD) (Verspoor, de Bot, & Lowie, 2011) although our general approach also relies on concepts and findings in second language

acquisition research. Last but not least, CALL is always mediated by computational technologies: In computer-mediated communication, learners interact with other learners of the same language (L2), with L2 instructors, and L1 speakers of that language *via* digital artefacts; in tutorial CALL (Heift & Schulze, 2015; Hubbard & Bradin-Siskin, 2004), learners interact directly with socially, culturally, and cognitively imbued digital artefacts. These digital artefacts are a central component of complex language learning processes.

So it is impossible to predict all future states of a CAS or the state in which the system comes to a rest, i.e., the end state of language learning. For example, we cannot predict with some certainty the exact actions and the communicative success of a learner in a specific language learning task and, even more so, the ultimate attainment of individual language learners in the early stages of their language learning already. However, since future states of CAS are a function of past and current states, it is possible to predict important characteristics of immediately adjacent states of the system – small steps in the language development of the learner – with some probability: How well a particular learner is going to perform the next step in a learning or task sequence, or with what aspects s/he will need help, can be inferred from our observation of prior behaviour and learning outcomes. (This is pretty much what teachers do relying on their experience and intuitions; we need to be able to model the underlying information structures, belief systems, and reasoning processes in learner-computer interactions.) Based on our sustained observation of large groups of learners as individuals, we can also identify states of the CAS – in other words, sub-processes or process segments – through which individual learners or learner types go frequently, or through which they never go, although these states are theoretically feasible.

Thus, the predictive power of complex systems theory is limited, certainly in such complex social systems as computer-learner interactions in language learning. However, rooted in its ontology, this theory has considerable explanatory power. For this, we need to consider appropriate *methods* in the CAS research paradigm, and we will do so in some detail at the end of this chapter.

Language emergent in use

To start our more detailed discussion, we can state that CAS theory intends to "describe and ultimately explain how language as a complex system emerges and develops over time, both as a social instrument in groups and as a private tool in individuals" (de Bot, Lowie, & Verspoor, 2005, p. 117). Emergence is a process in which larger patterns and regularities arise through the interaction of smaller entities. It is central to how a CAS functions and can largely be attributed to the

system's ability to self-organize. This self-organization implies constant change and adaptation that the system undergoes. Although this may seem to imply a lack of agency of the learners and their inability to impact the language learning process, it rather means that the system integrates – synchronically and diachronically – variables and components of the learners, the objectives they constructed and to which they are acting towards, the artefacts they chose and use and through which their actions are mediated, and contextual factors, such as the community and rules, which influence SLD processes. The learner is thus one, albeit important, actor in the CAS – under the dialectic of autonomy and heteronomy – who contributes to and influences the emergent language use of L2.

A usage-based conceptualization of language – in cognitive linguistics and its grammatical frameworks, such as construction grammar – presupposes that as individuals use and encounter the language, they begin to associate its usage with previous experience and construct a taxonomy of usages, informing them about subsequent potential uses. Rather than thinking of grammar as rule-based and language as generated from such rules, language is seen as a collection of patterns that are observed in iterative use. A usage-based view takes grammar to be the cognitive organization of one's experience with language. In the context of CALL, learners accumulate experience with various constructions in different, but also repeated, learner-computer interactions. Aspects of that experience, for instance, the frequency of use of certain constructions, or particular instances of constructions and their salience, have an impact on representation that are evidenced in speakers' knowledge of conventionalised phrases on the individual plane, and in language variation and change on the plane of speech communities (Bybee, 2006, p. 711). "[S]peakers track the frequencies with which variants are used by members of their community and they base their own production frequencies by aggregating this information over many successful interactions" (Blythe & Croft, 2009, p. 60). The main point of reference and the central unit of analysis is the construction, so one suitable framework is Construction Grammar (CG) (see also Schulze & Penner, 2008, on CG in ICALL). CG is an umbrella term for a number of approaches that all view construction as the central linguistic unit (Fischer & Stefanowitsch, 2006; Östman & Fried, 2004). A construction is "a form-meaning pair (F, M) where F is a set of conditions on syntactic and phonological form and M is a set of conditions on meaning and use" (Fischer & Stefanowitsch, 2006, p. 5; Lakoff, 1987, p. 467). Form, meaning, and usage are inextricably linked together, and one cannot be analysed without taking into account the other. Constructions represent linguistic signs of various sizes: from morphs to lexemes and multiword lexemes to abstract syntactic and semantic rules. The logical consequence of representing lexical items, larger linguistic patterns, as well as regular syntactic and semantic phenomena as constructions is the assertion that there

is a continuum from lexical sign to syntactic constructions. CG theorists argue that combining two or more forms usually does not result in a simple concatenation of the meanings the constituents have in isolation (Fried & Östman, 2004, p. 12). Consequently, CG assumes that form and meaning of a construction are not separate, independent modules, but are inseparable and stand in a complex relationship to each other. The constructions of a given language do not simply form an irregular list of all patterns possible in that language. Instead, they reflect the linguistic conventions that the speakers of the language know and form a "structured inventory" of conventions (Langacker, 1987, pp. 63–76). Constructions are also central in research on CAS in SLD (e.g., Ellis & Larsen-Freeman, 2009). In the early stages of SLD, fixed constructions are acquired and normally not yet analysed. For example, in Russian, the phrase "меня зовут" (my name is; me they call) is a passive-substitute construction that learners can only begin to analyse after many years of language use, but they begin to use it as a fixed construction on day one. Over time, item-based constructions emerge. Learners have chunked repeatedly used utterances, but they have not yet analysed the chunks. The slot after constructions, such as *I come from …* and *Ich komme aus …* gets filled by a wide variety of countries, cities, and towns. If this emergent analysis is combined with a focus on form, conscious reflection and noticing, or explicit instruction, then repeatedly used patterns are interpreted as abstract constructions. These can be expressed in rules of varying abstractness and scope, e.g., the German preposition *aus* requires a noun phrase in the dative case to its right. Abstract constructions function at a higher level and encompass various classes of words and grammatical constructions. It has to be noted that the transition of fixed via item-based to abstract constructions is not linear, not common to all constructions, and the speed and stages of transition vary from construction to construction and from learner to learner. Constructions, such as phraseologisms, stabile collocations, and semantically light verb constructions, remain fixed or item-based, e.g., *zwei Fliegen mit einer Klappe schlagen* and *to catch two birds with one stone*, *ein Bad nehmen* and *to take a bath*. Other constructions, such as subject-verb agreement in Indo-European languages, and the two nominative case-marked noun phrases of German copula verbs *sein*, *bleiben*, and *werden* (to be, remain, become), are introduced early on in the language instruction as abstract constructions so that they percolate throughout the SLD.

The emergence of the L2 in SLD in the form of an emerging, more complex taxonomy of constructions, as well as the complex, adaptive nature of SLD, pertains to the analysis and understanding of learner-computer interactions as CAS.

Second language development

Much second language acquisition research is largely based on assumptions about the linear development of language skills with an anticipated end state. This neglects the nonlinear, dynamic, and complex nature of language learning. What is learnt one day that cannot always be remembered and successfully applied the next, and prior learning results are not always the steppingstone for new language learning as sometimes the new internalisation destabilizes the old, or prior learning interferes with the success of current language learning processes. For example, the introduction of the reflexive pronoun *sich* and its declined forms (some of which are homonyms of declined personal pronouns) to learners of German makes some of them temporarily ignore most of what they learned about personal pronouns.

Although nonlinear sub-processes, such as developmental spurts, backsliding, and fossilisation, have been the research focus of SLA, they have often been treated as an anomaly and exception to the (linear) rule. However, such processes are evidence that an L2 is being acquired at varying speeds, which, of course, results in nonlinear developmental trajectories of individuals. Due to individual language-learner differences, this diachronic variation is compounded through the synchronic variability within groups of language learners.

Our current understanding of SLD is both rooted in and addresses limitations of past SLA research. After the focus on learner language as a static system in contrastive analyses of L1 and L2 (Lado, 1957) and error analysis (Corder, 1974), the interlanguage (Selinker, 1974, 1992) continuum – both across groups of learners and for each learner over time – was conceptualized as a huge variety space (Klein, 1986), spanning L1 and L2. Although the comprehensive description of nonlinear developmental trajectories of individuals through and in this variety space, as well as the synchronous variability of individual interlanguages in learner groups, has never been achieved, the study of interlanguage processes – transfer, overgeneralisation, and simplification – provided a new impetus for SLA research as it moved us away from static conceptualizations of learner languages to an understanding of learner language as a complex developmental process. Only the information-processing metaphors of comprehensible input (Krashen, 1982), comprehensible output (Swain, 1985), and learner uptake (for a discussion in CALL, see Smith, 2005) began to shift the focus of SLA research from language and text to the actions of a learner as an information processor. This approach was limited by its concentration on process-external (input) and process-externalised (output) variables. Often, the computation of learner uptake – the part of the input that had been internalised and was now observed in the output – was steeped

in conceptualizing the relationship of input, the language samples to which learners were exposed, and output, the utterances learners produced, as a linear one. The conceptualization of language learning as information-processing meant that the qualities of the learner as a social being with a fluid, multifaceted identity and a rich historical sociocultural pretext and context had to be ignored, reducing the complexity of the language learner and language learning to a few, still inherently textual, variables. A further focus shift toward learning processes came about through the proliferation of interactionist research in SLA (Long, 1996), which assigns a primary role to interaction in language learning and thus views learners as the agents. Yet they remain rather one-dimensional beings because questions concerning why certain interaction opportunities for language learning are not taken up or how specific interactions depend on social, historical, and cultural contexts have hardly been considered. A number of theories and methods – in the context of this book, most notably sociocultural theory (Lantolf & Poehner, 2014; Lantolf & Thorne, 2006; Swain et al., 2011) – consider the variables and components of learning processes and their context comprehensively and situate the complex, non-linear activities of language learning in cultural-historical contexts. Such theories are thus commensurable with CAS theories. The complex dialectical relationship of language learners as subjects to their activities' objectives is mediated by culturally imbued material and ideational artefacts, such as texts in learning materials, electronic dictionaries, online quizzes, wikis, and social media but also language per se and cultural experience. The rich context is considered through the integration of the surrounding community, social and institutional rules, and the division of labour among the collective learner subject into the system of the activity. Sociocultural theory, as a theory of the development of mind, views language learning as a complex dynamic system and learner language as a complex mediating artefact. Learner and artefact, learner and objective, and artefact and objective are in dialectical relationships. This means that their development, or emergence, is propelled by processes of the unity and conflict of opposites, the negation of the negation, and the passage of quantitative changes into new qualitative changes. Central are not only processes of mediation through artefacts in, for example, contextual help in computer environments, but also the dialectic of (cognitive) internalisation and (social) externalisation in (language) learning, and the dialectic relationship of co-learners in the dynamic zone of proximal development.

To sum up our excursion into theories of second language acquisition, we can say that whereas some theories conceptualize language learning as a social process and others as a cognitive, CAS theory strives towards analysing the two in conjunction with one another. From a CAS perspective, Larsen-Freeman (2002) has

argued that "we should be looking for how to connect cognitive acquisition and social use … Forcing us away from reductionism and towards holism" (2002, p. 33).

Characteristics of CAS

As we have attempted to show in the previous sections, conceptualizing second language development and/or learner-computer interaction as CAS is central to our research. Thus, a thorough understanding of the nature of CAS is an essential prerequisite. We base our discussion in this section on the set of CAS characteristics put forth by de Bot and Larsen-Freeman (2011); other theorists – and Larsen-Freeman earlier – have presented similar characteristics (see e.g., Larsen-Freeman & Cameron, 2008a; Sockett, 2013).

i. Sensitive dependence on initial conditions
ii. Complete interconnectedness
iii. Nonlinearity in development
iv. Change through internal reorganization and interaction with the environment
v. Dependence on internal and external resources
vi. Constant change, with chaotic variation sometimes, in which the systems only temporarily settle into "attractor states"
vii. Iteration, which means that the present level of development depends critically on the previous level of development
viii. Emergent properties

Sensitive dependence on initial conditions

When Lorenz (1993) describes the phenomenon behind the butterfly effect, he explains the sensitive dependence on initial conditions in CAS; he was the first to refute claims that small influences on CAS could be neglected, as though they would not cause noticeable effects. In the context of learner-computer interaction, two language learners who are otherwise similar in experience and proficiency can display very different trajectories through a CAS. In a graphical representation, their individual (and perhaps initially common) trajectories deviate from one another, or the curve bifurcates, i.e., a common curve splits at one point in two distinct directions. For example, phonological awareness and L1 literacy are known to be initial conditions that often impact SLD (de Bot, Lowie, & Verspoor, 2007), but the quality and quantity of their impact emerge in their interaction with the many other variables of the system and its context over time. Although the sensitivity to these initial conditions must be considered when analysing the

change, which occurs in a CAS, it must also be noted that the initial conditions are being reflexively altered as the system changes (Larsen-Freeman & Cameron, 2008a). So, we cannot rely on the initial conditions alone to explain all changes at all times, as the conditions that triggered initial and iterative change will reflexively change as the various components of the system continue to interact.

Of course, a challenge arises when attempting to determine which initial conditions are relevant, as the researcher's goal is to understand the change that occurs in the system by first examining the change, and then attempting to discern which conditions may have influenced said change. As de Bot and Larsen-Freeman (2011) mandate, "for our research … we need to have detailed information on the initial conditions if we want to be able to explain differences and similarities in learning outcomes" (p. 10).

In the case of digital game-based language learning, how learners interact in a traditional classroom, how often they play computer games, their experience with forms of digital media, or their desire to communicate with more proficient speakers – all are initial conditions that could influence the individual SLD while playing online games. Factors such as gender, age, and previous language learning experience in any of their L2s can impact their SLD, too, when gaming. According to Larsen-Freeman (1997), "a slight change in initial conditions can have vast implications for future behaviour" (p. 144); this applies to both the gaming and the language learning behaviour. Players who begin the game in an area which is heavily populated by other players, for example, have a better opportunity to interact, and interaction at this early stage of the game can have implications for a player's future social connections with other players, which in turn impacts the gaming learner's linguistic behaviour and possibly their SLD success.

Complete interconnectedness

It is not just the initial-condition variables, which are interconnected with other system-internal and contextual variables. A CAS is completely interconnected; the various components, which comprise the system, i.e., the actors, (digital) artefacts, and factors that determine or influence the process, are connected to one another. If one changes, the others will be impacted to at least some degree. In language, this means that pronunciation, lexis, syntax, meaning, and usage are all interconnected (de Bot & Larsen-Freeman, 2011). So are the variables of SLD. For example, complexity and accuracy of learner texts – both are components of language proficiency – interact. When learners write more complex texts with a more diverse range of more sophisticated linguistic constructions, the accuracy of the text may decrease due to error avoidance, among other factors. It has to be

noted, though, that these two variables are not in a linear relationship. Extrapolating from this small example, it becomes evident that a CAS with many interacting variables is in constant flux; if only one variable changes or is changed through an instructional intervention, the whole system changes through the interconnectedness of its variables and components. This also means that after changing a variable through a teaching intervention, we cannot predict the outcome. Instead, we have to continue observing the CAS as its variables continue to co-adapt and, most likely, continue to induce development through continued interventions, in which the same or other variables are changed.

Nonlinearity in development

Nonlinearity is integral to understanding SLD and learner-computer interaction in CALL (see Schulze, 2008, for a conceptualization of student modelling in ICALL based on CAS), as we have illustrated in previous sections. This characteristic is closely linked to the complete interconnectedness of the CAS, as when one variable impacts another, the SLD may deviate from curricular plans, not result in, or exceed, anticipated learning outcomes, and behavioural trajectories of learners learning a language at the computer often deviate from the steps that software and instructional designers were planning. For example, an affordance (see Chapter 3, this volume) in a digital learning environment, such as access to online lexical resources (dictionary, glosses, or corpus), and the provision of corrective metalinguistic feedback on learner input, will not result in the same quality and quantity of change for each learner, nor will its use at different times and in different contexts result in the same change. This is so because there is not necessarily a linear relationship between cause and effect and between condition and response, due to the complexity and nonlinearity of the CAS. From an analytical (and pedagogic) standpoint, this implies a shift away from affirmative stances of expecting things to happen to embracing the uncertainty in what might happen (Davis & Simmt, 2003). Variability – and therefore nonlinearity – should be appreciated, and, ultimately, "intra- and interindividual variability are important features that should be treated as data and be analysed" (Dijk Davis & Sumara, 2008, p. 62).

Change through internal reorganization and interaction with the environment

Due to the continual nonlinear process of change throughout a CAS, the system itself will reorganize as its many constituent pieces influence one another, especially in the complex interactions with the variables of the environment and

context of the CAS. Context becomes "the landscape over which the system moves, and the movement of the system transforms the context" (Larsen-Freeman & Cameron, 2008a, p. 68). In this regard, co-adaptation is a fundamental aspect of the reorganization and interaction with the environment. In the context of digital gaming, Gee (2006) argues that the "proactive production by players of story elements, a visual-motoric-auditory-decision-making symphony, and a unique real-virtual story produces a new form of performance art coproduced by players and game designers" (p. 61). The very nature of co-production in online gaming signifies the adaptive relationship between player and environment.

Internal and external resources

The internal and external resources of a CAS construct and maintain the system. Internal resources are within the language learner (de Bot & Larsen-Freeman, 2011), e.g., motivation and time to learn, ability to solve problems effectively or to use a computer. The external resources can include the spatial environment being explored or the material artefacts with which the learner interacts (de Bot & Larsen-Freeman, 2011). One might consider the development of young children: As they learn new cognitive and motor skills as internal resources, the variables of the external world around them will change, and both will adapt to one another (de Bot et al., 2007). For instance, in massively multiplayer online games, interaction between non-playing characters and live players is an example of external resources that the learner can utilize when navigating the game environment and learning an L2 at the same time. How the learner's internal resources interact with the external resources will largely define how she or he interacts with the game itself and, in turn, develops proficiency in the L2.

CAS are open systems in that they do not come to a rest at an equilibrium as long as external energy continually enters. In other words, changing external resources trigger, induce, and sustain the change of variables in the system in their interaction with internal resources. In SLD in learner-computer interactions, the external digital resources, such as electronic texts, learning resources, and instructional sequences in online learning environments – all affect change in the CAS, as long as the learner does not preclude them from entering the learning process, the CAS. In instructed language acquisition, it is, of course, the instructor who is the main external resource that induces and sustains change in the students' SLD.

Attractor states

The changing collective variables of a CAS can be operationalized, measured, and then plotted on a time-series graph. To achieve a better depiction of the change in the system, each value y_n can also be conceptualized as a function of $y_{(n-1)}$ and be plotted against this value. The resulting graph is a phase-space portrait. The time that passes between the variable y, having the value $y_{(n-1)}$, and then the value y_n – the lag time – is always identical. In the portrait, one can see that the CAS finds itself in certain states more often than in others and, in some (theoretically possible) states, never at all. Larsen-Freeman and Cameron (2008a) explain that "in the topological vocabulary of system landscapes, states, or particular modes of behaviours, that the system 'prefers' are called attractors" (p. 49). Similar attractors exist in learner-computer interactions: In digital gaming, the reliance on subtitles or other L1 cues in the game world can act as an attractor, being both useful currently and (potentially) hindering the learner's progression at a later state (Sockett, 2013). Figure 4.1 is a small phase space portrait of an individual's proficiency development over two terms. Proficiency is operationalized as textual complexity, accuracy, and fluency. It can be seen that the small, central attractor of proficiency (CAF) emerged through the trade-off effects of accuracy (low) and complexity plus fluency (higher).

The ostensibly chaotic nature of a CAS can make it difficult to figure out the system, but attractor states can provide useful confines within which a system can be analysed, as well as allowing a brief respite in order to determine at least one result of the system, thereby providing some clarity to the otherwise profound complexity. It must be noted as well that although a CAS may find itself in an attractor state, this does not mean that change is no longer occurring; rather, the

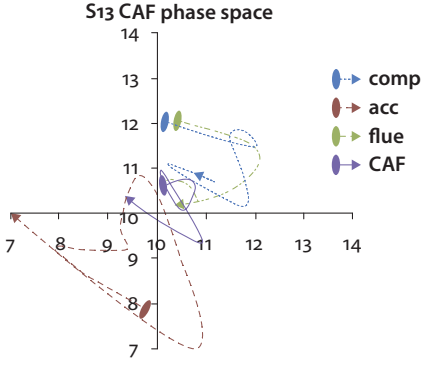


Figure 4.1 Proficiency development of one student over 8 weekly essays

degree to which the system is changing is not (yet) sufficient to transition the system out off the attractor state.

At the opposite end of the continuum, there are CAS states which appear to be possible, but the CAS has never been observed in these states; this would be the white space on the phase-space portrait. These state spaces can be called repellors. In designing and analysing learner-computer interactions, repellors are important, in that they enable both the designer and the researcher to significantly limit the search space for design solutions or analytical algorithms. Simply put, when we have no evidence that learners ever performed a certain interaction or wanted to avail themselves of a certain digital affordance, then it is very unlikely that this interaction or affordance needs to be considered; when we know that learners are attracted to erroneous gender-marking of German nouns, but are repelled by semantic errors (knowing what to mean), then computer feedback for the learner sentence *Die Kollege informiert dich morgen.* / The$_{\text{(fem or plural)}}$ colleague$_{\text{(masc sing)}}$ will inform you tomorrow. / will focus on asking the learner to use the appropriately gender-marked article *der* rather than changing the noun into the plural, or replacing it by its female counterpart, *Kollegin*, to achieve determiner-noun agreement and case-concord.

Iteration

CAS can be observed frequently in the same or similar states (attractors). In parts, this is so because iteration plays a crucial role in a CAS. It is mainly through the many iterations of the CAS that initial conditions gain their influence. In learner-computer interactions, the CAS goes through many small iterations of processes, such as pressing a particular button, making a lexical choice or a grammatical well-formed decision, and requesting learning help by clicking a hyperlink. All of these repeatedly introduce a small change in the CAS, resulting in significant change in the CAS after many iterations.

Emergent properties

Through the iteration, interconnectedness, and self-organization of the CAS, its properties emerge (compare the paradox of the heap ["Sorites paradox," 2014]). Ellis and Larsen-Freeman (2006) explain that "the patterns of language development and of language use are neither innately prespecified in language learners/users nor are they triggered solely by exposure to input" (p. 577). Rather, emerging language is impacted by the interaction with other individuals, societies, cultures,

and their (digital) artefacts; "language and culture are emergent phenomena of an increasingly complex social existence" (Beckner et al., 2009, p. 3).

CAS research in CALL

With our examples in the previous section, we have tried to show not only that CAS are useful in research but also that CAS characteristics are pertinent to learner-computer interaction and SLD. However, thus far, there has been little CAS research in CALL, although a number of scholars have stated the importance of such approaches and their appropriateness to CALL research. Colpaert (2013), for example, argues for an ecological paradigm shift in CALL (which is similar to a shift towards CAS), emphasizing that any single technology alone cannot be responsible for language learning, but rather, learning emerges from the various interacting components that exist in unison with one another. He claims that "no technology possesses an inherent effect on learning, nor on our brain" (Colpaert, 2013, p. 275), and indeed, rather than assume the technology itself has this potential, we should investigate the role of the technology within the CAS and the many other internal and contextual influences. In the following, we will illustrate the applicability of CAS theory in research on learner-computer interactions with selected examples from online and digital game-based language learning.

In the context of extramural language learning, Sockett (2013) observed a group of nine students learning English online informally over the course of three months. All graduate students in applied linguistics maintained blogs to document their experiences. Analysing the 35,000-word corpus, which was derived from their introspective writing, Sockett purports that the English-language learners' strategies can be expressly connected to the characteristics of a CAS, as outlined by Larsen-Freeman and Cameron (2008a), with strategies such as attempting to understand the communicative intentions of other players in online gaming and being exposed to language in authentic contexts that pertain to everyday life, albeit in the digital environment. Sockett and Toffoli (2012), adapting these characteristics, highlighted four primary aspects that are particularly relevant to extramural online language learning: sensitive dependence on initial conditions, attractor states, co-adaptation as a result of the internal reorganization of the system, and nonlinear development. In this study, they situate language learning with social technologies as CAS, moving away from a model of technocratic learner autonomy to one which considers the social roles other members of the online communities play. The informal learning which occurs while university students browsed the internet in their spare time is understood to be emergent in nature. Listening, reading, writing, and vocabulary building

were all in focus as elements of SLD, and they were enhanced by participating in informal online environments, but the development gains of each participant varied wildly due to the frequency and types of interaction that emerged within the various online environments.

In the context of gaming as a learner-computer interaction in CALL, Thorne, Fischer, and Lu (2012) investigate the role that texts in online multiplayer games have in forming what they refer to as complex semiotic ecologies. By analysing the complexity of specific texts which are produced by playing online multiplayer games, playing *World of Warcraft* (WoW) can be better understood as a CAS. Players utilized external resources, such as discussion boards and wikis about the game, to change the internal resources of the digital game environment. Interaction in the game was analysed using various measures of textual complexity (such as lexical sophistication and diversity, syntactic complexity, and readability) and compared to the complexity of text found in these external resources. Thorne et al. found that these external resources were just as rich as the language found within the game and concluded that "external websites function as keystone species within *WoW*'s broader semiotic ecology" (p. 296). They note the validity of analysing such online gaming as CAS, stating that "the reading of texts and the associated action sequences of players form complex and adaptive systems that reorganize themselves based on the contingencies of the immediate goal-directed activity at hand" (p. 298).

Zheng, Young, Wagner, and Brewer (2009), although not positioning their study within a CAS framework, analyse the interactions of their participants, specifically the concept of negotiation of action, as emerging meaning-making behaviour. Playing the synthetic immersive environment *Quest Atlantis*, participants engage in conversations with other players and non-playing characters within the game environment. As quests are undertaken, new goals emerge that are directly related to the internal and external resources of the system.

Focusing on social media and virtual environments, Liou (2012) conceptualizes the interactions in the virtual world *Second Life* as a complex adaptive system, understanding how the learners residing within this environment interact with the environment itself and its many tools (*Second Life* allows almost unlimited modes of content creation) while taking into consideration the affordances of the system. In this study, 25 EFL learners were instructed to perform specific tasks within *Second Life*, such as orienting themselves to the environment and doing peer review. Although the game environment was identical for each student, the external resources of the system, such as unstable internet connections, were alleged to have impacted the development potential of certain students who were either frustrated or could not participate at all, leading to communication breakdowns and the inability to complete tasks. Within the internal resources of

the CAS, users created objects within the game world that were utilized by other players, thereby further impacting the system.

Zheng (2012) also discusses language learning in *Second Life* and how the online environment espouses a conceptualization of CAS and encourages – what Zheng calls – eco-dialogical interaction, whereby "values guide the selection and revision of goals across diverse time-space scales, under which the sociocultural norm 'we' (laws or rules of phonology, syntax, or semantics) are nested" (p. 545). Zheng situates the movement of the player within a virtual environment as being directly related to coordination and cooperation amongst players, which, in turn, leads to communication and SLD. The various and diverse means by which players can complete tasks in the online environment and the ability to interact with other players in an effort to determine how to complete these various goals foster the emergent characteristics and the nonlinearity of SLD within the online environment. She specifically notes that "the meaning-making resources are distributed in virtual spaces, including the macro layout of the physical space, the static clue notes that were designed into the virtual space, dictionaries, and learners' own notes that were collected in their inventories" (Zheng, 2012, p. 555). While some of these aspects are specific to *Second Life*, such as collecting learners' notes in a virtual inventory, the remaining are applicable to any online gaming environment, and they are indicative of the many internal resources of the system.

Marek and Wu (2014) position their research within CALL instructional design, claiming that a CAS theoretical approach should be used. Taking into account as many factors as possible which could influence teaching and learning English as a foreign language (including student and school influences, both internal and external), a CALL ecology model is conceptualized, situating instructional design in CALL as being dependent on these internal and external resources, so that "technology used for CALL is not an end in itself, but a means to an end that is based on fully understanding the educational ecology, determining the desired outcomes, and selecting technology that is most likely to achieve those outcomes" (Marek & Wu, 2014, p. 571).

Methods of analysis of CAS

As we saw in these examples of previous CAS research in CALL, identifying the emergent properties of CAS, such as SLD and learner-computer interaction, and explaining how they emerged are the main goals of such research. In other words, CAS analysis is detecting, localizing, describing, explaining, and interpreting change. Therefore, in each investigation, we first and foremost identify the instantiations of the eight CAS characteristics for this system:

i. What are the initial conditions for this learner-computer interaction? What aspects of *change* in the interaction showed sensitivity to, or depended on, these conditions?
ii. What collective variables, actors, artefacts, and other components induced, influenced, and sustained *change* and development of which aspects of the learner-computer interaction? In which way are the variables, actors, artefacts, and components connected with each other?
iii. What are the trajectories of the process of learner-computer interaction, as a whole of (research-relevant) collective variables, specifically? Which (fractal) patterns of *change* can be identified in the trajectory of an individual and across individuals?
iv. What *change* occurred during the learner-computer interaction? What were the processes and outcomes of the corresponding self-organization of the CAS and of its interaction with the environment?
v. Which internal and external resources led to *change* in the learner-computer interaction, and how?
vi. What is the general nature of the *change* in the CAS? Which attractor and repellor states can be identified? What can these phase spaces tell us about the nature of the CAS?
vii. What are important iterative sub-processes of this learner-computer interaction? How does a particular iteration introduce *change* into the learner-computer interaction?
viii. What properties of the learner-computer interaction emerge in its course, and how do they *change*?

All eight question complexes require the definition and operationalization of CAS-essential and research-relevant variables. Although all variables may not receive equal attention in an analysis of a specific CAS, they are potential factors to consider. Indeed, attempting to analyse everything that occurs within a CAS may be challenging (see Marek & Wu, 2014) and has yet to be fully resolved, beyond admitting that it is an issue (Larsen-Freeman & Cameron, 2008a). We can, however, state that it is the goal of researchers working within a CAS framework to avoid reductionist analyses that attempt to ignore or exclude factors which are hypothesized to not play a role, and instead to consider tendencies, patterns, and contingencies (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008b; van Geert & van Dijk, 2002).

Two guiding principles are particularly important to the selection and application of appropriate methods: (1) long-term, multivariate analyses of language learning processes are necessary; neither reductionist snap shots in cross-sectional

quantitative studies nor isolated qualitative case studies are sufficient to investigate change in learner-computer interaction and SLD in CALL; (2) the complexity of CAS and, consequently, the difficulty with and the low likelihood of predicting their future states accurately mean that we need to identify (qualitative) retrodictive methods of analysis (Dörnyei, 2014). Retrodictive methods – an adjective neologism that denotes the opposite perspective of predictive – reverse the process of analysis so that the outcomes of the CAS are considered first, and then their development is traced back to determine which components and variables induced or caused change. Quantitative approaches (large data sets collected over a period of time and with high density and regular lag time [Larsen-Freeman, 2006; Verspoor et al., 2011]), metaphorical qualitative approaches (e.g., thought experiments [Larsen-Freeman & Cameron, 2008a]), and mixed methods, combining cross-sectional cluster analysis over time with the qualitative analysis of developmental trajectories and outcomes of the language learners – are all also possible. Through these methods, the multitude of interacting variables of the system and its context has to be considered. Of course, the large number of variables in the CAS and its rich context make their continuous observation, as well as their analysis, very challenging. To reduce the high number of degrees of freedom of the CAS, we adopt a technique from molecular dynamics: collective variables. "It is frequently the case that the progress of some … process can be followed by following the evolution of a small subset of generalised coordinates in a system. When generalised coordinates are used in this manner, they are typically referred to as reaction coordinates, *collective variables*, or order parameters, often depending on the context and type of system" (Tuckerman, 2008, n.p., our emphasis). Collective variables, such as proficiency and motivation, are thus dynamic configurations of smaller variables and are essential to describing the developmental change of the CAS. Collective variables have been introduced to, and employed in, applied linguistics research (see e.g., Larsen-Freeman & Cameron, 2008a) because they help to avoid reducing the number of variables through experimental and/or statistical elimination or isolation.

Essentially, all analysis of CAS is an analysis of their change over time. This means that a research design of experimental and control group is seldom necessary. Instead, the different states of an individual process are compared iteratively. Commonalities and differences matter in that both provide clues about from where and how the change originated and was influenced. These individual processes, the CAS of one learner's interaction with the computer, are compared iteratively with comparable states of the CAS of comparable learners.

Conclusion

CAS theory welcomes the variability of actors, components, and factors in the system and its context and the change that results. Davis and Sumara (2008) argue that "given the idiosyncratic characters, recursively elaborative, and ever-divergent possibilities of complex phenomena, accounts of complexity-informed research can never be offered as events to be replicated or even held up as models" (p. 42). Yet through their more realistic depiction of complex nonlinear processes in context, CAS offer new insight into learner-computer interaction. Such insights not only further research in CALL, but also provide a basis for considered, contextualized design decisions in the creation of online learning environments, digital artefacts, and learning materials and help identify sub-processes suitable for pedagogic interventions. They reconcile former ostensible contradictions through their consideration of complex phenomena and processes in specific contexts. And, most importantly, complex adaptive systems offer ways of inducing change of one aspect of the CAS and tell us that we should not expect a wholesale, linear result of that change, but need to continue to observe and analyse, before evaluating the change in the system and possibly inducing further change … *Ein fortgesetzter Versuch* (Wolf, 1979).

References

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., … Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59(1), 1–26. Retrieved from <http://cnl.psych.cornell.edu/pubs/2009-LACAS-pos-LL.pdf> doi:10.1111/j.1467-9922.2009.00534.x

Blythe, R. A., & Croft, W. A. (2009). The speech community in evolutionary language dynamics. In N. C. Ellis & D. Larsen-Freeman (Eds.), *Language as a complex adaptive system* (pp. 47–63). Hoboken, NJ: Wiley-Blackwell.

Bybee, J. (1998). The emergent lexicon. CLS34: The Panels. *Chicago Linguistics Society* (pp. 421–435). Chicago, IL: Chicago Linguistic Society.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733. doi:10.1353/lan.2006.0186

Chapelle, C. A. (1997). Call in the Year 2000: Still in search of research paradigms? *Language Learning & Technology*, 1(1), 19–43.

Colpaert, J. (2013). Peripatetic consideration on research challenges in CALL. In P. Hubbard, M. Schulze, & B. Smith (Eds.), *Learner-computer interaction in language education* (pp. 272–279). San Marcos, TX: Computer Assisted Language Instruction Consortium.

Corder, S. P. (1974). Error analysis. In J. P. B. Allen & P. Corder (Eds.), *The Edinburgh course in applied linguistics* (Vol. 3, pp. 122–131). London, United Kingdom: Oxford University Press.

Davis, B., & Simmt, E. (2003). Understanding learning systems: Mathematics education and complexity science. *Journal for Research in Mathematics Education*, 34(2), 137–167. doi:10.2307/30034903

Davis, B., & Sumara, D. (2008). Complexity as a theory of education. *Transnational Curriculum Inquiry*, 5(2), 33–34.

de Bot, K., & Larsen-Freeman, D. (2011). Research second language development from a dynamic systems theory perspective. In M. H. Verspoor, K. de Bot & W. Lowie (Eds.), *A dynamic approach to second languaged Development* (pp. 5–24). Amsterdam, Netherlands: John Benjamins. doi:10.1075/lllt.29.01deb

de Bot, K., Lowie, W., & Verspoor, M. (2005). Dynamic systems theory and applied linguistics: The ultimate "so what"? *International Journal of Applied Linguistics*, 15(1), 116–118. doi:10.1111/j.1473-4192.2005.0083b.x

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21. doi:10.1017/S1366728906002732

Dörnyei, Z. (2014). Researching complex dynamic systems: "Retrodictive qualitative modeling" in the language classroom. *Language Teaching*, 47(1), 80–81. doi:10.1017/S0261444811000516

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics: Introduction to the special issue. *Applied Linguistics*, 27(4), 558–589. doi:10.1093/applin/aml028

Ellis, N.C., & Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. In N. C. Ellis & D. Larsen-Freeman (Eds.), *Language as a complex adaptive system* (pp. 90–125). Hoboken, NJ: Wiley-Blackwell.

Fischer, K., & Stefanowitsch, A. (2006). Konstruktionsgrammatik: Ein Überblick. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 3–17). Tübingen, Deutschland: Stauffenberg Verlag.

Fried, M., & Östman, J.-O. (2004). *Construction grammar in a cross-language perspective*. Amsterdam, Netherlands: John Benjamins. doi:10.1075/cal.2

Gee, J. P. (2006). Why game studies now? Video games: A new art form. *Games & Culture*, 1(1), 58–61. doi:10.1177/1555412005281788

Gleick, J. (1987). *Chaos: Making a new science.* New York, N.Y.: Viking.

Heift, T., & Schulze, M. (in press). Tutorial CALL. *Language Teaching*.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Clevedon, United Kingdom: Multilingual Matters.

Hubbard, P., & Bradin-Siskin, C. (2004). Another look at tutorial CALL. *ReCALL*, 16(2), 448–461. doi:10.1017/S0958344004001326

Klein, W. (1986). *Second language acquisition*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511815058

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford, United Kingdom: Pergamon.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226458106.001.0001

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: The University of Michigan Press.

Lakoff, G. (1987). *Women, fire, and dangerous things. What categories reveal about the mind*. Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226471013.001.0001

Langacker, R. W. (1987). *Foundations of cognitive grammar (Vol. 1). Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langacker, R. W. (2008). The relevance of cognitive grammar for language pedagogy. In S. De Knop & T. De Rycker (Eds.), *Cognitive approaches to pedagogical grammar. A volume in honour of René Dirven* (pp. 7–35). Berlin, Germany: Mouton de Gruyter.

Lantolf, J. P. (2006). Language emergence: Implications for Applied Linguistics – A sociocultural perspective. *Applied Linguistics*, 27(4), 717–728. doi:10.1093/applin/aml034

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, United Kingdom: Oxford University Press.

Lantolf, J. P., & Poehner, M. (2014). *Sociocultural theory and the pedagogical imperative in L2 education. Vygotskian praxis and the research/practice divide*. New York, NY: Routledge.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18(*2), 141–165. doi:10.1093/applin/18.2.141

Larsen-Freeman, D. (2002). Language acquisition and language use form a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization: Ecological perspectives* (pp. 33–46). London, United Kingdom: Continuum.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619. doi:10.1093/applin/aml029

Larsen-Freeman, D., & Cameron, L. (2008a). *Complex systems and applied linguistics*. Oxford, United Kingdom: Oxford University Press.

Larsen-Freeman, D., & Cameron, L. (2008b). Research methodology on language development from a complex systems perspective. *The Modern Language Journal*, 92(2), 200–213. doi:10.1111/j.1540-4781.2008.00714.x

Liou, H.-C. (2012). The roles of Second Life in a college computer assisted language learning (CALL) course in Taiwan, ROC. *Computer Assisted Language Learning*, 25(4), 365–382. doi:10.1080/09588221.2011.597766

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & B. T. K (Eds.), *Handbook of second language acquisition*. San Diego, CA: Academic Press.

Lorenz, E. N. (1993). *The essence of chaos*. Seattle, WA: University of Washington Press. doi:10.4324/9780203214589

MacWhinney, B. (2006). Emergentism – Use often and with care. *Applied Linguistics*, 27(4), 729–740. doi:10.1093/applin/aml035

Marek, M. W., & Wu, W.-C. V. (2014). Environmental factors affecting computer assisted language learning success: A complex dynamic systems conceptual model. *Computer Assisted Language Learning*, 27(6), 1–19. doi:10.1080/09588221.2013.776969

Östman, J.-O., & Fried, M. (2004). *Historical and intellectual background of construction grammar. Construction Grammar in a cross-language perspective* (pp. 1–10). Amsterdam, Netherlands: John Benjamins. doi:10.1075/cal.2.010st

Schulze, M. (2008). Modeling SLA processes using NLP. In C. Chapelle, Y.-R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic assessment* (pp. 149–116). Ames, IA: Iowa State University.

Schulze, M., & Penner, N. (2008). Construction grammar in ICALL. *Computer Assisted Language Learning*, 21(5), 427–440. doi:10.1080/09588220802447727

Selinker, L. (1974). Interlanguage. In J. C. Richards (Ed.), *Error analysis: Perspectives on second language acquisition* (pp. 31–54). London, United Kingdom: Longman.

Selinker, L. (1992). *Rediscovering interlanguage*. London, United Kingdom: Longman.

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly*, 39(1), 33–58. doi:10.2307/3588451

Sockett, G. (2013). Understanding the online informal learning of English as a complex dynamic system: An emic approach. *ReCALL*, 25(1), 48–62. doi:10.1017/S095834401200033X

Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: A dynamic systems view of the informal learning of English in virtual online communities. *ReCALL*, 24(2), 138–151. doi:10.1017/S0958344012000031

Sorites paradox. (2014). *Wikipedia. The Free Encyclopedia*. Retrieved from <http://en.wikipedia.org>

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M., Kinnear, P., & Steinman, L. (2011). *Sociocultural theory in second language education: An introduction through narratives*. Bristol, United Kingdom: Multilingual Matters.

Thorne, S. L., Fischer, I., & Lu, X. (2012). The semiotic ecology and linguistic complexity of an online game world. *ReCALL*, 24(3), 279–301. doi:10.1017/S0958344012000158

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2007). Konstruktionsgrammatik und früher Erstsprachenerwerb. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 19–37). Tübingen, Deutschland: Stauffenberg Verlag.

Tuckerman, M. E. (2008). Reaction coordinates. Retrieved from <https://www.nyu.edu/classes/tuckerman/stat.mechII/lectures/lecture_9/node5.html>

van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. Hertfordshire, United Kingdom: Harvester Wheatsheaf.

van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behaviour & Development*, 25, 340–374. doi:10.1016/S0163-6383(02)00140-6

van Geert, P., & Steenbeck, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, 25(3–4), 408–442. doi:10.1016/j.dr.2005.10.003

Verspoor, M., de Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development: Methods and techniques*. Amsterdam, Netherlands: John Benjamins. doi:10.1075/lllt.29

Wolf, C. (1979). *Fortgesetzter Versuch. Aufsätze, Gespräche, Essays [Continued attempts. Articles, dialogues, essays]*. Leipzig, Deutschland: Reclam-Verlag.

Zheng, D. (2012). Caring in the dynamics of design and languaging: Exploring second language learning in 3D virtual spaces. *Language Sciences*, 34(5), 543–558. doi:10.1016/j.langsci.2012.03.010

Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal*, 93, 489–511. doi:10.1111/j.1540-4781.2009.00927.x

CHAPTER 5

CALL design and research

Taking a micro and macro view

Mike Levy and Catherine Caws
The University of Queensland, Australia / University of Victoria, Canada

This chapter aims to explore two areas of computer assisted language learning (CALL) work that have proved problematic over time. The first area relates to our understanding of the broader contextual factors that influence CALL activity; the second relates to our understanding of the nature of interactions when those interactions are mediated via technology in some way. Thus, we aim to consider external factors and their influence on CALL and internal factors as they pertain to mediated interactions in CALL contexts. In both cases, we argue that insights and techniques drawn from the fields of HCI and engineering can enrich our understandings and practices, especially in focusing areas of research and development more effectively, and in conceptualizing research and practice in the first place.

Keywords: normalization, CALL design, CALL activity, micro/macro view, sustainability

Introduction

In order to unite CALL theoretical frameworks (preceding chapters in this volume) with models of CALL research practices (following chapters in this volume), this chapter aims to explore two specific areas of CALL work that have proved problematic over time. The first area relates to our understanding of the broader contextual factors that influence CALL activity (as partially illustrated in Chapters 2–4), and the second relates to our understanding of the nature of interactions when those interactions are mediated via technology in some way (see Chapters 6–10). Thus, we aim to consider external factors and their influence on CALL, and internal factors as they pertain to mediated interactions in CALL contexts. In both cases, we argue that insights and techniques drawn from the fields of human-computer interactions (HCI) and engineering can enrich our

DOI 10.1075/lsse.2.05lev

understandings and practices, especially in focusing areas of research and development more effectively, and in conceptualizing research and practice in the first place. We define our two specific areas as the *macro* view (namely, the holistic view of CALL activity) and the *micro* view (namely, the more focused exploration of technology-mediated interactions).

After highlighting the main features of both micro and macro views and relating these features to prior research, we go into a deeper exploration of the ways in which these two specific views on CALL research and practice can help us to better frame, articulate and design contexts of learning mediated by a constantly changing and evolving technology.

Key contextual perspectives

The macro view

Such reflections lead to our first area of exploration, which we will refer to as the *macro view*. This perspective considers a whole suite of factors that are *external* to the particular student interaction (human-computer or human-to-human-via-computer). Such factors include, but are not limited to, availability and access to technology and specific training (in using the technology), the curriculum, levels of technical competence (teacher and students), the technological infrastructure of the school or university, the level of technical support, school policy and so on. To assist in this analysis, we will revert at times to relevant theory, especially that used to inform and guide a broader context of use (e.g., ecological CALL, activity theory, dynamic systems theory, as seen in previous chapters of this volume). In addition to these theories, we will consider a number of general terms that have been referred to in the CALL literature when thinking about the broader context, notably, *systems*, *integration*, and, relatedly, *normalization*. Technological innovation itself also plays an important role in our critical examination of CALL research design because it constitutes the very foundation of our domain.

Systems

Viewing a particular learning environment as a system emphasizes the relations between the parts and the whole. Levy (1997) took this view further and said that "elements of the system can *only* be conceptualized meaningfully if they are viewed as part of the whole" (p. 66). Consequently, any element that is not working properly within the system will affect the whole system, and, furthermore, any

external influences impacting the system will influence each element within it, to a greater or lesser degree. From this perspective, it becomes important, and even critical, to identify the key elements in the system and their effects upon it. As we will discover herewith, not all elements will be equally influential or important, a fact that makes system design even more reliant on proper conceptualization and engineering.

Various theoretical standpoints are consistent with this way of thinking. For example, activity theory proposes the activity *system* as the basic unit of analysis, where the activity system comprises a dynamic network of interacting and interdependent elements with its own cultural history (see Chapters 2–3, this volume). Other approaches, such as Design-Based Research (DBR), also point in this direction, such as Barab and Squire (2004) who argued for the need to "consider the larger systemic constraints in which the context of intervention is a part" (p. 12). Dynamic Systems Theory (DST) provides an equally solid grounding for the analysis of non-linear systems (see Chapter 4, this volume). Ecological perspectives on CALL also point in the same direction, the sense of an evolving whole, rather than a focus on any one particular component.

Integration

Integration has been a topic of discussion in CALL from its earliest years. For example, Robinson (1991) reported on the conclusions of two research studies that highlighted "the importance of integrating individual CALL work with the total program of language instruction, including the classroom, rather than configuring it as an independent, supplementary activity" (p. 160). Hardisty and Windeatt (1989) emphasized pre-computer and post-computer work as well as work at the computer. They valued the importance of integration not only at the lesson level but also at the curriculum level. Hillier (1990) concluded that student training, teacher training and class scheduling were the most important elements for integrating computer work into their program. *Flipped learning* or *blended learning* is indicative of other approaches and potential solutions to the fundamental question of successfully integrating in-class and out-of-class work, where the overall goal is to maximise time on task through work both with a teacher and without one. Ultimately, within a *normalized* state of CALL informed practice (Bax, 2003), we would infer that integration could become a seamless process, comparable to one that is required for any mechanical system to operate effectively. The question of integration really relates to the ways in which the various elements influencing the use of new technology in language learning are brought together and managed in order to create a successful CALL environment. We need to understand

more clearly what (from physical artefacts to human intervention) is involved in successful integration and how we might break down the concept into a number of practical ideas and strategies. One potential oversight to immediately acknowledge is that any technological innovation is not, should not be, simply an add-on. In the world of software engineering, add-ons constitute, by definition, *accessory* devices that are meant to increase a system capability. However, in order to function optimally, technological innovation needs to be infused and not merely added on. This point was reinforced by Postman in the early years of educational technology, who said that technological change is "neither additive nor subtractive" but "ecological" in that "one significant change generates total change" (as cited in Debski, 1997, p. 41).

The technology itself

The technology itself also exerts its influence, especially in the way it can perturb the system or interrupt the processes of technology integration through renewal and change. As a culture, we are susceptible to the lure of the latest technology, and our expectations of what might be achieved are often at odds with the realities. Such reactions to new technologies have been captured in Gartner's *Hype Cycle* model <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp> which articulates five distinct categories or stages that occur in the emergence of any new technology, namely: *technology trigger*, *peak of inflated expectations*, *trough of disillusionment*, *slope of enlightenment*, and *plateau of productivity*. This trajectory provides a sense of how unrealistic initial expectations can quickly lead to disappointment, and the realisation that through extended use and systematic evaluation over time, a more reasoned assessment of the technology may be found. We argue that such features apply just as much in the world of education as to the world at large (see also Buckingham, 2007; Lanham, 2006; Levy, 2007). In particular, recent studies in CALL have shown a disconnect between what teachers perceive as positive learning contexts and what students do, or between technologies that are at the forefront of CALL research studies and technologies that are used regularly outside of class (see Huang, 2013; Steel & Levy, 2013). Such a gap between the perceptions and realities of use and what is actually learnt in the process merits further studies as well as innovative research methods, in particular, those that focus on an iterative process and the recycling of results into the design of new contexts of learning (Caws & Hamel, 2013).

Normalization

Focusing upon the whole CALL environment through the idea of a system, or of an ecologically, potentially sustainable manifestation, through the importance of integrating the parts and through the role that technology plays leads us to an appreciation of the importance of a stable learning environment. In this regard, the concept of normalization is helpful. In 2003, Bax argued that we should aim for a state of normalization, and said, "This concept is relevant to any kind of technological innovation and refers to the stage when the technology becomes invisible, embedded in everyday practice and hence 'normalised'" (p. 23).

Bax (2003) continued with the bold statement that designing for a state of normalization could "structure our entire agenda for the future of CALL" (p. 24) (see also Bax, 2011; Chambers & Bax, 2006; Lafford, 2009) while also warning us that "normalisation of a technology can arguably at times have negative consequences" (Bax, 2011, p. 1). For example, it is not advisable to unilaterally adopt a new technology too quickly, before it has been rigorously evaluated. It is precisely within such circumstances that expensive resources are under-utilized. New innovations need to be subject to the acceptance of teachers and students, with a clear understanding and appreciation of their value. Yet normalization can be a suitable "end-goal" for CALL (Bax, 2003, p. 24). Bax argued for a kind of reverse engineering whereby, through research, we identify the factors that need to be accounted for in order to facilitate or lead to the normalized state. Thus an agenda for research and practice is articulated. Levy and Stockwell (2006) concluded, with caveats, that for language teachers and learners, "Normalization becomes a process of understanding the infrastructure, the support networks, and the materials and working effectively within them" (p. 234).

In consideration of these external factors (namely, *systems*, *integration*, *technologies*, and *normalization*) and with regard to the *macro view*, we will reflect on ways in which concepts and techniques from HCI and engineering can help us with the analysis of the learners' experience, the analysis of technologies and, ultimately, the overall CALL design. In particular, we will consider more intensively how ideas from *systems theory* (emphasizing sustainable systems that are constantly "corrected" through feedback), the notions of *reverse engineering* (highlighting a process of dissembling or reversing potential malfunction of a design, system or technology) and the *life-cycle* (paying particular attention to evaluation, and feedback in a view to *re*-design) might assist our understandings of how elements within a system might contribute to the workings of the system as a whole (see Levy, 1997, pp. 215–218).

The micro view

The second area we explore relates to the particular features of technology-mediated language learning and the resultant nature of interactions that occur within this setting. We refer to this area as the *micro view*. It is argued here that interactions in technology-mediated settings are fundamentally different from those in non-mediated settings, such as face-to-face (FtF). Moreover, we argue that even if a technology has reached a state of normalization as defined by Bax (2003), its use by learners will often show variations in efficacy and efficiency of use (see also Chapter 2, this volume). Such disparities in learners' interactions with technologies can be captured by way of microanalyses (see Part II, this volume).

Our area of focus here is well illustrated by Smith (2008) in his article "Methodological Hurdles in Capturing CMC Data: The Case of the Missing Self-Repair." Smith (2008) examined computer-mediated communication (CMC) between pairs of students learning German. The study is significant in Smith's alertness to the particularities of the context of interaction. He was one of the first to use video-capture data-collection techniques to record the language learners created in their private chat box before a message was sent and incorporated into a chat log (see also O'Rourke, 2008, 2012; Smith & Gorsuch, 2004). The CALL context in this example is uniquely different from the FtF context. The FtF context does not require or impose a two-step process as far as output is concerned: It is the interface design itself that imposes this constraint upon the user in text chat. This makes the communication process in the CALL context different, and the researcher has to be very cautious in assuming findings in one context and applying the findings unequivocally/directly to the other. Conversation analysis (CA) techniques can be very helpful in elaborating differences (see also Hutchby, 2001; Hutchby & Barnett, 2005; Hutchby & Tanna, 2008). The Smith study also illustrates the immense value of using a data-collection device or instrument that is capable of capturing a richer, more in-depth and complete picture of text production and the changes that typically occur as a text is edited and finalised privately, prior to the user posting the message in the public space.

Smith's (2008) study is relevant here because of its methods of data collection and analysis (see also Chapter 8, this volume). Instead of only using a chat log file as a data source – often unquestioned as the traditional approach – this study captured a video file of the whole interaction on screen. While the chat log file collected some of the examples of repair, it did not capture the *history* of the construction or all the instances of repair. As Smith (2008) said, "many CALL studies do not make use of existing technology in their data collection and analysis methods, which can severely limit the impact and relevance of their findings" (p. 85).

Generally speaking, in endeavouring to capture such data, we are trying to capture and understand *what students do*. Such a perspective can be traced back at the very least to the seminal volume by Winograd and Flores, published in 1986, *Understanding Computers and Cognition*, where they dedicated a whole chapter to this topic. As they said, "'Doing' is an interpretation within a background and a set of concerns" (Winograd & Flores, 1986, p. 143). We aim to develop this perspective further by arguing that several techniques may be drawn also from HCI. This connection between what learners actually do, as opposed to what we believe that they are doing, has also been the focus of previous CALL research, notably those that are informed by ergonomic approaches (see Chapter 2, this volume).

Thus, with the micro view, in considering how concepts and techniques from HCI and engineering can assist, we refer to relevant material in the domains of, (a) the users' (i.e., learners) experience (UX) as it applies to the users' behaviours, attitudes or emotions as they interact with a particular technology; (b) the user interface (UI) as it applies to the characteristics of the technology that facilitate the interactions (i.e., usability) between human beings or between a human being and a technology, and as it relates to users' needs (that typically have been identified at an earlier stage); and (c) human-computer interaction (HCI also coined LCI in this volume to emphasize the fact that the user is a learner) as it applies to the actual interactions between learners and technologies as exposed through research strategies and techniques that can help elaborate the nature of such interactions, such as user-walkthroughs or talk-aloud protocols as described by Hémard (2006). Ultimately, then, our focus is on designing and evaluating CALL contexts that are learner-centred.

Our chapter then continues with a deeper, more elaborated discussion of the issues. Section 2 will consider the macro level and the importance of considering CALL at the level of the system. In simple terms, such a view carefully considers a whole suite of factors that are external to the particular human-machine or human-to-human interaction. Ideas of integration and normalization will be discussed further with their implications. In thinking about how concepts and techniques from HCI and engineering can help, we consider more intensively ideas from systems theory, and the notions of reverse engineering, re-engineering, and, more generally, the concept of a life cycle. We will see that such notions are also inherently related to action research if we seek to delve into a deeper understanding of all the social, cultural and ethnographic factors that may affect the process of learning in a technology-mediated context. From prototyping to full implementation of learning systems, one can see the value of such iterative mechanisms of analysis and development.

In Section 3, we consider the micro level and the nature of mediated interactions. Through a number of examples, the current state of play will be described.

The various examples will also relate to Part II of this volume, which details several mechanisms for capturing and analysing learner-computer interactions. This will be followed again by a discussion of how techniques and strategies from HCI and engineering may further our understandings and practises. In considering how concepts and techniques from HCI and engineering can assist, we refer more specifically to the effects of learner interface design and learner experience, and we explore research strategies and techniques that can better inform us on the learning processes and practices, and ultimately on the level of normalization attained by a technology in its context.

Discussion: The macro level

Our discussion of the macro view begins with a more detailed analysis of the idea of normalization and its ramifications. This is a useful path forward because of the issues that were brought to light as a result of the discussion. The concept of normalization is predicated on the achievement of a relatively stable system. Consequently, this line of thinking potentially helps to expose those elements that tend to interfere or disrupt that ambition.

There are arguably a wide range of factors that militate against stability in contemporary CALL. These factors are of several kinds: social, cultural, economical, systemic, structural and even spatial. (For instance, access to technology will vary greatly within one country, depending upon the access and availability to wireless Internet.) Combined with these is the fact that new technologies appear at an alarming, increasing, rate and that many – though certainly not all – are quickly absorbed into everyday life (at least in most western nations). One only has to think of the latest smart phone. The wide-spread adoption of mobile phones in the wider world by young people, or the use of games with highly sophisticated graphics, leads to changing expectations when it comes to the technologies and software applications used in schools. Expectations are raised to a higher level. In stark contrast, educational institutions tend to have limited resources and are unable to match this rate of change. The result within the school environment may be a blanket ban on mobile phones, for example. Yet, language learners are independently using the powerful personal technologies (not necessarily for learning or studying) they now have at their disposal. In sum, numerous external factors impinge upon the teacher, her students and the nature of the classroom or learning environment.

Latterly, Bax (2011) has somewhat revised his concept of normalization in language education using a neo-Vygotskian perspective in order to take into account the multiple factors that may affect the interactions with the technologies.

In effect, by re-questioning the concept of normalization and realising that many variables will limit or slow down the process of reaching full effectiveness of a technology, we admit the complexity of the systems and the need for re-engineering the elements that comprise them. Thus, in the following section, we will take a closer look at the *macro* factors that are critical in our consideration of a state of normalization, the research agenda that derives from this process and the ways in which engineering and HCI can inform our research practices.

Critical factors

Recognizing the fact that "true integration of CALL within language learning and teaching" had yet to be achieved, Bax (2003) proposed a list of factors that critically affected our progress towards normalization (p. 11). We have already discussed some key issues, such as people's attitudes (teachers and administrators), or system issues (timetabling or access to technology) that led to his first list in 2003. Building on this list, Levy and Stockwell (2006) suggested a tentative start list of critical factors when normalization is the goal (p. 233):

1. Easy access to the appropriate technologies (hardware/software), when required
2. Acceptance by administrators that language learning has particular hardware/software needs
3. Reliable technologies and applications
    a. Technical support when needed
    b. CALL materials that are robust and easy to use
4. Reliable and willing partners in collaborative projects
5. Acceptance of CALL activity by staff and students *as normal practice*
    a. CALL materials that are relevant to the goals and needs of the students
    b. Training for staff and students

The ways in which these critical factors affect each other can be visualised in Figure 5.1 below.

Though we may be able to make informed guesses at what factors are likely to be more generally applicable, the relative influence or impact of individual factors will inevitably vary from place to place. Local issues will always play a very important part. Thus, the particular weighting of factors and the order of importance in any particular setting are likely to vary and be highly context specific.

Unfortunately, many of the factors involved are likely to lie well beyond the control or the direct influence of the individual language teacher. As an example, decisions concerning the location and distribution of computers within an

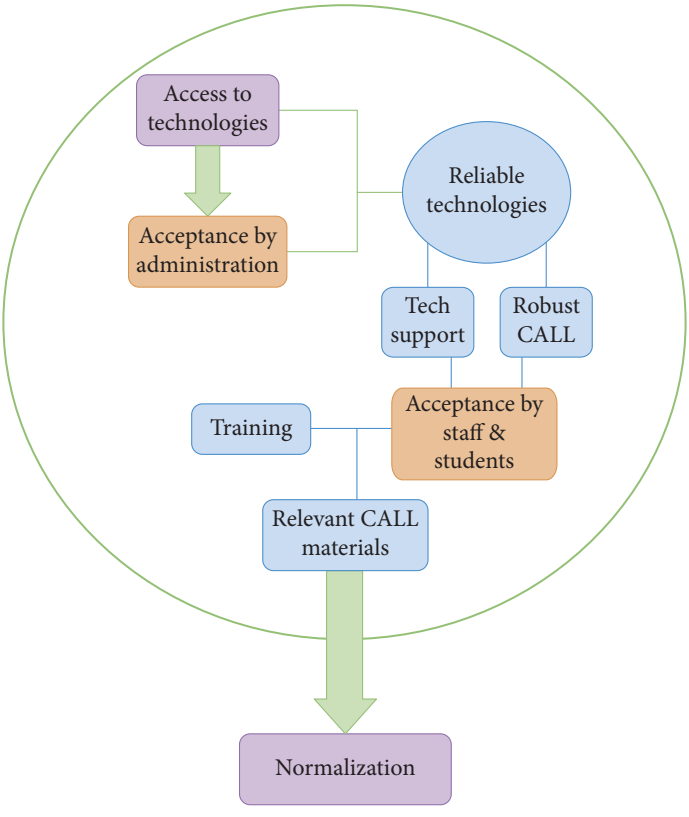


Figure 5.1 Critical factors towards normalization

institution that are highly likely to impact upon normalization, for example, are not usually made by language teachers. Yet questions of access are often a major concern. Further, without appropriate training, neither staff nor students can hope to incorporate CALL as normalized practice (see Hubbard, 2004).

CALL is also context specific. In any particular situation, certain factors will present themselves as pivotal concerns while others will be of less immediate relevance or importance. So in one setting, the question of access might be crucial; in another, a fixed and non-negotiable curriculum might be a major barrier to innovation. In a different setting, teacher training and attitudes might be central. Language teachers are very much working within a complex system of opportunity and constraint. Normalization, then, becomes a process of understanding the infrastructure, the support networks and the materials, and working effectively within them.

Chambers and Bax (2006) have furthered this line of inquiry in their article "Towards Normalisation." They discussed a wide range of obstacles to normalization besides the technology and the software, including, teacher training, administrative and pedagogical support, syllabus and curriculum integration, teacher attitudes, school culture, physical setting and location of computers, funding, leadership, accountability structures and so forth (see also Fishman et al., 2004; Levy, 1997). It is well worth noting that Chambers and Bax (2006) identified "syllabus integration" as the one overriding factor (p. 477), whereas Fishman et al. (2004) identified time constraints as a direct result of the impact of "standardized assessment" (p. 60).

To date, we believe Chambers and Bax (2006) have come closest to unravelling the complexities of normalization and context (p. 470). They endeavoured to divide normalization into some of its potential *constituents* described in this article as *issues*. For normalization to occur, Chambers and Bax isolated eleven particular issues, divided into four groups under the headings: (a) Logistics; (b) Stakeholders conceptions: Knowledge and abilities; (c) Syllabus and software integration; and (d) Training, development and support. By way of example, consider the first category given by Chambers and Bax (p. 470), logistics:

1. For normalization to take place, CALL facilities will ideally not be separated from "normal" teaching space.
2. For normalization to occur, the classroom will, ideally, be organized so as to allow an easy move from CALL activities to non-CALL activities.
3. For teachers to "normalize" computer use within their daily practice, they may need additional time for preparation and planning.

There have been few follow-up papers to Chambers and Bax (2006) on the same theme, as far as we are aware, although one such is described by Kennedy and Levy (2009), who gave examples of sustained activity in CALL over time. They said,

> For as long as we have been engaged in CALL projects, the characteristics of the institution's support for CALL have met the relevant criteria recommended by Chambers and Bax (2006, p. 477–478) as necessary for the normalization of CALL (issues 1, 2 and 10). First, we have "CALL facilities not separated from normal teaching space". … Second, the layout of the two main CALL-equipped classrooms is "organized so as to allow for an easy move from CALL activities to non-CALL activities"… Third, we have "provision of reliable technological support and encouragement." (p. 455)

The paper by Chambers and Bax has been one of the few to seriously consider the broader factors that are required for normalization. They discussed these issues in a very practical, teaching-oriented way. However, there is a broader point to be

made that is complementary. This perspective, once again, focuses on the broader picture through systemic approaches to research.

A research dimension informed by HCI and engineering

In thinking about *research* in the context of LCI, Salomon (1991) has advocated systemic approaches as most effective. These approaches for research very much align with our focus on practice so far, and the systems approach is highly compatible with recognized practices in HCI and engineering. Salomon's (1991) study of "complex learning environments undergoing change" began with the assumption that "elements are interdependent, inseparable, and even define each other in a transactional manner so that a change in one changes everything else and this requires the study of patterns, not single variables" (p. 10). This is further reinforced when Salomon (1991) said that with systemic approaches, the research is dealing with a "whole dynamic ecology" (p. 12), the "newly created classroom culture" (p. 13) and "authenticity" (p. 16). This fits very well with what we said earlier about normalization.

However, it would be an unfair representation of Salomon's paper, if the reader were left with the impression that systemic approaches were all that were needed. Salomon (1991) also said,

> The systemic study of complex learning environments cannot be fruitful, and certainly cannot yield any generalisable (applicable) findings and conclusions, in the absence of carefully controlled analytic studies of selected aspects in which internal validity is maximised. (p. 16)

Salomon (1991) further stated the following:

> For one needs to know what aspects of the complex setting deserve to be studied in greater detail under controlled conditions. The sources of such knowledge are one's detailed and systematic observations of the complex phenomenon. Without observations of the whole system of interrelated events, hypotheses to be tested could easily pertain to the educationally least significant and pertinent aspects, a not too infrequent occurrence. (p. 17)

Of course, the idea of the system permeates systemic approaches to research. In any particular teaching context, the teacher first needs to identify the elements of the system, as far as CALL is concerned. The system, then, has to be examined from within, considering how the various elements interact with one another, and then from the outside, considering what factors or elements are likely to impact the system or disturb its equilibrium. Systems are dynamic and are subject to change, so formal and informal evaluative studies will need to be ongoing.

More specifically, observations and techniques drawn from HCI and engineering offer many lessons and suggestions that can help us refocus our attention on CALL activity and research in a more holistic manner. We will start our argument by referring to excerpts from Norman's (2013) influential book *The Design of Everyday Things*. This book is fundamental in understanding the way in which *design* affects almost every aspect of our daily lives. A key point is that design is *purposeful*. It involves *planning ahead* and *anticipating* actions and responses in the myriad contexts of use. Good design is also effective in facilitating, supporting and optimising the completion of the tasks or functions that the technology has been designed and built to serve. It does not matter if the technology is simple or complex (from a door to a spacecraft); these basic design qualities still apply. In particular, Norman bases the principles of design on psychology, cognition, action, or interaction, which are also inherently critical aspects of learning.

The core of our argument is, essentially, that we need to carefully craft the *design* of technology-mediated language learning contexts because design will have a direct impact on normalization. Such a perspective is consistent with Levy's (2002) view of the role of the language teacher as designer, as he has explained:

> Viewing the language teacher as a designer brings to the foreground some critical insights. The first and most important of these is that the language teacher in creating a product or plan of action operates within a set of interrelated constraints. Constraints, often associated with the limited time and resources available to the teacher and the student, typically include: the number of contact hours pre-determined for a course; lesson times and durations; preparation time; access to new technologies and to software; development budget; technical support; ancillary learning materials and so on. (p. 77)

The influencing factors mentioned in this extract also overlap with the factors mentioned later (see Chambers & Bax, 2006, and discussion later in this paper). Norman (2013) also made some important observations on the essential qualities of good design. His introductory statement is quite revealing for our argument here. Norman (2013) stated that "Good design is actually a lot harder to notice than poor design, in part because good designs fit our needs so well that the design is invisible, serving us without drawing attention to itself" (Preface, para. 2). Likewise, Bax (2003, 2011) referred to normalization as a stage where the technology is so infused in the learning context that it has become *invisible*. Chambers and Bax (2006) added, "In this light, our aim as CALL practitioners is to achieve such a seamless linkage between the computer and our teaching that the computer becomes as unremarkable in our daily practice as the pen and book" (p. 466). When a design works well, in language teaching/learning as in other disciplines or areas, we are not drawn to the particulars of the detail of the design. Rather,

we are provided with a working environment that is highly compatible with our goals and intentions, and the task at hand (in our case language learning and teaching), which, in turn, provides a setting where we can simply get on with the job smoothly, without complication and with maximum effect (i.e., learning).

Overall, *design* is a crucial element of engineering and HCI research, and we will see that several elements of these disciplines transfer directly into our examination of CALL research both at the macro and micro levels. We will focus here on the macro level and revisit the four elements that we had isolated earlier (see Section 1).

Systems

In thinking about parallels between learning environments and systems, our practices can be enriched further by a range of techniques that are common to systems engineering, an interdisciplinary discipline of engineering that is based on the "fundamental idea that a system is a purposeful whole that consists of interacting parts" (INCOSE, 2015, p. 5). Like systems engineers, we (language learning engineers) need to address issues of reliability (of the various systems' elements), logistics (as seen above in Chambers & Bax, 2006), and coordination of "teams" (all the actors that play an integral part of the activity systems, as seen in activity theory, Chapter 4, this volume). The idea of the life cycle is also valuable when contemplating new technology implementations and their lifespan (see Levy, 1997, p. 216). In systems engineering, there is a requirement that all the identifiable aspects of a system are taken into consideration and constitute a whole. As we have seen earlier, CALL research and practices need to adopt a similar perspective in order to come closer to a state of normalization. Such an ecological perspective on CALL leads us to our second macro element, the *integration* of CALL into the language learning context.

Integration

We commented earlier on CALL integration as a focus of many research studies. CALL integration is often contingent on external factors, such as the physical settings of the teaching/learning space or the timetabling of courses. Likewise, when we referred to *logistics* as per Chambers and Bax (2006), we notice that the factors that the authors identified as being in favour of normalization (such as CALL facilities being close to teaching space or classrooms allowing an easy move from CALL activities to non-CALL activities) all have to do with *design* of the learning environment (spatial or virtual).

Integration plays a role at many levels: classroom (spatial), lessons and curriculum, program or people training. In all cases, an optimum integration of these elements relies on some aspect or form of design. When integration fails, relying on troubleshooting (namely, a form of re-engineering) is a fitting solution. Troubleshooting is a particularly effective method to look for causes of processes that have failed, and it is commonly applied in the maintenance of complex systems (see Chapter 2, this volume). As discussed earlier, successful integration of technologies in language learning contexts rests upon a delicate balance of many complex and diverse elements. In that regard, integration is also related to the notion of complex systems, derived from the field of complexity theory (see Cameron & Larsen-Freeman, 2008; also Chapter 2, this volume), whereby aspects such as change and heterogeneity constitute central elements. Conversely, successful integration of new technologies to learning environments also relies on users' capability to adapt to change and alter their behaviours towards what constitutes learning. Within this complex system, the technology plays a crucial role.

The technology

As explained earlier, the technology itself can seem disturbing, luring, or, ideally, neutral if already fully and seamlessly embedded in interaction practices. Norman (2013) made an interesting distinction between the affordances (see also Chapter 3, this volume) of the instrument (namely, the actions that the instrument permits) and the signifiers (namely, the signs discovered by users of what can be done with the instruments) (Chapter 1, para. 1). He goes on by explaining that in the case of complex devices, a user will often need some form of instruction in order to better manipulate the device (see also Hubbard, 2004).

The design of the technology (as well as its integration within a complex system) will highly influence the success or failure of a particular instrument. It is often the case that when a technology "fails," we have expected too much of it, such as in the case of automated translation. We have failed to be *humble* about the power of the machine.

In many cases, we also fail to appreciate the human and social factors that influence the success of an activity or an interaction with a technology (see Norman, 2013). In other instances, the technology has been developed in house and too little focus went into its design. In this particular case, as well as in the case of popular technologies in the private sphere that are introduced in the public educational sphere (such as Facebook), principles of HCI can greatly help us assess the situation. In particular, HCI is based on some of these principles:

- Positive interactions with a technology is the direct result of "good" design
- All design is *re-design*, based on observations of users and analysis of their activities

These basic principles invite us to pay close attention to our users/learners in order to design technologies that support their needs, that accommodate their learning styles and that are compatible with their environment. As seen in Chapter 2, this volume, an HCI-inspired ergonomic user-centred approach to evaluating technologies has already been used in CALL research. For instance, focusing on the potential of hypermedia, Hémard (2006) rightly noted that in spite of the "perceived potential" of hypertext, many systems provided poor navigational architecture or interactivity, resulting in learning environments that learners did not feel motivated to use (p. 25). He added, "At the root of this problem lies the fact that the designers' model of how their electronic environment ought to behave is not matched by the learners' mental model of interactivity on the web and how this can help them achieve their learning goals" (Hémard, 2006, p. 25). Hémard referred here to Norman's design framework. Likewise, Ward (2006) recommended the application of sound software design principles to CALL design. She explained, "Software is not designed and built for software engineers alone (nor should it be) – it is an outward looking process that should be driven by user needs. Software design principles are based on the fact that the software will be built to cater for user demands in a myriad of different contexts" (Ward, 2006, p. 131). Overall, we can see that applying design principles reaches beyond the design of the software or app only, that is in isolation. The design must also do its best to represent the many contexts of use.

Normalization

At various points in this chapter, while talking about the macro view, we have discussed the importance of breaking down the whole into parts, and identifying key influencing factors, and then considering how those parts contribute to the workings of the whole, as in the design of the whole learning environment. In a sense our point of focus moves from the whole to the parts and back again. Bax (2003, 2011) argued for a kind of reverse engineering whereby, through research, we identify the factors that need to be accounted for in order to facilitate or lead to the normalized state. Reverse engineering is a process that typically applies to a product; however, considering the many elements that need to be assembled in order for effective CALL to occur and for normalization to be achieved, we argue that by trying to disassemble the CALL learning context into identifiable chunks, we can better analyse the design features that need improvements. Likewise,

evaluation methods commonly used in HCI can be directly applied to the process of re-designing CALL, with a goal towards normalization. Specific examples of data collection and analysis will be used when focusing on the micro level. However, what is particularly striking in terms of analogy is that good HCI depends upon a careful investigation of users' needs and goals in order to design interactions that are enjoyable and connected within a whole systems.

Discussion: The micro level

The nature of mediated interactions

In understanding mediated interactions, it is essential to capture the detail, and to avoid falling into the trap of assuming that research findings deriving from co-located, face-to-face interactions can be transferred straightforwardly and simply to mediated learning contexts, such as synchronous computer-mediated communication (SCMC). As noted by Levy (2000), "For the CALL researcher, technology always makes a difference; the technology is never transparent or inconsequential" (p. 190) (see also Levy, 2006).

The Smith (2008) study is significant in his alertness to the particularities of context at the micro level. In practice, mediated interactions are often likened to face-to-face interactions with (we will argue) insufficient evidence regarding the grounds upon which they may correctly be regarded as similar or comparable. The CALL context in this example is uniquely different from the FtF context. The FtF context does not require or impose a two-step process as far as output is concerned; it is the interface design in the technology-mediated context that imposes this constraint upon the user in text chat. This makes the communication process in the CALL context different, and the researcher has to be very cautious in assuming findings in one context and apply unequivocally/directly to the other. This study also illustrates the immense value of using a data-collection device or instrument that is capable of capturing a richer, more in-depth and complete picture of text production and the consequent interaction as it moves from the private to the public space (see Chapters 7–9 for more examples of data-collection devices and micro-analysis of processes).

Working along similar lines, O'Rourke (2008, 2012) has provided two complementary articles that focus chiefly on method. These articles describe approaches that advocate important roles for qualitative research, although not necessarily on its own. The use of eye-tracking is one feature among other devices for capturing the specific features of an interaction (e.g., Tono, 2011). O'Rourke (2012) argued that "eye-tracking can bring us closer to the first-person experience of SCMC"

(p. 306). He argued that "the practical mechanics of co-construction work very differently in SCMC than in speech" and added that "these differences are attributed ultimately to the temporal relationship between linguistic production and perception and to the differing nature of the 'communicative space' in the two modes" (O'Rourke, 2012, p. 306). Further, in another related article, O'Rourke (2008) said, "Smith & Gorsuch (2004) demonstrate convincingly, based on audio and video recordings, that facial expression, body posture, audible self-speech, and general direction of gaze can provide important information on difficulties in SCMS production that is either completely unavailable from output logs, or can only be weakly inferred" (p. 234). In fact, in research to date, O'Rourke (2008) argued that many aspects of the learning environment relevant to learning in SCMC had been neglected, namely:

1. users' paralinguistic and non-linguistic behaviours – gestures, spoken utterances, posture, etc.;
2. interactional tempo, both globally (whether a session is generally characterized by rapid or more leisurely exchanges) and locally (response latency, i.e., the length of gaps between particular turns);
3. drafting processes – i.e., editing of input prior to sending; and
4. attentional focus – i.e., what users are actually attending to at a given moment. (p. 233)

It is well worth contemplating the possibilities for CALL research across these four areas. To date, some data-collection devices focus more on detailing the actions of the individual human user (e.g., eye-tracking) and some more on the technology in its response (e.g., screen capture) – the overall objective, of course, would be to capture as full a record as is possible of the whole interactive process from all sides in real time. For example, one might consider a distance SCMC collaboration, in tandem learning for example, with two students interacting at a distance. It would be interesting to capture data from both students, strictly in sequence in real time, of what occurred, including what was constructed before a message was sent by each participant, simultaneously, and how exactly the messages collided and were responded to, with particular attention given to the precise order of events. When one reflects on the nature of the interaction online and at a distance, one can begin to uncover the immense differences between synchronous online interactions and FtF interactions when the participants are present in the same physical space. The trap is to oversimplify and, as O'Rourke (2008, p. 233) has said, to "neglect" differences that may turn out to be very significant. Further work might consider the possibility of two students working together at the computer and the language used between the students as well as that online (e.g., Levy & Gardner, 2012).

Otherwise, the differences O'Rourke suggests have been overlooked help point us in the direction of a CALL research agenda and new, innovative research studies that address these issues (see also Hamel, 2012, and Chapter 7, this volume).

Just to continue with O'Rourke's (2012) analysis for a moment, he explained why output logs are "impoverished" (p. 236), and why typically they entirely exclude the *private space* in which students construct their utterances during text chat. He concluded, "If we wish to understand the moment-by-moment reality of communicating in real time by text – a reality that affects cognitive, affective and social dimensions of behaviour – we need to 'zoom in' and examine the texture of interactions with SCMC systems as experienced by the individual" (O'Rourke, 2012, p. 247). Several studies in Part II of this volume will address this need by proposing some practical tools and methods to help us understand *how* learners interact with systems. Ultimately, this need can be partially answered through a careful application of principles derived from the fields of interaction design and experience design. Norman (2013) emphasizes this need when he claims:

> the focus [of interaction design] is upon how people interact with technology. The goal is to enhance people's understanding of what can be done, what is happening, and what has just occurred. Interaction design draws upon principles of psychology, design arts, and emotion to ensure a positive, enjoyable experience.
>
> (The complexity of modern devices, para. 3)

Mediated interactions: Temporality

At a micro level, particular forms of interactions, such as technology-mediated communications, are strongly influenced by external, uncontrollable, factors. As an example, it is helpful to think of technology-mediated communication on a time-scale, from slow (asynchronous) to fast (synchronous), then adding face-to-face communication with participants in the same physical space (FtF-SPS) as a comparison at the faster end (see Levy & Stockwell, 2006, pp. 97–99). In past times, traditional hard-copy letters – a form of technology-mediated communication – could take months to pass from the author to recipient. Such forms required a particular kind of forward thinking, resulting in particular forms of *interaction*, input, output, and *negotiation of meaning*. Authors had to look forward and backward and *predict* what would be relevant and timely when the letter was finally received. The time available, relatively, was not an issue. Compare this to, say, a Skype conversation, where interactions are *almost* simultaneous. But, interestingly, there is evidence to suggest that there remain important differences with FtF-SPS, especially when one begins to consider turn-taking behaviours, overlaps, interruptions, etc., perhaps also on the language forms themselves. As

Skehan argued in 1998, all things being equal, exerting greater time pressure on learners will mean that there is "less time for attention to form both in terms of accuracy or complexity" (as cited in Levy & Stockwell, 2006, p. 167). Adding on, Levy and Stockwell (2006) said,

> Time pressures themselves vary greatly from one form of CMC to another. Asynchronous CMC, of course, allows the learner far more time to think about a response, as well as providing sufficient time to consult resources such as dictionaries or grammar reference books, or even to seek assistance from other people. (p. 98)

The issue of time applies to many other technologies that are currently used within language learning contexts, and as such further research studies need to be developed in which the concept and impact of time is taken into account. The literature on planning, for example, shows that pre-task planning has the potential to significantly influence the language produced in the task that follows (see Skehan & Foster, 1997, 2001). Thus, the issue of time is a most significant one regarding different forms of mediated communication as well as different forms of interactions with CALL instruments.

Generally speaking, in endeavouring to capture interactional data, we are trying to capture and understand "what students do" (see Chapter 2, this volume). Such a perspective can be traced back at the very least to the seminal volume by Winograd and Flores published in 1986, *Understanding Computers and Cognition*, where they dedicated a whole chapter to this topic. As they said, "'Doing' is an interpretation within a background and a set of concerns" (Winograd & Flores, 1986, p. 143). Raby (2005) also made a good case for direct observations of learners while working on computer-mediated tasks, and the value of user-centred ergonomic approach (see Chapter 2, this volume).

Evaluating interactions according to HCI and software engineering methods

Just as we have seen at the macro level, at the micro level, methods and practices inherited from HCI and engineering have contributed very positively to the CALL research agenda. There are many factors that CALL researchers have pointed out and that are becoming more and more current in today's research practices and methods (see Part II, this volume). At a time when the development of hypermedia language learning applications was increasing at a fast pace, a handful of CALL researchers recognized an inherent need to apply strict software engineering and design principles. One such researcher was Hémard (1997), who remarked that, "little help in the form of design and technical support [was] being made available

to individual authors with little or no design expertise" (p. 9), hence the need to make better use of principles and guidelines from user-interface software engineering. One important factor that is inherited from this field is the collection of "empirical data drawn heuristically from experience in user-interface design" (Hémard, 1997, p. 10). We will see effective examples of such data collection in Part II of this volume.

The principle on which we base our argument is our view that learning environments as systems are highly dynamic, with many influential elements that need to be better observed, evaluated and eventually re-designed. In the field of HCI, for instance, an iterative process of prototyping, feedback, rapid testing, and evaluating through direct manipulations helps software engineers to distinguish good design from bad design. In a more generic way, our goal is to distinguish between the *expected* performance or behaviour of the instrument or the learner and the *effective* performance or behaviour as defined by the *observed* behaviour (Raby, 2005). Hémard (2004), commenting on the explosion of web technologies and the lack of proper critical evaluation of their interactive potential, stated that "the way forward involves adopting a more reflective and iterative approach towards existing online CALL design and practice supported by the systematic evaluation of the usability and effectiveness of its delivery" (p. 503). The main goal of proper evaluation is to test the usability of a technology, namely its potential to allow learners, as Karat said in 1997, to "achieve specified goals with effectiveness, efficiency and satisfaction" (as cited in Hémard, 2004, p. 503).

In an attempt to classify and assess various methods to enhance CALL design, development and research, Hémard (2004) referred to several quantitative and qualitative methods that are directly inspired by HCI. In regard to the micro approach, we note in particular the following methods proposed by Hémard (2004, pp. 505–506):

- Real life observation during a CALL activity to gather qualitative information on learners' interactions
- Check-list to collect input from peers or experts (to evaluate a system, for instance)
- User walkthrough using a CALL lab, equipped with data capture technology to collect real qualitative recording of users' interactions (see also Chapters 7–9, this volume)
- Focus groups to provide additional qualitative feedback from discussions
- Tracking using interaction logging tools to analyse users' behaviours
- Usability tests to collect valuable, direct and accurate feedback from users' performances

Part II of this volume will present a sample of these techniques in various learning contexts using learning software, or during computer-supported learning tasks. In all cases, we will see that by delving into precise learner-computer interactions, we tend to further illuminate what occurs in the "private" space, hence helping CALL researchers and designers make better predictions on common errors, successes, quality of input and output, and affordances of CALL activities and instruments.

Conclusion

With an aim to connect theoretical aspects of learner-computer interaction (LCI) as seen in the first part of this volume, to the practical applications of CALL research that Part II of this volume will present, this chapter has dealt with two critical aspects of CALL design, research and practices: the external factors that influence CALL and the internal factors that concern mediated interactions in CALL contexts.

According to Fishman et al. (2004), the primary reason research on technology innovations has had relatively low impact in everyday practice in K-12 schools is because it has not focused sufficiently on issues of how innovations function at the level of systems (p. 69). Likewise, authors such as Bax have argued that we need to investigate more fully the system barriers or constraints that impede the creation of a more normalized, integrated role for technologies in the language classroom.

Theoretical perspectives, such as activity theory, allow for boundary investigation and analysis, with a focus on the realities of sustained application. There is further overlap with recent work on ecological CALL (see Lafford, 2009; van Lier, 1998) and complex systems (see Larsen-Freeman & Cameron, 2008, p. 244) because a similar embedded stance is adopted.

As far as the concept of normalization is concerned, one can certainly understand why some stability may be appealing – particularly for administrators – because of the costs involved, in every sense. However, as we have discovered, a state of normalization (both at the micro and macro levels) requires the involvement of all parties involved in educational settings (from the user to the instructor and staff).

CALL researchers also need to question and carefully test any research findings set by theories emanating solely from the context of face-to-face learning interactions. A simple extrapolation from FtF to technology-mediated learning settings is often an over-simplification and reduces the importance of contextual factors at all levels, macro and micro. In other words, the idea that theories and

constructs derived from face-to-face research designs and settings can be simply applied without complication to mediated settings should be more seriously questioned. In our opinion, face-to-face interactional settings should be considered different to mediated ones until proven otherwise, not the reverse. Technology makes a difference.

Finally, more empirically-based studies in technology-mediated contexts are needed. Research here requires a highly perceptive response to the subtle differences that distinguish technology-mediated communicative exchanges with those where participants are both co-located and face-to-face simultaneously. Only then can we begin to understand the key differences between the two settings, and the particular role that a mediating technology might play.

References

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1–14. doi:10.1207/s15327809jls1301_1

Bax, S. (2003). CALL – past, present, future. *System*, 31, 13–28. doi:10.1016/S0346-251X(02)00071-4

Bax, S. (2011). Normalisation revisited: The effective use of technology in language education. *International Journal of Computer assisted Language Learning and Teaching*, 1(2), 1–15. doi:10.4018/ijcallt.2011040101

Buckingham, D. (2007). *Beyond technology: Children's learning in the age of digital culture*. Cambridge, United Kingdom: Polity Press.

Caws, C., & Hamel, M.-J. (2013). From analysis to training: Recycling interaction data into learning processes. *OLBI Working Papers*, 5, 25–36. doi:10.18192/olbiwp.v5io.1116

Chambers, A., & Bax, S. (2006). Making CALL work: Towards normalisation. *System*, 34, 465–479. doi:10.1016/j.system.2006.08.001

Debski, R. (1997). Support of creativity and collaboration in the language classroom: A new role for technology. In R. Debski, J. Gassin, & M. Smith (Eds.), *Language learning through social computing* (pp. 41–65). Melbourne, Australia: ALAA and the Horwood Language Centre.

Fishman, B., Marx, R. W., Blumenfeld, P., Krajcik, J., & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences*, 13(1), 43–76. doi:10.1207/2fs15327809jls1301_3

Gartner Inc. (2015). *Gartner Hype Cycle*. Retrieved from <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp>

Hamel, M.-J. (2012). Testing aspects of the usability of an online learner dictionary prototype: A product- and process-oriented study. *Computer Assisted Language Learning*, 25(4), 339–365. doi:10.1080/09588221.2011.591805

Hardisty, D., & Windeatt, S. (1989). *Computer assisted language learning*. Oxford, United Kingdom: Oxford University Press.

Hémard, D. (1997). Design principles and guidelines for authoring hypermedia language learning applications. *System*, 25(1), 9–27. doi:10.1016/S0346-251X(96)00057-7

Hémard, D. (2004). Enhancing online CALL design. *ReCALL*, 16(2), 502–519. doi:10.1017/S0958344004001727

Hémard, D. (2006). Evaluating hypermedia structures as a means of improving language learning strategies and motivation. *ReCALL*, 18(1), 24–44. doi:10.1017/ S0958344006000310

Hillier, V. (1990). Integrating a computer lab into an ESL program. *CAELL Journal*, 1(1), 23–24.

Huang, H. (2013). E-reading and e-discussion: EFL learners' perceptions of an e-book reading program. *Computer Assisted Language Learning*, 26(3), 258–281. doi:10.1080/09588221.2012.656313

Hubbard, P. (2004). Learner training for effective use of CALL. In S. Fotos & C. Browne (Eds.), *New perspectives on CALL for second language classrooms* (pp. 45–68). Mahwah, NJ: Lawrence Erlbaum Associates.

Hutchby, I. (2001). *Conversation and technology: From the telephone to the Internet*. Cambridge, United Kingdom: Polity Press.

Hutchby, I., & Barnett, S. (2005). Aspects of the sequential organization of mobile phone conversation. *Discourse Studies*, 7(2), 147–171. doi:10.1177/1461445605050364

Hutchby, I., & Tanna, V. (2008). Aspects of sequential organization in text message exchange. *Discourse and Communication*, 2(2), 143–164. doi:10.1177/1750481307088481

INCOSE (2015). *Systems engineering handbook: A guide for system life cycle processes and activities*. Compiled and edited by Walden, D., Roedler, G, Forsberg, K., Hamelin, R. D. and Shortell, T.

Kennedy, C., & Levy, M. (2009). Sustainability and CALL: Factors for success in a context of change. *Computer assisted Language Learning*, 22(5), 445–463. doi:10.1080/09588220903345218

Lafford, B. (2009). Towards an ecological CALL: Update to Garrett (1991). *The Modern Language Journal*, 93, 673–696. doi:10.1111/j.1540–4781.2009.00966.x

Lanham, R. A. (2006). *The economics of attention: Style and substance in the age of information*. Chicago, IL: University of Chicago Press.

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford, United Kingdom: Oxford University Press.

Levy, M. (1997). *Computer assisted language learning: Context and conceptualization*. Oxford, United Kingdom: Clarendon Press.

Levy, M. (2000). Scope, goals and methods in CALL research: Questions of coherence and autonomy. *ReCALL*, 12(2), 170–195.

Levy, M. (2002). CALL by design: Products, processes and methods. *ReCALL*, 14(1), 129–142. doi:10.1017/S0958344002000617

Levy, M. (2007). Climate change in CALL: From realigning the goals and technology options to breaking the "hype cycle." Plenary Address. Japan Association for Language Teaching (CALL SIG), 1–3 June, 2007, Waseda University, Japan.

Levy, M., & Gardner, R. (2012). Liminality in multitasking: Where talk and task collide in computer collaborations. *Language in Society*, 41(5), 557–587. doi:10.1017/S0047404512000656

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer assisted language learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. (Kindle ed.). New York, NY: Basic Books.

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer-mediated communication and language learning. *Computer assisted Language Learning*, 21(3), 227–251. doi:10.1080/09588220802090253

O'Rourke, B. (2012). Using eye-tracking to investigate gaze behaviour in synchronous computer-mediated communication for language learning. In M. Dooly & R. O'Dowd (Eds.), *Researching online foreign language interaction and exchange: Theories, methods and challenges* (pp. 305–341). Bern, Switzerland: Peter Lang.

Raby, F. (2005). A user-centered ergonomic approach to CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 179–190). New York, NY: Lawrence Erlbaum Associates.

Robinson, G. (1991). Effective feedback strategies in CALL: Learning theory and empirical research. In P. Dunkel (Ed.). *Computer assisted language learning and testing* (pp. 158–165). New York, NY: Newbury House.

Salomon, G. (1991). Transcending the qualitative-quantitative debate: The analytic and systemic approaches to educational research. *The Educational Researcher*, 20(6), 10–18. doi:10.3102/0013189X020006010

Skehan, P. (1998). *The cognitive approach to language learning*. Oxford, United Kingdom: Oxford University Press.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3), 185–211. doi:10.1177/136216889700100302

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9781139524780.009

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning and Technology*, 12(1), 85–103.

Smith, B., & Gorsuch, G. J. (2004). Synchronous computer-mediated communication captured by usability lab technologies: New interpretations. *System*, 32(4), 553–575. doi:10.1016/j.system.2004.09.012

Steel, C., & Levy, M. (2013). Language students and their technologies: Charting the evolution 2006–2011. *ReCALL*, 25(3), 306–320. doi:10.1017/S0958344013000128

Tono, Y. (2011). Application of eye-tracking in EFL learners' dictionary look-up process research. *International Journal of Lexicography*, 24(1), 124–153. doi:10.1093/ijl/ecq043

van Lier, L. (1998). All hooked up: An ecological look at computers in the classroom. *Studia Anglica Posnaniensia*, 33, 281–301.

Ward, M. (2006). Using software methods in CALL. *Computer Assisted Language Learning*, 19(2–3), 129–147. doi:10.1080/09588220600821487

Winograd, T. & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Reading, MA: Addison-Wesley.

PART II

Data and elicitation technologies and techniques

CHAPTER 6

Learner personas and the effects of instructional scaffolding on working behaviour and linguistic performance

Trude Heift
Simon Fraser University, Canada

This chapter examines data-driven learner personas and instructional scaffolding in the form of preemptive feedback in an ICALL environment. Ninety-three beginner learners of L2 German participated in a study by performing a sentence completion task as part of their regular course assignments throughout a semester. On the basis of their access to help throughout the study, participants were classified into three distinctive learner profiles, or personas: No Help, Sporadic Help, and Frequent Help personas. The study then investigated the effects of access to different amounts of help on the learners' working behaviour and linguistic performance. Study results indicate that the three learner personas showed significant differences in their working behaviour and linguistic performance, but by investigating the effects of the instructional scaffolding the CALL system provided, results suggest that two learner personas are sufficient to capture learners' differences. With the ultimate goal of understanding learner personas and instructional scaffolding as it relates to learning outcomes, satisfaction and success in CALL, this paper provides possible explanations of these study results and suggests areas for future research and development.

Keywords: scaffolding, help options, learner-computer interactions, German as a second language

Introduction

Compared to just a few years ago, the principal obstacles to computer-assisted instruction are no longer of technological nature. Instead, we are still wrestling with central pedagogical questions that have occupied the field of second language acquisition (SLA) for decades. One of these questions concerns the diversity of language learners. How can we devise ways of individualised instruction suited

DOI 10.1075/lsse.2.06hei

to a variety of learners by, at the same time, addressing the needs of individual learners? In what ways can and should CALL be individualised?

This chapter aims to address these questions by examining the effects of help access on the working behaviour and linguistic performance of 93 beginner L2 learners of German. For this study, our CALL environment displayed preemptive feedback in the form of lexical and grammatical hints specific to a learning activity with the goal to assist our L2 learners of German during task completion. Preemptive feedback is a type of instructional scaffolding, and, in contrast to reactive feedback, it initiates a focus-on-form phase so that learners receive relevant metalinguistic information before difficulties arise. This not only may lead to more successful task completion but also may reduce potential frustration by marking critical features in the language task (see Ellis, Basturkmen & Loewen, 2001).

By exploring data on the frequency of learners' help access of the preemptive feedback that our CALL program provided, we cluster our learners into different learner types, or personas (see Colpaert, 2004; Cooper, 1999; Levy & Stockwell, 2006; Nielsen, 2013) and then examine their subsequent working behaviour and linguistic performance while completing a set of form-focused L2 activities. More specifically, we examine whether our distinct learner personas look up correct answers without giving it a try and also inspect their error patterns.

In the following, we first situate our study in related literature on learner modelling, learner personas and preemptive feedback. We then describe our study participants and research methodology. The results section provides an examination of the effects of help access of preemptive feedback on the learners' working behaviour and performance. Our discussion of the results focuses on computational and pedagogical implications of the findings. The chapter concludes with opportunities for further research.

Literature review

Learner modelling

With the goal to explain why speakers choose, consciously or subconsciously, their forms of speech, SLA research has focused on learner variability, or interlanguage (IL) variation, by examining linguistic, psycholinguistic, and sociolinguistic factors and constraints. More generally, this body of research has provided evidence of extensive variability in learner language that can be attributed to individual differences, and task and external variables, all of which are said to affect the learners' learning processes.

That CALL activities can meet the individual differences and needs of learners is a claim that has been made since the earliest work on CALL. The goal was to use the computer to support classroom instruction in a way that would provide individualisation to meet learners' needs by identifying specific areas of knowledge, providing learning activities on these areas of knowledge for learners to complete, and tabulating learners' successes and errors within the knowledge categories. This work has become more sophisticated as developers of Intelligent Computer Assisted Language Learning (ICALL) applications explore more sophisticated exercise types requiring Natural Language Processing (NLP). NLP allows for a more delicate and potentially more useful analysis because the computer can analyse learners' language rather than simply categorising learner responses on selected-response items. NLP programs are also useful for modelling what the student knows based on the evidence found in his or her writing, and such models can be used for making suggestions about useful areas of instruction.

Learner modelling as an area of inquiry has been the focus and goal of intelligent language tutoring systems (ILTSs). An ILTS can adapt and tailor instructional materials and content to its users with AI techniques that are used to model the individualised learning experience and guide pedagogical decision-making. The goal here is to create learning programs that come closer to natural language interaction between humans than has been the case in traditional CALL applications. For this, the ILTS constructs a so-called learner model, which is a description of the learner's current skill level along with the student's learning styles and preferences relative to the learning task. Commonly, these models make some assumptions about the learner by determining her or his current knowledge state, which requires the ILTS to observe and record the learner's interaction with the learning system. Measuring learner knowledge, however, is a highly complex task due to a number of variables that have to be considered in assessing and capturing the individual differences that warrant individualisation at any given point in time. A number of learner models have been described and implemented, and, most commonly, they are used to generate individualised feedback and unique learning paths for each learner.

One of the challenges of learner modelling, however, refers to the fact that it is impractical for an ILTS to accommodate the different skills, preferences, and needs of each and every learner. How, then, can we best individualise instruction?

Learner personas

With the concept of learner personas, we can capture and cluster similarities and differences among learners and then model the learning process in areas relevant

to individual learners and learning situations. Personas are archetypal users of a learning tool that represent the needs of larger groups of users in terms of their goals and personal characteristics. They are described and pieced together based on relevant information from knowledge about real people.

The usefulness of personas in defining and designing interactive applications is based on ideas by Alan Cooper, the father of Visual Basic, and expressed in his book entitled *The Inmates Are Running the Asylum* (Cooper, 1999). As part of his goal-oriented interaction design to real-world business environments,[1] which places an emphasis on the users' (work) goals, such as workflow, contexts and attitudes of the persona, Cooper (1999) believed that, in contrast with iterative user prototyping, the most powerful method is to make up "pretend users and design for them" based on in-depth ethnographic data (p. 123). According to Cooper (1999), these personas should be established during the initial conceptualization phase of software design, or to express his view on the development process of personas, more generally: "To deliver both power and pleasure to users, you need to think first conceptually, then in terms of behaviour, and last in terms of interface" (p. 23).

Lilley, Pyper and Attwood (2012) have made a distinction between ad-hoc personas and data-driven personas. Ad-hoc personas are defined during the conceptualization phase and based on pre-conceptions of what software designers think users might be like. In contrast, data-driven personas are established through data collections from actual users. This might include data on their demographics, gathered through user surveys and/or concurrent system interactions. Nielson's (2013) process model, for instance, contains ten steps split into four different parts to define data-driven personas: "Data collection and analysis of data, personas descriptions, scenarios for problem analysis and idea development, and acceptance for the organization and involvement of the design team" (p. 10). Accordingly, the model covers the entire process from the preliminary data collection, through active use, to continued development of personas.

Although both ad-hoc and data-driven personas are fictitious, with the concept of personas, software designers aim at defining and grouping similarities and differences among users by considering user demographics as well as design features that go with those. Naturally, the only time they really matter is when those demographics directly affect user behaviours and performances. The differences among distinct personas must then be based on deeper issues, for instance, what users do (actions or projected actions), why they do them (goals and motiva-

1. For other approaches to personas (e.g., the role-based, engaging, and fiction-based perspectives), see Nielsen (2013).

tions) and not as much on who the users are (see also Calabria, 2004). Once the similarities and differences have been determined, user interaction can be modelled in areas relevant and appropriate to a particular learning tool and/or environment.

In our current investigation, we are interested in determining the effects of instructional scaffolding in the form of preemptive feedback on the learners' working behaviour and linguistic performance. Unlike Cooper's (1999) approach to defining personas with ethnographic data, however, we base our definition and classification of learner personas on the learners' concurrent interactions with the system; that is, we construct data-driven personas. Accordingly, we first capture our learners' interactions with the system, and from those we establish our personas with respect to the preemptive feedback they received and consulted. An important question here is: How many personas do we construct? According to Nielsen (2013), the number depends on how different the users are, but, as Cooper (1999) emphasized, the number should be reasonably small to keep them distinct. In any case, once the personas are defined, the CALL system then models each learner according to the characteristics of his or her persona. The information about each persona should be dynamic in the sense that it changes over time and adjusts to learners as they progress in their understanding of the subject matter. Possibly, this knowledge can also be negotiated with learners and manipulated accordingly.

From a software-development perspective, the design of our data-driven personas follows our general approach of a cyclical process of development, implementation and evaluation to software engineering (see Colpaert, 2004). Such a holistic and cyclical approach to software engineering is generally preferred (see Caws, 2013; Hubbard, 2011) because each and every stage during the lifecycle delivers output that serves as input for the subsequent stage. Accordingly, and based on observations of learner interactions with the CALL system and subsequent data analyses, our personas are likely to be revised and/or adjusted as learner behaviour and performance change over time.

Preemptive feedback is one area that lends itself well to exploring the concept of learner personas and a topic that, despite its great potential for individualising the language learning process during task completion, has not received much attention in CALL system design. The following section contextualizes preemptive feedback within existing literature.

Instructional scaffolding: Preemptive feedback

Over the past decades, CALL systems have given increasingly more importance to pedagogical, user-centred designs by emphasizing, among other aspects, pedagogical interventions that enhance the learner-computer interactions during task performance. One way to assist learners with a task is to provide scaffolding in the form of hints and reminders that coach learners about their work and progress. Indeed, scaffolding in CALL has commonly been employed in the form of help options for task completion and also learner feedback.

The term scaffolding originates from the work of Jerome Bruner (1983) who defined it as "a process of 'setting up' the situation to make the child's entry easy and successful and then gradually pulling back and handing the role to the child as he becomes skilled enough to manage it" (p. 60). These ideas are strongly associated with sociocultural theory (see Lantolf & Thorne, 2006), and, applied to CALL, scaffolding is generally understood as the instructional assistance provided by a CALL program during learner-computer interactions.

The notion of scaffolding has also been adopted in research on technological support for learning, which has become increasingly important in pedagogical, i.e., user-centred designs (for a more extensive overview, see Quintana et al., 2004). In these contexts, the intention is that the support not only assists learners in accomplishing tasks but also enables them to learn from the experience. Moreover, in this framework, scaffolding refers to ways in which the software tool itself can support learners as opposed to only teachers or peers.

Previous research in this context has shown that the use of scaffolding can guide students in knowledge construction, knowledge integration, and knowledge representation during their work on performing learning tasks (e.g., Chang & Sun, 2009; Van Merriënboer et al., 2003). Moreover, studies have also presented evidence of the cognitive benefits of scaffolding, particularly in eliciting self-explanation, self-questioning, self-monitoring, and self-reflection during learning (e.g., Ge et al., 2005).

Scaffolding in computer-aided environments can, however, be achieved in a number of ways. Guzdial (1994), for instance, has outlined three roles software could play in scaffolding:

1. communicating processes to learners
2. eliciting articulation from learners to encourage reflection
3. coaching learners with hints and reminders about their work

While the distinct roles of scaffolding described in 1 and 2 are mainly concerned with the students' cognitive processing of a task, its role outlined in 3 refers to

guidance on student input during task completion, which, in CALL, has most commonly been implemented as learner, or corrective, feedback.

Indeed, a large body of research both in face-to-face settings as well as in CALL environments has focused on learner feedback. From an SLA perspective, these studies mainly have focused on research exploring the interaction hypothesis (Long, 1991) and the input-interaction-output model (Gass & Selinker, 2001). For instance, Long and Robinson (1998) identified two kinds of responses to learner input, with the goal to draw the learner's attention to form: reactive and preemptive. Reactive focus on form is also commonly referred to as corrective feedback, error correction, or negative evidence/feedback, and it supplies learners with either explicit or implicit negative evidence. It generally occurs in reaction to learner errors, which are then addressed by, for instance, the teacher or a CALL program. In contrast, preemptive feedback draws attention to potentially problematic areas in the task by initiating a focus-on-form phase so that learners receive relevant meta-linguistic information before difficulties arise. One of the goals here is to reduce potential frustration by marking critical features in the language task to increase task completion (see Ellis, Basturkmen & Loewen, 2001).

Preemptive feedback may also assist in providing learners with explicit knowledge, which, as Ellis (1993) has argued, constitutes a valid goal for instruction because it helps improve performance through monitoring and facilitating acquisition through noticing. According to Schmidt's (1994) Noticing Hypothesis, language learners are limited in what they are able to notice, and the main determining factor is that of attention. Schmidt (1994) argued that attention is not only necessary for acquisition to take place, but noticing is also a conscious process in that "attention also controls access to conscious experience thus allowing the acquisition of new items to take place" (p. 176). Accordingly, form-focused instruction that induces learners to pay conscious attention to forms in the input can assist interlanguage development.

The effects of preemptive feedback in CALL have hardly been studied empirically and are thus speculative and deserve closer investigation (see Ellis, 2001, & Farrokhi et al., 2008, for face-to-face studies). One likely reason for this lack of research might be that preemptive feedback requires some kind of error analysis that makes predictions about the most likely error(s) that may occur with a given exercise. While language instructors, based on their teaching experience, may be able to predict errors intuitively and fairly accurately, a CALL program either needs to encode this knowledge manually, which is a very onerous task, or needs to consult a learner corpus for a specific set of learning activities. In an attempt to assist learners during task completion, Heift (2013) designed a learner corpus from previous users and investigated different types of preemptive feedback of varying specificity with the goal of drawing the learners' attention to the most

common errors in a given exercise. Her findings indicate that, for her beginner and early intermediate L2 learners of German, both types of preemptive feedback were significantly more effective than not providing any assistance before students attempted to complete a task. Moreover, the beginner learners significantly outperformed the early intermediate students, and by considering the two types of preemptive feedback in relation to different error types, the study suggests that at an intermediate level, students are more likely and/or seem more able to pay attention to multiple pieces of information contained in the preemptive feedback.

The current study addresses this general lack of CALL research in the areas of learner personas and preemptive feedback by examining the effects of instructional scaffolding on the learners' working behaviour and linguistic performance during a form-focused language learning activity. The following section outlines our research questions and methodology.

Our study

Research questions

The current study focuses on the following two research questions:

1. In what ways does the working behaviour of the different learner personas of help access of preemptive feedback vary, as measured by their answer look-up behaviour?
2. In what ways does the linguistic performance of the different learner personas of help access of preemptive feedback vary?

In the following, we describe our research methodology by detailing our study participants, data collection and analysis.

Study participants

The 93 L2 learners of German who participated in the study were all registered in a beginner L2 German course in Fall 2013 at a Canadian university. As determined by their previous exposure to German and/or a university placement test, the study participants had no prior knowledge of German. At the beginning of the semester, all study participants consented to a possible anonymous analysis of their data for research purposes. A background questionnaire, which we administered at the beginning of the course, revealed that 55 students were female and

38 male. The learners were all proficient in English, with native languages varying from English, Chinese, Korean, Farsi, Russian and Polish.

Data collection

The data were collected with the built-in tracking system of *E-Tutor* (www.e-tutor.sfu.ca). E-Tutor is a web-based intelligent CALL (ICALL) system for L2 learners of German that covers the content of the first three university courses of German during which the main components of the L2 grammar are generally taught. The system follows the grammatical and vocabulary sequence of *Deutsch: Na klar!* (Di Donato, Clyde, & Vansant, 2004), a textbook commonly used in North America for L2 learners of German. The fifteen chapters in E-Tutor each provides a variety of learning activities that allow students to practice chapter-related vocabulary and grammar. In addition, students can practice their pronunciation, listening comprehension, reading and writing. The system also contains cultural information on Germany and its people with chapter-related texts, authentic pictures and audio recordings. E-Tutor is commonly used in conjunction with regular face-to-face instruction whereby students complete the learning activities as part of their homework assignments.

Unlike more traditional CALL systems, E-Tutor uses Natural Language Processing to provide a linguistic analysis of learner input and to generate error-specific feedback. This parsing technology allows the system to perform a linguistic analysis of the input and then inform the learner of the exact source of an error, mainly with respect to lexical and grammatical errors. The system also tracks the learners' linguistic knowledge over time by keeping a very detailed record of their behaviours and performances (for a more detailed description of the system, see Heift, 2010). From a research perspective, and given the complexity and ongoing classroom use of the system, E-Tutor lends itself very well to investigate a variety of CALL-related topics and issues. For this reason, the system has been used in a number of studies that investigated learner-computer interactions, such as learner feedback and learner modelling (e.g., Heift, 2004, 2008; Heift & Rimrott, 2012).

For the purpose of this study, we consider learner data from the build-a-sentence activity type (see Figure 6.1), which students completed as part of their regular homework assignments throughout the semester.

In the build-a-sentence learning activity, students are given a prompt and asked to construct a sentence by applying the correct inflections (e.g., for articles, verbs) and word order. For instance, consider Example (1), which displays the prompt and the correct answer for the exercise given in Figure 6.1.

Figure 6.1 Build-a-sentence activity in E-Tutor

(1) Prompt: wo / du / (def. article) / Rock / kaufen?
Where / you (sg.) / skirt / buy?
Answer: Wo kaufst du den Rock?
Where do you buy the skirt?

In Example (1), students need to apply the correct word order for German question formation, supply the correct article for the accusative case of the direct object (*den*) and inflect the verb *kaufen* for second person singular (*kaufst*).

The interface, which is similar for all learning activities, consists of an exercise prompt, followed by an input field with three buttons: CHECK allows students to submit the answer for answer processing, SOLVE allows learners to look up possible answers for a given exercise and SKIP advances to the next exercise. For pedagogical reasons, the error checking process of *E-Tutor* is iterative; that is, the system identifies and communicates one error at a time to the learner. Once the learner has revised the input, s/he resubmits the sentence for further analysis. The iterative error-correction process continues until the sentence is correct, or until the learner clicks the SOLVE button, thus *peeking* at the answer while no longer giving it a try. E-Tutor tracks all user interactions with the program by

also recording a detailed description of their errors and correct responses. This is possible due to the NLP component that is part of the system.

In the lower half of the user interface, the system displays the learner feedback (Feedback tab). In addition, students can look up their performance for each of the exercises (History tab). This is especially useful if students take several iterations before achieving a correct answer. Students can also obtain grammar help and perform dictionary look-ups (Grammar Help and Dictionary tab, respectively). Finally, students can examine the error profile for each exercise based on our learner corpus, which we discuss in the following section.

Preemptive feedback

To construct the preemptive feedback for the exercises contained in the E-Tutor, we created a learner corpus consisting of several million responses submitted by roughly 5000 previous learners who had completed the activity types of the E-Tutor between 2003 and 2008. We conducted an extensive statistical analysis for these millions of entries, and, for each exercise, activity type and chapter, we produced a ranked list of errors based on prior students' performance during those years. For each error profile, we then generated preemptive feedback that the system displays when students start an exercise (see Figure 6.1: "Tip: Be careful with article inflection (definite article)").

For the exercise given in Example (1), for instance, we determined the following error ranking:

1. Correct responses: 36%
2. Errors: 64%
    a. 41.8% article inflection
    b. 7.2% verb inflection
    c. 5.6% extra/missing words
    d. 4.4% word order
    e. 3.9% spelling
    f. 1.1% capitalisation

The statistical analysis revealed that 64% of the roughly 5000 student responses for this particular exercise were correct while 64% contained an error. Of the incorrect responses, 41.8% contained a wrong article inflection (e.g., *der* instead of *den*), followed by an incorrect verb inflection (e.g., *kaufen* instead of *kaufst*), an extra or missing word (5.6%), word order (4.4%), a spelling mistake (3.9%), and, finally, wrong capitalisation (1.1%). Accordingly, the preemptive feedback for the exercises in E-Tutor is based on an error ranking that is created from the error

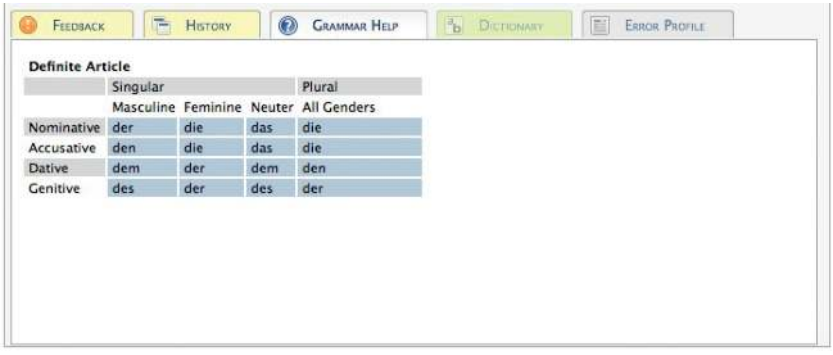


Figure 6.2 Link to definite articles

profiles of thousands of previous users. It reflects the most common errors unique to each individual exercise and activity type.

The preemptive feedback of E-Tutor also displays links to an inflectional paradigm or rule explanation in the case of a grammatical hint. For spelling mistakes, the system links to the E-Tutor's dictionary, which contains approximately 20,000 entries. For instance, for the example provided in Figure 6.1, the ICALL system displays the following preemptive feedback: "Tip: Be careful with article inflection (definite article)". When the student clicks on the link *definite article*, the system generates the declensions of the German definite articles, as given in Figure 6.2.

Data analysis

The build-a-sentence activity, which we considered for this study, contained twenty individual exercises for each of the four chapters that students completed throughout the semester.

The study participants' help access was determined by counting the instances when students clicked on the help link that was provided as part of the preemptive feedback (e.g., the link *definite articles* displayed in Figure 6.1). For each student, we then divided the total number of help access (clicks) by the total number of exercises that students completed (= 80). For the peeks, we counted the total number of times a student clicked the SOLVE button, thus peeking at the answer instead of working out the correct sentence by themselves. For each student, we then divided that number by the total number of exercises. For the errors, we counted the total number of errors for each student and exercise on first submissions, i.e., before students received any system hints on their error(s), and then divided that number by the total number of exercises.

For the inferential statistics, we applied one-way ANOVA with pairwise comparisons as post-hoc tests. Bonferroni was used to adjust for multiple comparisons. In the case of two groups, we applied an independent sample *t*-test. For all tests, an alpha level of 0.5 was used.

Results

Due to the iterative correction process of E-Tutor, the total number of submissions per student is naturally higher than the total number of exercises. In addition, the total number of submissions varies among our study participants because students committed a different amount of errors. Accordingly, we collected a total of 8,540 sentence submissions for the build-a-sentence learning activity and the four chapters that the 93 study participants completed throughout the semester. This averages to 91 submissions per student in total or 2.3 submissions per student and exercise. This is in accordance with previous studies undertaken with E-Tutor, where we generally found that it takes students on average 2-3 submissions to achieve a correct answer.

In order to answer our two research questions, our first goal was to establish distinct learner types based on their help access of the links that the preemptive feedback in E-Tutor provided. For this, we collected interaction data from 123 students who were enrolled in the beginner course of L2 German. In examining the data, we discovered that 31 of the 123 students never clicked on any of the links of the preemptive feedback that *E-Tutor* provided. Naturally, we were interested in the working behaviour and linguistic performance of these students as one of our help access personas. We then examined the data of the remaining 82 students and found a somewhat natural split between their amount of help access at around 20% of overall help access. To end up with equal sample sizes in the three groups and thus to increase the statistical power and reliability of the data, we then randomly selected, by using MS Excel's random function, 31 students from the pool of students who accessed help less than 19% of the time and those who accessed it more than 21% of the time. This resulted in a total count of 93 study participants, 31 students per group.

Accordingly, our investigation described here considers the working behaviour and linguistic performance of the following distinct learner types; the first group includes learners who never accessed any of the links that our preemptive feedback displayed, and thus we refer to them as the *No help group*. In the second group, we see learners who occasionally accessed the links, and we call them the *Sporadic help group*. The final group consists of learners who accessed the help

Table 6.1 Help access for the three personas

	Mean	Std. deviation	Minimum	Maximum
No help group (n = 31)	.0000	.0000	.0000	.0000
Sporadic help group (n = 31)	.0937	.0568	.0120	.1943
Frequent help group (n = 31)	.3333	.1081	.2167	.5833

links far more often than the remaining two groups, and we refer to them as the *Frequent help group*.

Table 6.1 specifies the help access for the three distinctive groups. It indicates that the help access that our study participants sought throughout their language practice over the semester ranged from 0% to 58.3%. The No help group never clicked on the links that our preemptive feedback provided, while the Sporadic help group on average accessed the links 9% of the time, followed by the Frequent help group with 33.3% of the time.

Research questions 1 and 2: Working behaviour and linguistic performance

Our first research question investigated whether our three distinct learner personas peeked at a correct answer for an exercise rather than working through the learning activity and providing the answer by themselves. Table 6.2 displays the descriptive statistics for the three learner personas. It shows that the No help group peeked at the correct answer most often (20.9%), followed by the Sporadic help group (7.4%) and, finally, the Frequent help group (7.2%). For the inferential statistics, one-way ANOVA indicates a main effect of peeks ($F(2, 90) = 6.761$, p = .002). To determine inter-group variation, we applied a follow-up Bonferroni test, which shows a significant difference between the No help and the Sporadic help group (p = .005), and between the No help and the Frequent help group (p = .007). No significant difference was found between the Sporadic and the Frequent help group (p = 1.000).

Our second research question examined our learners' linguistic performance. The data in Table 6.2 show that the No help group committed the most errors

Table 6.2 Peeks and error rates for the three personas

	Working behaviour		Linguistic performance	
	Mean	Std. deviation	Mean	Std. deviation
No help group (n = 31)	0.2098	0.2783	0.5661	0.1600
Sporadic help group (n = 31)	0.0744	0.0727	0.5193	0.1474
Frequent help group (n = 31)	0.0720	0.0528	0.4371	0.1504

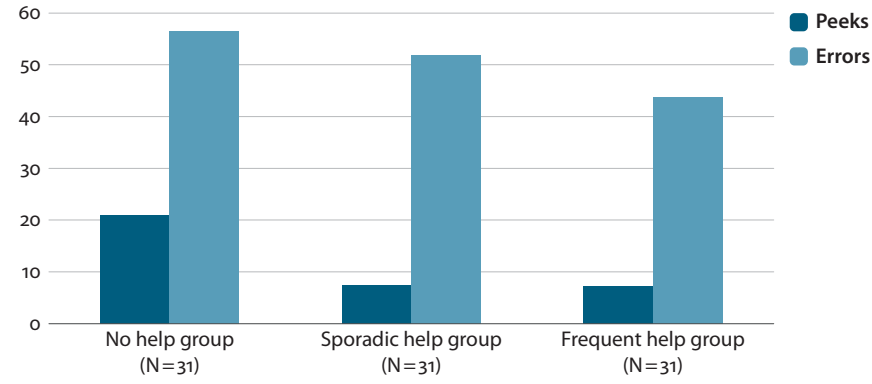


Figure 6.3 Peeks and errors for the three personas

on the E-Tutor exercises (56.6%), followed by the Sporadic (51.9%) and, finally, the Frequent Help group (43.7%). The percentages imply that students in general needed about two submissions to arrive at a correct answer. As for the inferential statistics, again, we applied one-way ANOVA and found a main effect of linguistic performance ($F(2, 90) = 5.669$, $p = .005$). To investigate inter-group variation, Bonferroni indicated a significant difference between the No help and Frequent help group ($p = .004$), while no significant differences between the remaining groups were found (No help and Sporadic help group, $p = .693$; Sporadic help and Frequent help group, $p = .110$).

The chart given in Figure 6.3 summarizes our findings with respect to the learners' working behaviour and linguistic performance, grouped by our three learner personas.

The following section discusses these findings in more detail.

Discussion

Our study results indicate significant differences in the learners' help access and their subsequent working behaviour and linguistic performance. As for their working behaviour, we observed significant differences between the No Help and both the Sporadic and Frequent help access personas, but no significant differences between the Sporadic and the Frequent help personas were found. With regards to the learners' linguistic performance, a significant difference between the No Help and the Frequent help personas was noted, while the differences between the remaining groups were comparable with respect to linguistic performance. These results make a number of pedagogical and computational suggestions.

From the perspective of scaffolding and learner personas, our results suggest that with respect to the interaction variables we investigated, two instead of three personas might have been sufficient given that we found no significant differences between the Sporadic and Frequent help groups in their working behaviour and performance. Accordingly, a broader, less fine-grained, classification of help access may seem appropriate for individualising the learning process as it relates to the preemptive feedback E-Tutor provided, at least for our study participants and the environment in which they were tested. This is in accordance with Cooper's (1999) suggestion of keeping the number of personas reasonably small to keep them distinct. Naturally, if the groups do not exhibit different behaviours and/or performances, there is little need for the design and implementation of different personas.

To test the concept of a reduction in personas, we ran a subsequent analysis to investigate the significance levels for our learners by splitting them into only two help access groups of 46 and 47 study participants each. Naturally, the differences between the two groups became more pronounced.

Table 6.3 displays our results and indicates that the persona with little help access not only peeked at the correct answer more often (16.9%) than the persona with lots of help access (6.8%) but also committed more errors (56.5% versus 44.7%).

A subsequent independent samples *t*-test confirmed that the two groups are significantly different in both factors under investigation: Working behaviour ($t(91) = 2.813$, $p = .006$) and linguistic performance ($t(91) = -3.799$, $p < .001$). These results highlight the fact that a multitude of factors has to be considered when defining personas. In our case, and only due to the analysis of the effects of help access on additional variables, we were able to observe that two personas are sufficient to describe the learners' working behaviour and linguistic performance in this particular aspect of the learning process.

In taking a broader view, our interaction-based research, coupled with a data-driven approach to personas, underlies our general and cyclical approach to software engineering, which is also central to CALL ergonomics (see Chapter 2, this volume). Research in CALL ergonomics relies on the observation of user

Table 6.3 Peeks and error rates for two personas

	Working behaviour		Linguistic performance	
	Mean	**Std. deviation**	**Mean**	**Std. deviation**
Little help access ($n = 46$) (Mean = .0158, STD = .0266)	0.1697	0.2366	0.5658	0.1475
Lots of help access ($n = 47$) (Mean = .2661, STD = .1312)	0.0689	0.0651	0.4478	0.1519

behaviour during CALL activities by paying close attention to the relationship between the user and the tool, whereby the user plays a key role (e.g., Bertin et al., 2010; Raby, 2005). The importance of this cyclical process is underscored by existing research but also by our current study. For instance, some ergonomics research on CALL systems, which appeared to be user-friendly at first glance, showed that learners were not always performing well because the technology had not been adapted to their needs (e.g., Caws, 2013; Hamel, 2012). In our case, and although the initial classification of our data-driven personas was based on interaction-based research, by investigating the effects of help access on learners' working behaviour and linguistic performance and thus examining additional interaction-based data, we learned that a more coarse-grained division of personas (i.e., two as opposed to three learner types) is sufficient. This is motivated by the fact that no significant differences between the Sporadic and No Help group in their linguistic performance were found, thereby suggesting that, with regards to linguistic performance, the Sporadic Help group clearly falls in between the two remaining groups. This also supports Caws and Hamel's (2013) approach in that any data collections as well as their interpretations have to be recycled into new learning processes and technological design.

CALL ergonomics also places a strong emphasis on learning processes, as opposed to outcomes. Our research suggests that while, in the end, emphasis might play a key role, learning outcomes also tell us something about the learning behaviour that leads to successful learning. In our study, and by learners not taking advantage of the scaffolding that our preemptive feedback provided, the learning behaviour and processes clearly impacted the learning outcomes. Study participants with the least help access looked up the answers most often by also committing the most errors.

Conclusion

This chapter investigated learner personas and preemptive feedback in the context of L2 German in a CALL environment. By grouping our study participants into three significantly different personas of help access, we were able to observe distinctive working behaviours and linguistic performances among the three groups. More specifically, our findings indicate significant differences between the No help and the two remaining personas. The No Help group peeked at the correct answer significantly more often while also committing significantly more errors. In contrast, we did not observe significant differences between the Sporadic and Frequent help groups with regards to their working behaviour and linguistic performance.

From a pedagogical perspective, our study suggests that preemptive feedback not only leads to more successful task completion but also may reduce frustration given that learners commit fewer errors. As observed in previous studies (e.g., Heift, 2013), learners seem quite concerned about making errors, independent of whether or not the errors contribute to their course grade. Preemptive feedback cuts down on the number of errors, and this may lead to a more positive learning experience. Our study further suggests that the preemptive feedback in E-Tutor might provide part of the meaningful interaction we are seeking in CALL by minimizing learner errors from the start, thus reducing learner frustration and enhancing their L2 development. Finally, while the personas we identified provide justification to individualize the learning process, the better learning outcomes we noted with the personas that made use of the preemptive feedback suggest that those learners that generally tend to not seek help should possibly be encouraged to do so. This process can certainly become an integral part of the CALL program in the form of learner modelling. For instance, the CALL application can draw the attention to the preemptive feedback of those learners that perform poorly and ignore it. Moreover, even the preemptive feedback itself can be closely modelled and become more individualized by displaying only those hints that are relevant to particular students. For instance, for those personas who have a good understanding of word order, as indicated by their past performance history, any preemptive feedback hints on word order can be omitted for them.

From a computational perspective, the preemptive feedback in E-Tutor is based on extensive statistical analyses of a learner corpus, which resulted in an error ranking for each activity type and exercise. These error rankings are fundamental to the preemptive feedback E-Tutor displays to the learner. However, preemptive feedback can certainly be achieved and implemented without a learner corpus by relying on language teachers to predict the most likely errors although this would make the classification and process less empirical and more onerous. With regards to designing personas, our data suggest that by examining learner-computer interactions, we are able to observe and identify different working behaviours and linguistic performances and capture and cluster similarities and differences among language learners accordingly. This allows us to individualize the learning process more effectively than trying to adjust to the needs of each and every learner. However, we also learned that it is important to examine the complex interactions of several variables rather than treating each in isolation to come up with the optimum number of personas for a given learning situation. The concept of personas, however, can be easily expanded to capture additional factors, which may impact other learning processes by examining those during learner-computer interactions.

References

Bertin, J. C. and Gravé, J. P. (2010). In favor of a model of didactic ergonomics. In J. C. Bertin, P. Gravé and J. P. Narcy-Combes (Eds.), *Second language distance learning and teaching: theoretical perspectives and didactic ergonomics* (pp. 1–36). Hershey, PA: Information Science Reference. doi:10.4018/978-1-61520-707-7.ch001

Bruner, J. (1983). *Child's talk*. New York, NY: Norton.

Calabria, T. (2004). An introduction to personas and how to create them. *KM Column*. Retrieved from <http://www.steptwo.com.au/papers/kmc_personas/>

Caws, C. (2013). Evaluating a web-based video corpus through an analysis of user interactions. *ReCALL*, 25(1), 85–104. doi:10.1017/S0958344012000262

Caws, C., & Hamel, M.-J. (2013). From analysis to training: Recycling interaction data into learning processes. *OLBI Working Papers*, 5, 25–36. doi:10.18192/olbiwp.v5i0.1116

Chang, W., & Sun, Y. (2009). Scaffolding and web concordancers as support for language learning. *Computer Assisted Language Learning*, 22(4), 283–302. doi:10.1080/09588220903184518

Colpaert, J. (2004). *Design of online interactive language courseware: Conceptualization, specification and prototyping. Research into the impact of linguistic-didactic functionality on software architecture* (Unpublished doctoral dissertation). University of Antwerp, Antwerp, Netherlands.

Cooper, A. (1999). *The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity*. Indianapolis, IN: SAMS.

Di Donato, R., Clyde, M. D., & Vansant, J. (2008). *Deutsch: Na klar! An introductory German course*. New York, NY: McGraw Hill.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, 27(1), 91–113. doi:10.2307/3586953

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, 35, 407–432. doi:10.2307/3588029

Farrokhi, F., Ansarin, A., & Mohammadnia, Z. (2008). Preemptive focus on form: Teachers' practices across proficiencies. *The Linguistics Journal*, 3(2), 7–30.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course* (2nd ed.). New Jersey, NJ: Lawrence Erlbaum Associates.

Ge, X., Chen, C. H., & Davis, K. A. (2005). Scaffolding novice instructional designers' problem-solving processes using question prompts in a web-based learning environment. *Journal of Educational Computing Research*, 33(2), 219–248. doi:10.2190/5F6J-HHVF-2U2B-8T3G

Guzdial, M. (1994). Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments*, 4, 1–44. doi:10.1080/1049482940040101

Hamel, M.-J. (2012). Testing aspects of the usability of an online learner dictionary prototype: Process and product oriented analysis. *Computer Assisted Language Learning*, 25(4), 339–365. doi:10.1080/09588221.2011.591805

Heift, T. (2004). Corrective feedback and learner uptake in CALL. *ReCALL*, 16(2), 416–431. doi:10.1017/S0958344004001120

Heift, T. (2008). Modeling learner variability in CALL. *Computer Assisted Language Learning*, 21(4), 305–321. doi:10.1080/09588220802343421

Heift, T. (2010). Developing an intelligent language tutor. *CALICO*, 27(3), 443–459. doi:10.1558/cj.v27i3.443-459

Heift, T. (2013). Preemptive feedback in CALL. In A. Mackey & K. McDonough (Eds.), *Interaction in diverse educational settings* (pp. 189–207). Amsterdam, Netherlands: John Benjamins. doi:10.1075/lllt.34.14ch10

Heift, T., & Rimrott, A. (2012). Task-related variation in computer assisted language learning. *Modern Language Journal*, 96(4), 525–543. doi:10.1111/j.1540-4781.2012.01392.x

Hubbard, P. (2011). Evaluation of courseware and websites. In L. Ducate & N. Arnold (Eds.), *Present and future perspectives of CALL: From theory and research to new directions in foreign language teaching* (2nd ed., pp. 407–440). San Marcos, TX: Computer Assisted Language Instruction Consortium.

Lantolf, J. P., & Thorne, S. L. (2006). Sociocultural theory and second language acquisition. In B. van Patten & J. Williams. (Eds.), *Explaining second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer assisted language learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lilley, M., Pyper, A., & Attwood, S. (2012). Understanding the student experience through the use of personas. *Innovation in Teaching and Learning in Information and Computer Science*, 11(1), 4–13. doi:10.11120/ital.2012.11010004

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam, Netherlands: John Benjamins. doi:10.1075/sibil.2.07lon

Long, M., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–63). Cambridge, United Kingdom: Cambridge University Press.

Nielsen, L. (2013). *Personas – User focused design*. London, United Kingdom: Springer. doi:10.1007/978-1-4471-4084-9

Quintana, C., Reiser, B., Davis, E., Krajcik, J., Fretz, E., Duncan, R., … Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337–386. doi:10.1207/s15327809jls1303_4

Raby, F. (2005). A user-centered ergonomic approach to CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 179–190). Mahwah, NJ: Lawrence Erlbaum Associates.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11–26.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, 38(1), 5–13. doi:10.1207/S15326985EP3801_2

CHAPTER 7

Video screen capture to document and scaffold the L2 writing process

Marie-Josée Hamel and Jérémie Séror
University of Ottawa, Canada

This chapter explores the potential of video screen capture (VSC) as a technology that can provide new insights when investigating learner-computer interactions in CALL research, and that can play a mediating role in second language (L2) writing pedagogy. Arguments are put forward as to why CALL researchers and language educators should be interested in this accessible and flexible tool. Three studies are described to consolidate these arguments. The first one, a usability study, investigates L2 learners' dictionary search processes in the context of the design of an online dictionary prototype. The second study examines the composition processes and strategies of L2 writers. The third study looks at the pertinence and added value of integrating VSC in the L2 writing class. Affordances of VSC arose from these studies. VSC emerged as a powerful documentation tool enabling the collection of process-oriented learner data and new forms of dynamic corpora. It also emerged as a retrospection tool capable of supporting L2 writers in their literacy development and as a scaffolding tool to provide multimodal feedback on L2 written output.

Keywords: video screen capture, L2 writing process, learner-computer interaction, dynamic learner corpora, affordances

Introduction

In the field of digital literacies, a number of technological innovations are transforming how individuals can engage in meaning making activities (Lea, 2013; Stapleton, 2010). Technologies such as tablets, voice recognition and motion capture are but a few examples of these technologies. Indeed, homework assignments, quizzes, exposure to authentic input through reading and listening tasks, composition practices, and even conversations are all pedagogical tasks that are increasingly occurring electronically in digital spaces, i.e., through computer-mediated

DOI 10.1075/lsse.2.07ham

environments, most often on some kind of digital screen. These technologies stand out for their potential to radically transform human-computer interaction (HCI) or what the field of CALL refers to as learner-computer interaction (LCI).

This chapter focuses on the affordances of one specific new technology particularly suited to document and exploit what happens on the digital screen when learners interact with a computer: video screen capture (VSC) technology. Drawing on the authors' own experiences with this technology and its application in three separate research projects, it will be argued that in the context of CALL, VSC offers researchers, instructors and their students a powerful means of achieving new insights and opportunities to enrich our understanding of the link between second language literacy skills development and computer-mediated language tasks (Barbier & Spinelli-Jullien, 2009).

In what follows, we will provide an overview of the nature of VSC and its appeal to researchers and educators. We will then briefly present three studies which have used VSC to document and/or enhance both students' and instructors' literacy practices in the context of CALL design and L2 pedagogy. These will be used to illustrate the nature of the data that can be collected through VSC, how it can be analysed and the types of insights it can lead to. Finally, we will highlight the unique affordances of VSC and their implications for how VSC can help advance CALL research and the design of CALL pedagogy and teacher training.

What is VSC?

VSC technology has emerged in the last few years as an increasingly popular tool used to create audio-visual documents that can help computer users share images and movies of what they do on their computer screens. In essence, VSC refers to the use of software that will allow one to record a movie of on-screen actions, which occur as an individual interacts with a computer (or a mobile device screen).

VSC is perhaps best illustrated by the growing number of self-help videos which can be found on YouTube where experts explain step by step how to use a piece of software or how to accomplish a complicated task on a computer. These movies offer an over-the-shoulder effect similar to one-on-one instruction (Carr & Ly, 2009).

Screen-recording videos are often accompanied by a voice-over recorded by the author of the recording. This voice-over provides off-screen commentary and explanations of what occurs on the screen. To create this voice-over, VSC users can choose to record their voices simultaneously as they record their screens or later in a subsequent stage as they edit the video. Additionally, audio tracks

can also be added which include all sounds generated by the computer itself (i.e., mouse clicks, the active pressing of a button, the sound of audio and video recordings played on a computer, etc.). Interestingly for research, the audio track can also at times capture indirect external sounds (such as typing noises, ambient music or the sound of pages being flipped).

Finally, many screen capture software programs allow users the choice to include in their recordings additional sources of video input in the form of the images recorded by their computers' webcams. If this option is selected, the videos produced show both what is happening on the screen as well as, often in a smaller window, a video of the user's face as he or she is interacting with the computer.

As such, through a combination of moving images and sounds, users of VSC can share with others audio-visual recordings of their actions in digital environments (everything from mouse clicks and windows closed to the text they write). Whereas in the past doing this might have required producing a document with typed detailed descriptions of onscreen events combined with static pictures (screenshots) of a computer screen, users can now relatively easily record, archive and share specific moments on their screens.

To create VSC, a number of software applications are now available. While some of these are free (e.g., Jing, Screencast-O-Matic, and CamStudio), software programs which typically offer more features (e.g., editing functions) are available for purchase (e.g., Snagit and Camtasia Studio). VSC is offered as a standard function through the QuickTime software pre-loaded on Apple computers, and increasingly VSC technology is designed to work seamlessly with popular software programs such as PowerPoint and Adobe Connect. Recently, VSC applications have also been developed for mobile devices (e.g., Screen chomp, Explain Everything).

These VSC programs offer users a great deal of choice, allowing them to select the area of their screen they want to capture (full screen or a selected window only) and what they want to capture (video only, video and sound, mouse clicks, webcam, etc.). In the majority of cases, videos produced can be saved in a number of popular formats (MP4, AVI or Flash videos) with the choice of either high or low screen resolutions.

Many of these VSC programs also permit individuals to distribute online the screen capture videos they produce in the form of *screencasts* (screen capture videos distributed online). Videos can be uploaded to the Internet and then shared easily with peers by sending out a URL link to the uploaded video or by using an HTML code to embed the uploaded video in a website.

Why be interested in the use of VSC?

In the past decade, whereas the use of VSC has been popularized by software companies and creators of instruction manuals, this software has attracted the attention of various individuals who seek to take advantage of its ability to show at a distance what they are doing on their screens (Carr & Ly, 2009; Peterson, 2007). Librarians, for instance, have turned to VSC to enhance their interactions with library users seeking help with the use of library resources (Price, 2010). VSC has also been widely adopted in the gaming community as gamers show off their skills and abilities by uploading screen-captured movies of themselves completing particularly difficult sections or elements of a video game (Gow, Cairns, Colton, Miller, & Baumgarten, 2010).

In the field of language education, VSC has slowly gained popularity as both a research and an educational tool (Drumheller & Lawler, 2011; Geisler & Slattery, 2007; Jones, Georghiades, & Gunson, 2012). Indeed, VSC offers educational researchers new ways to investigate processes associated with the various outcomes and products produced by learners through LCI tasks. It appeals in particular to researchers who are interested in detailed descriptions of the mediated nature of language and literacy development in digital spaces. This is explored in greater detail in the following section.

Exploring the mediated nature of language development in digital spaces

The ability to document and investigate LCI through VSC appeals to those researchers who frame learning within a task-based approach and who draw on sociocultural theories of language development, whereby the engine for learning goes beyond the transmission of information from teachers to students. Within these frameworks, the focus rather is on learning as the result of interactional discourses (Gibbons, 2003). These discourses are generated as learners participate in language-mediated activities and tasks that allow users to interact with the language, produce it, and refine their knowledge of its conventions and rules (Duff, 2010; Lantolf & Thorne, 2006; Vygotsky, 1978).

Through its ability to document events which occur as students interact with their computers, VSC is particularly well suited to explore this mediation process. Moreover, since VSC allows one to capture what occurs in digital spaces, it enables one to address the need to explore how the migration of everyday literacy practices into digital spaces is transforming literacy development both in and out of the classroom (Lotherington & Jenson, 2011; Stapleton, 2010; Yi, 2014).

Learners' movement away from physical pen-and-paper interactions towards forms of writing in digital spaces has enriched but also rendered the act of writing and the processes associated to the writing development more complex. As new generations of students learn to read, write and interact with computers and tablets, there have been growing calls to better understand how computer-mediated tasks and interactions affect students' abilities to engage with and produce texts. What, for example, is the impact of activities such as writing out a text long hand, looking up words in a physical dictionary, and revising and editing a printed version of a draft? All are literacy practices whose traditional forms are being gradually displaced by new practices mediated through digital technologies.

Studying these changes has been identified as an important research field. For example, in their book *Digital Writing Research: Technologies, Methodologies and Ethical Issues*, McKee and DeVoss (2007) identified a number of new areas of exploration emerging from digital writing research. These include emergence of digital communities, the notion of *ethos* and the use of ethnographic practices, as a means of exploring what occurs in digital communities. Stressing how changes in the writing context have resulted in "processes and products of digital writing" which are often "different from paper-based processes and products" (p. 9), they also stress the value of research that can capture and account for the links between writing processes associated with digital texts, the activity of learning, and multimodal spaces. VSC allows the exploration of these new types of technology-mediated processes and their roles in shaping students' understandings of *literacy processes* and development.

A tracking tool well suited for usability tests

As a tracking "see-me-in-action" tool, VSC is also attractive to all who are interested in observational research and who seek to monitor users' on-screen activities (Chun, 2013; Fischer, 2007). It can be used as an alternative or in conjunction with key logging programs to produce detailed records of users' screen activities for further analysis, and VSC can be used to conduct usability tests (Van Waes, Leijten, Wengelin & Lindgren, 2012).

Usability is a concept borrowed from HCI, a property conferred to any artefacts used by humans to accomplish specific tasks. Bevan (2009) referred to usability as *quality in use*, highlighting its process-oriented nature. Usability tests (Kuniavsky, 2003; Rubin & Chisnell, 2008) are experiments, interventions typically conducted iteratively (several times and at various development stages) with (typically a small number of) representative users (selected on the basis of prior user profile analyses) invited to perform specific tasks with or involving the use

of the artefact under development. Such interventions aim to recreate "authentic" task situations (based on prior task analyses) to observe how users:

1. Behave in such situations, in particular when interacting with the artefact;
2. Are successful at completing the given tasks; and
3. Are satisfied with the artefact for the given tasks.

Hence, a core part of any usability test is observing and evaluating various aspects of the artefact under development. For example, a usability test can be used to investigate how language learners make use of an online dictionary to address linguistic issues they are experiencing when working with texts (Hamel, 2012). Such usability tests allow CALL designers to identify aspects of the dictionary that might be redesigned so as to make its relevance and benefits for users more explicit.

VSC enables one to capture and thus observe exactly what the user is doing at the computer, hence its user-centred objective nature. In running usability tests, VSC offers a practical and relatively simple way to investigate the link between various processes and the success or failures that students are able to achieve as they complete a language learning task. VSC enables researchers to associate the actions seen on-screen to explanations (e.g., gathered from questionnaires and interviews) of why *certain students produced a text in the way that he or she has.* This is an insight which is often missing in the literature on composition studies and second language writing, where much of the work is conducted with the analysis of static, predominantly final drafts produced by students (Séror, 2013). Faced with the end product of writing, researchers and instructors are left to infer the reasons behind the qualities found in a text.

With VSC, inferences made about strategies used by students when they interact with the computer can be deducted on the basis of direct behaviour observations and the degree to which these have impacted the quality of the language output that was produced. Séror (2013), for instance, highlighted how students' composition processes were linked to students' strategic use of visuo-spatial elements (Olive & Passerault, 2013; see Chapter 9, this volume) found in digital spaces and within specific software programs (i.e., the colour and size of a window, the positioning of windows, the ability to customize the fonts and margins of a page, and the use of annotation features) when interacting with a word processor as part of their work with a text.

Similarly, researchers can verify visually, for instance, the amount of time a student actually spent revising a text before handing it in. We can explore what strategies the student employed when engaging in this revision process. Finally, we can identify what specific resources the student turned to when doing this.

These are but some of the questions that can be investigated thanks to process data, such as the type collected with VSC, which, combined with other types of data elicitation methods (such as questionnaires, interviews, talk-aloud protocols and stimulus recalls), can provide researchers (and teachers) with a more complete and accurate portrait of the language learner and his or her learning trajectory.

As a result, we can obtain a nuanced understanding of what differentiates successful and less successful language learners and their results on a language learning task. We can also importantly take into account more closely the learners' backgrounds, habits and needs, leading to recommendations grounded in authentic user-based practices regarding the best digital resources and interfaces for language learners and the design of new CALL applications.

An appealing tool for educators

Educators have also begun to explore the use of VSC. As with researchers, it is the "show and tell" qualities of VSC and its ability to produce permanent records of LCI that have attracted educators seeking to produce artefacts that can document and scaffold literacy development.

In some of the earliest applications of VSC for pedagogical aims, educators have created videos to provide multimodal feedback to students. In the videos, instructors annotate, comment and modify students' texts, offering visual, audio and dynamic dimensions to their feedback designed to scaffold students' learning and enhance what have traditionally been pen-and-paper comments placed in the margins of students' papers (Jones, Georghiades & Gunson, 2012; Mathisen, 2012; Séror, 2012).

Recently, VSC has also been used to produce video clips that are shared with students to review specific pedagogical objectives and resources (e.g., providing an overview of a grammar point) (Gormely & McDermott, 2011). This ability for educators to produce short video clips that students can watch at home is at the heart of an increasingly popular concept of *flipping the classroom* by providing information and teaching opportunities outside of the classroom so that more time can be spent in the class working on applications of the knowledge distributed to students (Khan, 2011; Toppo, 2011).

As suggested above, VSC represents an innovative tool that can be used to explore LCI and its relationship to literacy dimensions and the design of CALL tools promoting language development. Its focus on user interactions and the ability to create digital traces of language learners' actions lend themselves to usability studies which are well suited for studies of computer-mediated literacy processes (Degenhardt, 2006; Geisler & Slattery, 2007; Park & Kinginger, 2010).

In the next section, we look at three examples of how VSC has been used to explore the impact of LCI in the design of *online dictionaries* (Hamel, 2012), the development of *writing processes* (Hayes & Flower, 1980; Séror, 2013), *metacognitive awareness* (Hacker, Keener, & Kircher, 2009) and *learner autonomy* (Benson, 2001; Dion, 2011; Little, 2007).

Following brief descriptions of each project, we will seek to draw out the key lesson learned from the projects, focusing on the recommendations that have emerged from our use of VSC as a means of researching and enhancing LCI tasks.

Description of the research projects

Study 1: Investigations of *learner-task-dictionary* interaction

As part of a CALL research and development (R&D) project, Hamel (2012, 2013) employed VSC to conduct a series of usability tests on an online dictionary during its prototyping phases.

The VSC tool *Camtasia* was used to document and observe on-screen the *learner-task-dictionary* interaction. The aim was to measure the quality of this interaction, i.e., its usability, for the purpose of improving the design of an online dictionary (its interface and content).

Adopting an ergonomic approach to CALL design research (see Chapter 2, this volume), i.e., a learner-centred approach, Hamel drew on the concepts of usability (see above) and tools and techniques employed in the web engineering and interface design industry to measure the *quality in use* (Bevan, 2009). Usability tests were employed in this research as an elicitation method in order to get LCI data that would inform the design of her online dictionary.

These concepts were integrated with the use of VSC technology to facilitate the direct observations and process-oriented analyses of students' interactions with the online dictionary prototype. Language tasks were created which optimized conditions for the dictionary to be solicited during their completion process (Hamel, 2012). These were semi-authentic, corpus-driven micro-tasks for which learners had to translate, revise, construct or reformulate identified collocations, in sentence and text-wide contexts. VSC was crucial in capturing this LCI and students' solicitations of the online dictionary functions to engage in the process of constructing collocations.

The process and product-oriented LCI data collected through this study were used to directly inform both measures of efficiency and the effectiveness of the dictionary being studied. A set of parameters, based on visible on-screen actions, was devised to measure *efficiency* focusing on efforts and time at task while on

language accuracy to measure *effectiveness*. Pre- and post-task questionnaires given to the students also allowed the researchers to collect more indirect measures of learner background, experience and satisfaction with the tool. The results were triangulated, and some correlations were found between reported and observed behaviours (Hamel, 2013), namely between prior exposure to a variety of online resources and success.

Study 2: Investigating L2 writers' composition processes and strategies

Séror's (2012, 2013) research drew on the use of VSC to document and investigate undergraduate university students' composition processes and strategies as they completed authentic writing assignments in their second language. Inspired by the need for more detailed representations of the moment-to-moment actions, decisions and composition processes enacted by L2 students as they learned to write for university classes and ultimately to master the complex series of processes associated with the production of academic texts (e.g., Roca de Larios, Manchón, Murphy, & Marin, 2008; Sasaki, 2000; Victori, 1999), participants were equipped with the VSC tool Screencast-O-Matic (SOM) <http://screencast-o-mastic.com/>. Participants were instructed to record whenever they composed and completed assignments in their writing classes on their own computers. These recordings were largely conducted outside of the classroom and provided rare insights into the writing processes that underlie L2 writers' production of academic texts in authentic settings outside of the classroom.

Created unobtrusively as writers composed and completed assigned writing tasks on computers, these records were analysed in conjunction with retrospective interviews conducted to explore students' specific composition strategies, individual performances and their perspectives and justifications of the various behaviours observed in the recordings of their writing sessions.

Data-analysis procedures for the study triangulated both the video records and student interviews with a research log, field notes, and informal conversations with focal students and their instructors. A quantitative analysis of the sequences of events found in the visual records of students' composition processes was juxtaposed with a qualitative analysis of students' own perspectives of the composition processes and strategies underlying their writing.

Drawing on the work of Park and Kinginger (2010), each recording was coded for *transactions*, instances which expressed an immediate need on the part of the writer and his or her efforts to respond to a problem as identified through a series of visual signals in the screen recordings (for example, a pause, followed by the deletion of a word and the insertion of a new word, followed by another pause before continuing to write another sentence).

Study 3: Exploring the pedagogical pertinence and added value of the integration of VSC

Hamel, Séror and Dion (2015) collaborated in an on-going research project focusing on the pedagogical pertinence and added value of the integration of VSC in the second language (L2) writing class. Built on prior investigations of the digital traces of language learners using computers (Degenhardt, 2006; Geyser & Slattery, 2007; Hamel, 2012; Hamel & Caws, 2010; Park & Kinginger, 2010; Séror, 2013), the study's objective was to investigate how second language writing instructors might integrate VSC in their classroom activities and tasks to scaffold learners' writing development and design more effective, better suited and more personalized pedagogical interventions.

By means of case studies in two university L2 writing classrooms (N = 36), the research focused on the innovative practices linked to the adoption of VSC by two experienced second language writing teachers over the course of a semester. A key objective was to document these teachers' use of VSC for pedagogical purposes as well as to document the process and product of writing tasks by students as they completed these tasks, both in authentic classroom settings and outside the classroom as part of homework activities. Screencast-O-Matic was used as the VSC tool. A corpus of 200 screen recorded videos was collected and analysed (quantitatively) based on a taxonomy of functional and cognitive parameters devised from visible and audible (inter)actions identified in the videos.

In addition to the VSC recordings produced by students and instructors, classroom artefacts (e.g., task descriptions, journal entries), student questionnaires and teacher interviews were analysed to explore how the tool was used and adopted by instructors in these courses, its impact on the quality of the work produced by students and the perspectives expressed by both students and instructors as they reflected on the value of this tool for their language development.

Affordances and opportunities associated with VSC

Drawing from our own experience as practicing researchers with VSC in CALL and second language literacy development, we believe it is possible to identify a number of interesting affordances (see Chapter 3, this volume) associated with VSC. We will illustrate these below in an attempt to provide practitioners and researchers in CALL with *strong clues* about how VSC could be *fruitfully employed in design* (Norman, 1998) in *meaningful* ways (Gibson, 1997).

A tool that is accessible and easy to use

All of the above mentioned studies revealed that VSC as a tool was, on the whole, easily accessible and easy to learn to use by the researchers, language learners and language instructors. This quality is well illustrated in the research projects 2 and 3. In both cases, Screencast-O-Matic (SOM) was specifically chosen for its reliability and ease of use. SOM is a free Web-based application that does not require any specific software to be installed on the computer used for the recording. This made it a highly accessible resource for both students and their instructors and it meant that VSC could be produced in a variety of settings (recordings could be created when students were working in a lab, on a home computer or even when working on a library computer). The free version of the program allows individuals to create screen recordings of a maximum of fifteen minutes. A professional version available for a monthly fee was used in the second study and allowed participants to record their screens for as long as they wished. Once a recording is complete, users can easily save this video on a hard drive and/or upload it to a server, which can then be used to share links with other students in the class or with their instructors. Training individuals to use VSC tools has also proven, in our experience, to be relatively simple. Tools, such as Camtasia and SOM, essentially reproduce the near universal *record*, *play* and *rewind* interface found on both analogue and digital video and sound recorders.

Study 3, for instance, involved training instructors and students to use SOM through a series of workshops focused on research and teaching practice. Among attendees were the two teachers who volunteered for the project. In addition to providing training to the teachers interested in using SOM in their classroom, at the start of the semester (week 2), a researcher visited both of the instructors' writing classes and offered hands-on demonstrations of the use of SOM to their students. This demonstration helped familiarize students with the tool and also allowed the instructors to explain how the tool would be used to complete a number of the writing tasks that would be assigned to the students over the course of the semester. Students and their instructors were also provided with support material (*How to use SOM in 12 easy steps*) created by the researchers to allow students to review, at home and later on in the semester, the various steps involved in the use of SOM. The email address of a research assistant was also distributed to students and the instructors. This research assistant was presented as a resource that student and teachers could contact to ask questions and to troubleshoot any problems. As with the other studies we conducted, there were few user-related problems. The main issue which emerged were difficulties experienced by students who had to install/update their web browser's Java prior to being able to run the SOM application.

A powerful documentation tool

VSC in the three projects described above was a powerful documentation tool (Fischer, 2007), producing rich empirical records of observables generated in real-time, in both controlled as well as naturalistic settings. We stress again here that the data collected through VSC is presented in the form of screen recorded videos which include sound, image, movement and a full range of colours and modes that are essential aspects of the language learning experience in digital spaces. Researchers who use VSC can thus benefit from seeing all visible on-screen actions done by learners as they interact with texts and engage in textual meaning making.

This data is made even richer if students have opted to use their computer webcam and microphone to capture their voices and images as they engaged in LCI. This occurred, for instance, in the case of study 3 when students engaged in a writing task and chose to reflect on it in this way. Figure 7.1 shows an example of one student who chose to activate the webcam when video screen capturing her text revision process. In this extract, she is verifying a grammatical rule (about the use of gerunds in French) in a printed resource, reading it aloud and making a hypothesis about whether it applies to the text segment she has identified as problematic.

Such data is clearly important when looking at complex processes such as literacy practices, including the strategies employed by students and their underlying cognitive processes.

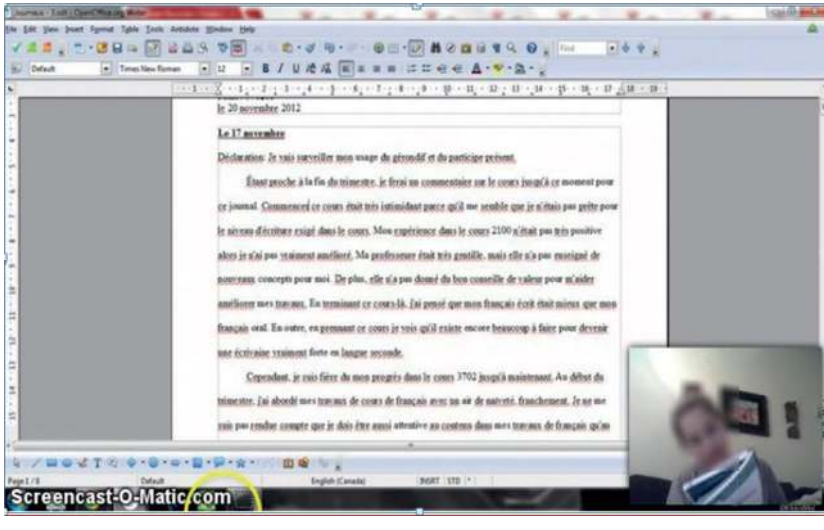


Figure 7.1 Student using VSC with webcam to document her revision process

In the case of writing, the second study produced recordings of students' writing sessions, which captured all of the activities linked to the realization of a written assignment from the first word to the last one. This allowed us to witness the multiplicity of decisions made as students moved from their original outlines, to a first draft and, finally, to a submitted text with all the micro actions that came in between (e.g., looking up a word in an online dictionary or struggling to produce a French accent). For the third study, students recorded up to fifteen minutes of their writing process (each video clip lasted twelve minutes, on average). In that short period of time, a high density of actions, both visible and audible, were observed – on average, 85 per video clip, which showed students being well invested in their writing task while revealing several types of strategies, such as focalization on form, hypothesis making, text repair, or drawing on prior knowledge.

VSC produces rich data that can be analysed in multiple ways

In all of the research projects mentioned above, it should be noted that the analysis of the data was facilitated by the use of Morae (techsmith.com), a specialized usability testing software program designed for the (distance) observation, capture, management, annotation and qualitative and quantitative analysis of VSC videos.

This program facilitates the insertion of annotations (tags) in the visual records produced by VSC. Markers and codes can be predefined and then attached as tags to the videos. These help identify parameters that can later be compiled and explored for general statistical trends within the program itself or through other programs by exporting the data into Excel files, for instance, for subsequent/further statistical analysis. Figure 7.2 shows a screenshot of Morae used to conduct usability tests with dictionaries in study 1.

Annotations added to videos with Morae can then be used to recreate timelines of events and to provide statistics regarding the quantity and general tendencies associated to key events in the data. One can calculate, for instance, the following:

1. The presence and durations of pauses taken by a student when composing;
2. The average duration of a lexical search, how students start a word search, which key word(s) they use and even which dictionary rubrics they look-up;
3. Instances when students engage in various steps of the writing process (producing text, vs. editing, vs. planning); and
4. The number of times students access online resources during their writing process and the type of resources they access (dictionaries, conjugators, translators, etc.).

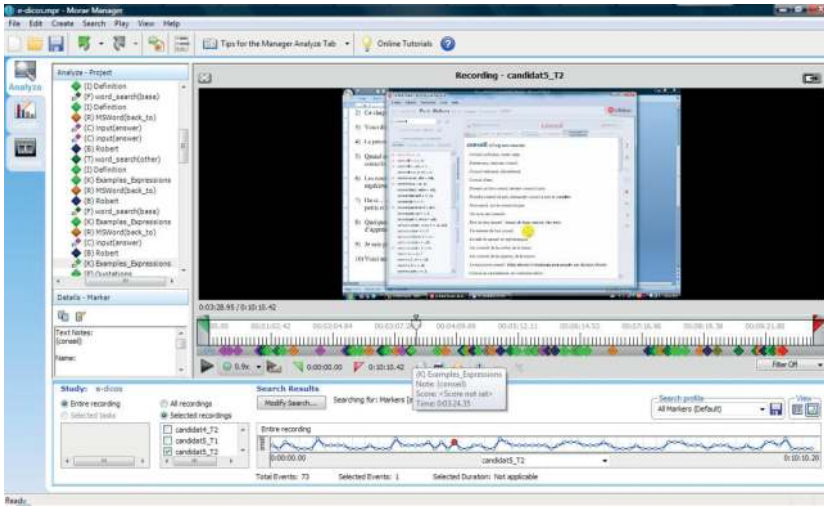


Figure 7.2 Morea, a usability test management software

Notes can also be added to the data with Morae at various points, allowing us to insert analytical memos and links to external data sources.

A tool like Morae facilitates the analysis of VSC. It does not impose on researchers a theoretical perspective or approach. Both grounded data-driven and theory-driven approaches can be applied to the nature of the data collected through the use of VSC. The approach adopted will depend on the researchers' epistemological orientation, research questions, methodological design, theoretical perspectives and pedagogical goals. In this sense, VSC remains flexible and can be used for various purposes (researching learners' information searches, researching writing processes, researching pedagogical reflective tasks, looking at peer editing, etc.).

In the case of the three studies which are the focus of this chapter, the following elements illustrate the types of analytical lens through which the VSC data collected was analysed.

Thanks to the annotation functions of Morae described above, it was possible to produce detailed timelines of task processes present in the VSC data. These timelines allowed the researchers to identify steps involved as students revised their texts. This process included selecting specific text segments, attempting to repair these segments, searching in online resources such as dictionaries, justifying in some cases the decisions made, etc.

Much like the detailed transcripts produced in conversation analysis research, these timelines offer valuable insights regarding the sequencing of events which

underlie specific events of interest and can provide hints regarding the important interplay between these events.

Guided by the question "why this now?" for instance, data emerging from study 2 helped identify moments when students turned to Internet-based sources, identifying both larger patterns of behaviours (e.g., greater use of external resources at the end of the writing session as students revised their texts), as well as unique moments linked to specific events and strategies (e.g., students' preference for specific dictionaries linked to their desire to work in their L1 or L2, depending on the lexical item they were looking up).

Similarly, LCI data from study 1 provided valuable task path sequences (i.e., navigation paths) of learners' interactions with the online dictionary when attempting to construct, reformulate or translate collocations. Hamel (2012, 2013) observed that weak learners tended to "waste" time at the beginning of their search for lexical information, hesitating about which keywords to input. Several learners looked for examples before they looked for meaning (definitions). In a series of synonyms provided in the dictionary, most learners selected high-frequents and L1 cognates over more idiomatic equivalents.

Figure 7.3 shows a 35-second task path sequence from study 1 of a participant searching for a synonym of the collocate "grande" (great) in the dictionary, starting from his search with the keyword of the collocation "joie" (joy) and finding "inépuisable" (endless) as a possible equivalent.

In study 3, using the same timeline approach, thanks to the parameters annotated in real-time in the video, it was possible to identify and reconstruct an attempt by a learner to repair a collocation as he also reflected on this task. During

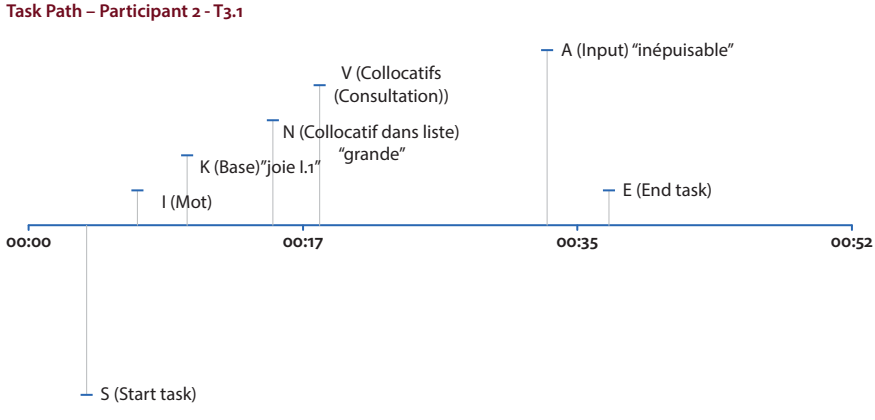


Figure 7.3 Task path sequence of a participant searching for a collocate in a dictionary

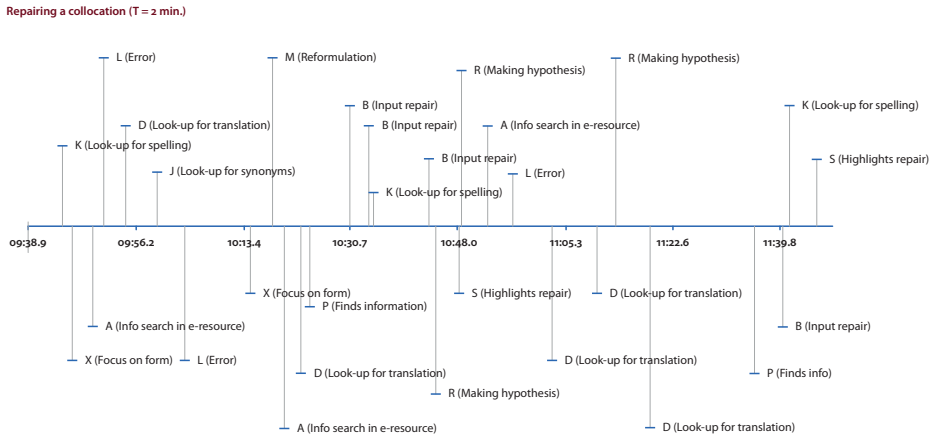


Figure 7.4 Action sequence showing an attempt by a learner to repair a collocation

this two-minute process, 30 actions were recorded. Figure 7.4 shows this action sequence.

Lifting the veil on hidden processes

The detailed step-by-step records provided by the VSC data helped bring to light aspects of students' literacy practices which have in the past traditionally remained invisible and thus unnoticed and/unverifiable in the absence of VSC (Geisler & Slattery, 2007). This ability for VSC to lift the veil on students' processes represents a key affordance of this tool.

In the case of study 1, for instance, one could see how language learners navigated their way to the various choices made, more or less efficiently, in electronic dictionaries. As they moved from one micro-task to another, some students learned to optimize their search paths in the dictionary whereas others did not.

Similarly, in study 2, it was interesting to note the role that students' L1 actually played in their L2 composition processes. Whereas these students' final drafts were, by the very nature of the task, completely written in French, VSC data allowed one to note how often writing in the L1 had in fact helped scaffold this L2 writing (e.g., a student wrote her first draft of her text in English before translating it into French).

This type of data (and the insights that can be generated from its analysis) is well suited for ethnographic studies of digital literacies that highlight the value of the direct observation of students' literacy practices and LCI. It also makes important contributions to the field of CALL ergonomics (see Chapter 2, this volume) by allowing focus on the quality of the user-task-tool interactions at the computer, on the mediations with the task and the tools and on the various choices, paths (optimal, efficient, etc.) students make and take as they use tools for L2 writing.

VSC as a means of exploring efficiency, effectiveness and user attitudes

Another advantage stemming from the detailed maps and portraits offered by VSC is that one can focus on the efficiency, effectiveness and user satisfaction experienced by users as they interact with texts in a digital environment. These criteria reflect those standardly used for usability tests to measure the quality of user-task-tool interactions (see Chapter 2, this volume). One can look at efficiency as a measure of efforts (calculated as a function of actions taken over a defined time period). For example, in study 1, as detailed above, parameters of efficiency

were included as coding annotations for the VSC data collected. These parameters coded the degree of efforts expanded by students when dictionary functions (such as performing a keyword search, looking up a synonym) were solicited by a learner during the task process.

Investigations of students' efficiency can also be used to explore the notion of *errors* made by users as a result of their task-tool interaction. In the case of study 1, for instance, such errors occurred in the LCI corpora, emerging from the usability tests with the online dictionary. They highlighted problems related to its interface *accessibility* (difficulties finding/using a function of the dictionary) and its content *comprehensibility* (difficulties understanding information provided by the dictionary, such as definitions), both having negative effects on its *learnability* (difficulties learning how to use the dictionary).

Errors, moments of struggles or transactions as students worked through the problem solving nature of composing their texts also emerged in study 2. In this case, these errors helped identify developmental aspects which need to receive particular attention in the design of writing pedagogy and the conceptualization of what students need to learn and the skills they need to develop in order to become good writers (e.g., many students need to be taught explicitly how to produce French accents on their keyboards or strategies for the effective use of grammar and spellcheck software, such as *Antidote*).

One can also look at effectiveness and the degree of user satisfaction/contentment associated with specific actions taken with a specific tool or achieved through the use of a specific strategy.

This can involve direct measures of effectiveness through objective measures of what can be seen on the screen (e.g., on-screen actions, task results). Hamel (2012) measured the quality and the quantity of the language output produced by language learners as they interacted with their dictionaries. A successful language output corresponded to an accurately constructed collocation, produced by a learner as a task outcome. Study 2 explored whether a student found an accurate way of expressing an idea after a moment of struggle in her writing was signalled by both greater than average pauses in her writing process and the interruption of text production to look up linguistic information in an online resource.

Effectiveness can also be investigated through users' self-reports (e.g., answers to questionnaires, interviews), data elicited to ask users to judge/comment on the degree to which they believe they have or have not been successful at achieving specific goals. This illustrates how data produced through VSC can also be analysed in conjunction with additional data sources to add to the richness of the accounts produced with VSC data.

In study 1, questionnaires were used to investigate learners' perceptions of the online dictionary used for the completion of a collocation task (Hamel, 2013). These questionnaires also investigated the feelings and judgments attributed by a student to the results obtained at the end of a task (for example, a high or low score, a positive or a negative comment given to a dictionary feature or a comment on the value of a personal performance). The questionnaires revealed that learners tended to underestimate their performance, found the task difficult and the dictionary essential for its completion – this, despite a high effectiveness score obtained by the majority. These findings corroborate Fischer's (2007) claim that there is often a discrepancy between observed and reported learner data.

Study 2 used interviews to ask learners to comment on various aspects of their writing. Questions, amongst others, explored students' degree of satisfaction with the texts they had produced and students' perceptions of the usefulness of the various resources consulted while writing. It was also possible to match VSC with stimulated recall interviews (Gass & Mackey, 2000). Students were asked to revisit and watch selected excerpts of their screen capture videos and to narrate their task processes and explain what was going through their minds as they engaged in the actions captured in the video. With this type of approach, the general goal is to elicit information about users' perception of a task, its realization and their feelings about what they are capable of achieving (Raby, 2005).

In study 3, both questionnaires and semi-formal interviews were used to investigate students' and instructors' perception as to the use of VSC as a pedagogical tool in their writing classrooms. It is also possible, as was done in the case of study 3, to ask students to engage in a reflective task (see more about this below), where students are asked to watch themselves again, reread the texts they have composed and comment in writing or, through the creation of a new VSC, on what they have noticed and learned about themselves as (L2) writers.

Additionally, students can be asked to record their thoughts at the same time as they complete a task and work on a computer (another type of task which emerged in study 3). This activity asks students to comment in a think-aloud fashion, which provides an aural track of the thoughts and ideas that accompany what occurs on screen. In one instance, this task was assigned to students who were working as a group, generating extremely interesting data on the types of interactions and discourse produced by students as they engaged in a text-planning activity.

Pre- and post-task questionnaires can also serve to gather information about the participants as well as to collect information about data, other than students' sense of satisfaction with the LCI tasks they have been asked to record with VSC. Demographic questions, questions about students' previous educational experiences and technology usage, as well as questions about their attitudes towards

their second language, are but a few examples of variables that can then be used to explore possible correlations with events and actions noted in the VSC recordings. As mentioned above, study 1, for example, showed that there was a strong correlation between task success and learners' prior experiences with a variety of other lexical resources.

The final element that can be used to help contextualize the events and actions seen in the VSC recordings includes the collection of any relevant textual materials connected to the digital texts captured through the VSC (e.g., copies of handwritten notes students use as they work on the computer, copies of the task descriptions handed out by instructors) and observational data in the form of field notes. In the case of studies 2 and 3, field notes and reports were kept as well as copies of course outlines, assignment descriptions and handwritten notes provided by participants in the study. This material provides valuable hints which can enhance the analysis of students' actions and can be triangulated with the various data sources mentioned above to produce detailed accounts and establish the relationships between learners' on-screen actions, their attitudes and backgrounds, as well as the context and resources associated with the specific literacy practices and CALL tools and applications being studied. This is well in line with an ergonomic approach to the analysis of LCI (see Chapter 2, this volume).

VSC and the ability to create new forms of corpora

Our research experiences with VSC and the richness of the data these projects produced suggest that there is great potential in VSC's capacity to produce rich audio-visual corpora of language learners and their educators as they engage in LCI tasks (i.e., composing a text, using an online dictionary, reflecting on a text, providing feedback to students). Indeed, a key affordance of the tool lies in the fact that while recordings produced by VSC can be analysed individually, these can also be collated and compiled to produce multimodal LCI corpora that allow for the cross-case analysis of individuals' literacy practices in digital spaces.

Such corpora represent new and exciting forms of empirical data which, once anonymized, could contribute to learner corpus projects that might be shared with others (see Chapter 10, this volume). The resulting database of observable processes could then be exploited to better capture the fluid and ever-evolving nature of literacy practices. It could be used in teaching interventions as well as for teacher training. Similarly, the corpus could become a source of valuable materials to be integrated into presentations, webinars and online tutorials and meetings.

VSC as a retrospection tool

Another key affordance emerging from our studies is the tool's potential to serve as an aide to retrospection (see Chapter 9, this volume). By allowing users to capture and later replay their interactions with the machine as they engaged in a task, VSC allows users to view these interactions in a more detached and reflective way than what is possible at the time one is actually completing the task. In this sense, the VSC recordings were labelled by one of the instructors in study 3 as a tool that can serve as a mirror offering new ways to view and understand their own behaviours and literacy practices. This instructor took full advantage of this affordance and encouraged her students to revisit their notes and texts, but also the VSC recordings which they had produced over the course of the semester, when studying for her course.

Within both research contexts as well as within the context of a classroom, VSC facilitates language learners' metacognitive awareness and strategic awareness. This retrospection affordance benefits the students/participants as well as the teachers/researchers who gain insights into students' developing knowledge and skills as they gain experience with a targeted LCI task.

VSC as an important and powerful scaffolding tool

A key affordance of this tool emerged from the work of Hamel, Séror and Dion (2015), focusing on the tool's ability to scaffold learners and enhance CALL pedagogy. The findings from this project highlighted the multiple and varied ways in which VSC can be integrated into the language classroom through the (re)design of L2 writing tasks.

At the start of the project, when discussing its aim and the potential applications of the use VSC has in L2 writing classrooms with instructors, time was spent brainstorming what a VSC-mediated L2 writing task might look like. This process took into consideration notions of syllabus design, course objectives and the nature of writing tasks previously assigned to students by all instructors present. Ultimately, the two instructors who participated in the study produced a number of tasks which integrated the use of VSC and responded to their personal needs and teaching styles.

The FLS instructor favoured VSC-mediated tasks that focused predominantly on text revision, aspects of the text genre and the desire to develop students' text agency. These tasks were designed to be completed as individual homework assignments in students' homes. The students' roles were to revise and assess their writing, reflect on revision and develop an awareness of themselves as writers.

For her part, this instructor saw her main role as being an *assessor* who provided feedback (evaluation, comments) on the text process and product as well as on the degree of agency and metacognition observed in students' reflections.

The tasks designed by the ESL instructor targeted specific components of the writing process, with objectives focusing on helping students experience and master subcomponents of the writing process, such as brainstorming, text planning and editing errors when revising. In contrast to the FSL instructor's tasks, his tasks were designed to occur within the classroom environment and took advantage of the fact that some of his classes were offered in a lab, equipped with computer stations. He favoured positioning the students in the role of *thinkers* when using SOM. Individual feedback focusing on the VSC recorded by students was not provided directly to them. However, the recordings were discussed in the class as a whole, although the actual videos were not shared or viewed by peers. Rather, students were encouraged to watch the videos on their own to help them reflect on their writing processes.

The ESL instructor further reinforced his focus on the writing process through the use of modelling. He presented the students with relevant text samples of his expected written outcomes and engaged students in peer work and editing so that stronger students might help weaker students by modelling optimal processes and strategies. Interestingly, he extended this modelling practice by asking an expert writer to produce a VSC, which could be shown to students as an example of how advanced writers complete writing tasks. This video clip served as an authentic, multimodal exemplar for the students, helping to reinforce the validity of the steps and processes promoted in his writing class.

In interviews discussing their experiences with VSC, both instructors noted that they had found VSC useful for monitoring, supporting and accompanying language learners as they worked independently through the various stages of writing associated with a specific task. Importantly, both instructors also identified the potential of building a database of their own students' VSC with the option (granted consent from students) of exploiting this small corpus for pedagogical purposes. Video extracts (e.g., action sequences, as seen above) might be chosen to illustrate best practice, common problems experienced by students and their solutions or to share with others resources that fellow language learners have identified and successfully used.

Instructors also commented on the ability to communicate with students in a multimodal medium that can be delivered outside the traditional context of the classroom. In their opinion, while integrating VSC into their classrooms did require transforming their teaching practices and a significant investment of time and energy to redesign writing tasks they had used in the past, VSC offered new and exciting ways of achieving the class objectives.

Whether it is for providing feedback, demonstrating the most effective way of looking up collocations in a dictionary or to illustrate in concrete and dynamic ways how students write, VSC offers a range of applications which can be used to support L2 literacy development. Such videos would be particularly useful in the context of courses offered online and in blended environments.

Conclusion

This chapter has illustrated the affordances of VSC and its role in the design of CALL research and pedagogy, drawing on examples of the use of VSC and its applications in three research projects.

Its affordances present a number of promising avenues to be further explored as researchers continue to discover ways to take advantage of the tools' documentation function as well as its dynamic, multimodal nature.

Our research has highlighted the potential of VSC for the investigation of computers and digital spaces, particularly the role it plays as a mediating tool which increasingly shapes the literacy development and experiences of users. Undoubtedly, these affordances play a role in helping shape what it will mean to teach digital literacies and to promote the competencies required of students, citizens of a modern, technologically connected world (Yi, 2014). Further work is needed to explore and document the full range of applications of VSC for research and pedagogy. As Levy (2013) has reminded us, a design-based CALL agenda should explore *usability, scalability* and *sustainability*. Hence, the integration of VSC should be carefully planned and scaffolded with training, and embedded in feedback. Creative and collaborative usage (sociocultural mediation), the development of communities of practice of teachers and learners experienced in using VSC, as well as technology experts, will represent important ways of further refining our understanding of the affordances of this exciting technology.

References

Barbier, M. L., & Spinelli-Jullien, N. (2009). On-line tools for investigating writing strategies in L2. *German as a Foreign Language*, 2(3), 23–40.

Benson, P. (2001). *Teaching and researching autonomy in language learning*. London, United Kingdom: Longman.

Bevan, N. (2009). Extending quality in use to provide a framework for usability measurement. In M. Kurosu (Ed.), *Human Centered Design* (pp. 13–22). Proceedings of HCI International 2009, San Diego, California, USA. Berlin: Springer. doi:10.1007/978-3-642-02806-9_2

Carr, A., & Ly, P. (2009). "More than words": Screencasting as a reference tool. *Reference Services Review*, 37(4), 408–420. doi:10.1108/00907320911007010

Chun, D. (2013). Contributions of tracking user behaviour to SLA research. In P. Hubbard, M. Schulze, & B. Smith (Eds.), *Learner-computer interaction in language education. A festschrift in honor of Robert Fischer* (pp. 256–262). San Marcos, TX: Computer Assisted Language Instruction Consortium.

Degenhardt, M. (2006). CAMTASIA and CATMOVIE: Two digital tools for observing, documenting and analysing writing processes of university students. In L. Van Waes, M. Leijten, & C. Neuwirth (Eds.), *Writing and digital media* (pp. 180–186). Leiden, Netherlands: Brill.

Dion, C. (2011). Tools to enhance second language writing autonomy: Can we do things better? In D. Gardner (Ed.), *Fostering autonomy in language learning* (pp. 64–76). Gaziantep, Turkey: Zirve University. Retrieved from <http://ilac2010.zirve.edu.tr/>

Drumheller, K., & Lawler, G. (2011). Capture their attention: Capturing lessons using screen capture software. *College Teaching*, 59(2), 93–93. doi:10.1080/87567550903252793

Duff, P. A. (2010). Language socialization. In N. Hornberger (Ed.), *Sociolinguistics and language education* (pp. 427–454). Bristol, United Kingdom: Multilingual Matters.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behaviour in CALL. *Computer Assisted Language Learning*, 20(5), 409–442. doi:10.1080/09588220701746013

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Geisler, C., & Slattery, S. (2007). Capturing the activity of digital writing: Using, analysing, and supplementing video screen capture. In H. A. McKee & D. N. DeVoss (Eds.), *Digital writing research: Technologies, methodologies, and ethical issues* (pp. 185–200). Cresskill, NJ: Hampton Press.

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, 37(2), 247–273. doi:10.2307/3588504

Gibson, J. (1997). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gormely, K. A. Y., & McDermott, P. (2011). Do you jing? How screencasting can enrich classroom teaching and learning. *Language and Literacy Spectrum*, 21, 12–20.

Gow, J., Cairns, P., Colton, S., Miller, P., & Baumgarten, R. (2010). Capturing player with postgame commentaries. *CGAT Conference 2010*. Retrieved from <http://ccg.doc.gold.ac.uk/papers/gow_cgat10.pdf>

Hacker D. J., Keener M. C., Kircher J. C. (2009). Writing is applied metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 154–172). New York, NY: Routledge.

Hamel, M.-J. (2013). Questionnaires to inform user tests in CALL. *International Journal of Computer Assisted Language Learning and Teaching*, 3(3), 56–76. doi:10.4018/ijcallt.2013070104

Hamel, M.-J. (2012). Testing aspects of the usability of an online learner dictionary prototype: A product and process-oriented study. *Computer Assisted Language Learning*, 25(4), 339–365. doi:10.1080/09588221.2011.591805

Hamel, M.-J., & Caws, C. (2010). Usability tests in call development: Pilot studies in the context of the Dire autrement and Francotoile projects. *CALICO Journal*, 27(3), 491–504. doi:10.1558/cj.v27i3.491-504

Hamel, M. J., Séror, J., & Dion, C. (2015). *Writers in action! Modelling and scaffolding second-language learners' writing process*. Higher Education Quality Council of Ontario. Retrieved from <http://www.heqco.ca/SiteCollectionDocuments/Writers_in_Action_ENG.pdf>

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jones, N., Georghiades, P., & Gunson, J. (2012). Student feedback via screen capture digital video: Stimulating student's modified action. *Higher Education*, 64(5), 593–607. doi:10.1007/s10734-012-9514-7

Khan, S. (2011). Let's use video to reinvent education. [Video file] Retrieved from <http://www.ted.com/talks/salman_khan_let_s_use_video_to_reinvent_education>

Kuniavsky, M. (2003). *Observing the user experience. A practioner's guide to user research*. San Francisco, CA: Morgan Kaufmann.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, United Kingdom: Oxford University Press.

Lea, M. R. (2013). Reclaiming literacies: competing textual practices in a digital higher education. *Teaching in Higher Education*, 18(1), 106–118. doi:10.1080/13562517.2012.756465

Levy, M. (2013). Design-based research and the quest for normalization in CALL. In J. Rodriguez & C. Pardo-Ballester (Eds.), *Design-based research in CALL* (Vol. 11, pp. 31–40). CALICO Monograph Series. San Marcos, TX: Computer Assisted Language Instruction Consortium.

Little, D. (2007) Language learner autonomy: Some fundamental considerations revisited. *Innovation in Language Learning and Teaching*, 1(1), 14–29. doi:10.2167/illt040.0

Lotherington, H., & Jenson, J. (2011). Teaching multimodal and digital literacy in L2 settings: New literacies, new basics, new pedagogies. *Annual Review of Applied Linguistics*, 31, 226–246. doi:10.1017/S0267190511000110

Mathisen, P. (2012). Video feedback in higher education: A contribution to improving the quality of written feedback. *Nordic Journal of Digital Literacy*, 7(2), 97–116.

Norman, D. A. (1998). *The psychology of everyday things*. New York, NY: Basic Books.

Olive, T., & Passerault, J. M. (2013). The visuospatial dimension of writing. *Written Communication*, 29(3), 326–344. doi:10.1177/0741088312451111

Park, K., & Kinginger, C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning & Technology*, 14(3), 31–50. Retrieved from <http://llt.msu.edu/issues/october2010/parkkinginger.pdf>

Peterson, E. (2007). Incorporating screencasts in online teaching. *The International Review of Research in Open and Distance Learning*, 8(3), 1–4.

Price, J. B. (2010). Screencasting on a shoestring: Using jing. *The Reference Librarian*, 51(3), 237–244. doi:10.1080/02763871003792030

Raby, F. (2005). A user-centered ergonomic approach to CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL Research Perspectives* (pp. 179–190). New York, NY: Lawrence Erlbaum Associates.

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, 17(1), 30–47. doi:10.1016/j.jslw.2007.08.005

Rubin, J., & Chisnell, D. E. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Indianapolis, IN: Wiley.

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3), 259–291. doi:10.1016/S1060-3743(00)00028-X

Séror, J. (2012). Show me! Enhanced feedback through screencasting technology. *TESL Canada Journal*, 30(1), 104–116.

Séror, J. (2013). Screen capture technology: A digital window into students' writing processes/Technologie de capture d'écran: une fenêtre numérique sur le processus d'écriture des étudiants. *Canadian Journal of Learning and Technology*, 39(3), 1–16.

Stapleton, P. (2010). Writing in an electronic age: A case study of L2 composing processes. *Journal of English for Academic Purposes*, 9(4), 295–307. doi:10.1016/j.jeap.2010.10.002

Toppo, G. (2011, October 7). 'Flipped' classrooms take advantage of technology. *USA TODAY*.

Van Waes, L., Leijten, M., Wengelin, Å., & Lindgren, E. (2012). Logging tools to study digital writing processes. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 507–533). New York, NY: Taylor & Francis.

Victori, M. (1999). An analysis of writing knowledge in EFL composing: A case study of two effective and two less effective writers. *System*, 27(4), 537–555. doi:10.1016/S0346-251X(99)00049-4

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Yi, Y. (2014). Possibilities and challenges of multimodal literacy practices in teaching and learning English as an additional language. *Language and Linguistics Compass*, 8(4), 158–169. doi:10.1111/lnc3.12076

CHAPTER 8

Using eye-tracking technology to explore online learner interactions

Ursula Stickler, Bryan Smith and Lijing Shi
The Open University, UK / Arizona State University, USA /
London School of Economics and Political Sciences, UK

This chapter sets out to introduce the use of eye-tracking to investigate language-learner computer interaction. By recording the gaze focus of a computer user engaged in an on-screen task, eye-tracking aims to provide information on cognitive processes. This allows the researcher to speculate about what learners are thinking while engaged in, for example, synchronous online language learning. After briefly presenting the history and different fields of eye-tracking research, the authors present two recent eye-tracking studies in *SCMC* (*Synchronous Computer-Mediated Communication*). The potentials and challenges of eye-tracking for researching language learning are discussed, as well as the methodological options of quantitative and mixed method studies. The last section, conclusions, encourages novice researchers to carry out their own eye-tracking projects, reflecting on methodological, practical and pragmatic issues.

Keywords: eye-tracking, research method, online language learning, stimulated recall, noticing, feedback

Introduction

To arrive at a real picture of learners' interactions with computers, they need to be studied from different perspectives, taking into account the different modalities used. Limiting too early what we are investigating can lead to a loss of information. For example, by studying only output data of chat logs, any information on self-corrections, hesitations, and other learner actions prior to sending off their chat contributions can be lost (Smith, 2008). By focusing on just the screen, the mouse and the keyboard, we can miss out on all the different scaffolds and support tools that learners use, even if their main focus is interaction via a computer

DOI 10.1075/lsse.2.08sti

(Suzuki, 2013). Eye-tracking is a useful method to gather data on users engaged in learning with a computer, adding another dimension to the picture that cannot be easily provided by alternative methods, such as video screen capture. This chapter will consider reasons for choosing this technique rather than an alternative, and it will consider how this choice is linked to the underlying research questions.

The previous chapter(s) have shown that capturing what learners do while engaged with a computer screen by using screen recordings or video capture can provide a plethora of information. And depending on the researcher's skill in analysing and interpreting these data, one may end up with a pretty good understanding of what learners are actually doing. But the question remains whether they are actually concentrating on the language, on their errors, on the instructions given by the tutor, on the "pretty pictures," or whether they are just dreaming off and looking somewhere else completely. Teachers of online classes are likely interested in whether learners actually take on board the support and the corrections offered.

There are various options for getting closer to this kind of information:

- one can ask the learners,
- one can test the learners' recollection afterwards and draw one's own conclusions, or
- one can try and capture where learner attention is focused during the task.

For the third option, tracking the gaze focus of a learner can be helpful. Although it is by no means completely accurate, the eye-mind hypothesis (Just & Carpenter, 1980) claims that in reading, the reader focuses the eye on the word just processed. In other words, the focus of one's gaze at a certain time correlates to the focus of one's attention (Duchowski, 2003). This might be totally untrue in certain situations (a case in point might be a boring language class where learners make an effort to stare at the board, but their thoughts are somewhere else completely). However, there is a strong likelihood that during concentrated tasks, as for example in an online language learning task where students have to drag images on to the appropriate vocabulary item given, the eye focus really is an indication of mental focus.

Based on this eye-mind hypothesis, many researchers have used eye-tracking in various ways to gain a clearer understanding of learners' thinking (Anderson, Ferreira, & Henderson, 2011; Just & Carpenter, 1976). The authors of this chapter have specifically applied the technique to language learning during synchronous online activities, such as synchronous text chat and multimodal online tutorials, involving online communication between two or more participants.

In general, our area of research is Synchronous Computer-Mediated Communication (SCMC), as opposed to the more frequently researched asynchronous

or single user online tasks, such as reading or watching videos with subtitles (Caffrey, 2008; Winke, Gass, & Sydorenko, 2013). This chapter will provide readers with an overview on how eye-tracking has been used in the past, encourage scholars to consider eye-tracking as an option for research projects, present two cases of using eye-tracking in our own research, and evaluate the benefits and challenges of eye-tracking for researching language learning online. Especially for novice researchers, we have added a section intended to aid reflection and decision about setting up a first research project using eye-tracking.

Justification of eye-tracking research: Background and personal stories

Our main motivation for using eye-tracking developed as we, along with other CALL researchers, became increasingly concerned with our reliance on what has been referred to as *impoverished* data (O'Rourke, 2008). CALL researchers and teachers using CALL tools are often too quick to assume that because a particular tool has certain affordances, the learners actually exploit these affordances fully. Likewise, we are often quick to ground our assumptions about the nature of CALL on results from one or two studies, sometimes decades old, a trend which can lead to a perpetuation of assumptions about learner behaviour and learning gains. It was the convergence of these two issues that prompted the second author to explore how we might overlay more methodological rigor in our studies of learner interaction in CALL environments. Perhaps it *is* true that SCMC interactions are like conversations in slow motion (Beauvois, 1992) and that this slower pace affords more processing time for learners to notice less salient features in the input. However, the research actually demonstrating this was sparse, and we seemed comfortable with a rather large leap of faith. Further, we were normally satisfied using chat transcripts of learner interaction as evidence of what learners actually did during SCMC chats. This is despite the fact that tools, such as screen capture and key stroke-logging technology, were readily available (for a history of CALL research see Bax, 2003).

Essentially, all of this comes down to the necessity to track learner behaviour. Fischer (2007, 2012) has pointed out that without knowing what students really do when they use a particular program, CALL researchers and developers run the risk of operating in a theoretical vacuum. This is obviously important when trying to evaluate claims of the effectiveness of certain software components, as they relate to language learning. At a minimum, we need to know whether or not students use them and, if so, in what manner. Fischer also demonstrates that there is very often a poor correlation between students' reported and actual use of specific CALL program components. For example, Fischer (2007) found

that students were at best not consistently aware of what they did as they used a particular program, which calls into question the reliability of their perceptions of the value of the program's components. If students' self-reports on the use of program features are unreliable, then their judgments of the instructional value of those features must be considered suspect, as evidenced by the absence of any relationship between perceptions of value and component use. We would argue that there might be even a worse correlation between what learners are supposed to do (as required by the task) and what they choose to do.

Tracking techniques can provide essential information in this regard, but while tracking techniques can tell us what students do, they cannot tell us why they do it. To get at the latter question, we need to employ appropriate retrospective and introspective methodologies in tandem with such tracking. In terms of human-computer interaction (HCI), the tracking research has shown us that students often use the software quite differently from how developers intended (Pujola, 2002), that there is much individual learner variability in interaction with CALL programs and in the amount of material learned (Chun & Plass, 1996; Collentine, 2000). On the brighter side, Heift (2007), in her discussion of learner personas in CALL (see also Chapter 6, this volume), outlines the importance of understanding how learners most effectively use the learning tools that we construct for them. Through tracking learner interaction with E-Tutor, she was able to identify three learner personas: adamants, browsers, and peekers, which were closely aligned with varying degrees of target language proficiency. This finding allowed several data driven hypotheses and decisions about CALL systems design, as it relates to individualized foreign language instruction. Tracking learner behaviour also allowed Chun and Payne (2004) to show the relationship between working memory capacity and the reported behaviour of learners looking up words in a multimedia application.

The next section of this chapter will present how the authors became interested in using eye-tracking in CALL research. Bryan Smith's main interest is in human-human interaction via computers. In one of his first attempts at providing a more robust record of what learners are doing in task-based SCMC, Smith (2008) found that using only the chat output log file underreports by over six-fold the amount of self-repair learners engage in when compared with a slightly "truer" record available from the screen capture record. This leads us to a fundamentally different interpretation of the chat interaction and has implications for instructed SLA. For example, based on the output logs alone, one may very well get the impression that the text-based medium does not greatly affect leaners' likelihood to attend to their own output. In follow-up work, it was discovered that a more detailed record provided access to key information about the effects of "interruptions" by the interlocutor on the output produced by learners (Sauro

& Smith, 2010; Smith & Sauro, 2009). Learners were also found to produce more complex or sophisticated language immediately after they delete a portion of their own text before sending it on to their interlocutor. Such a finding contributes to the SLA discussion on post-production monitoring. These data are there for the picking – we just need to employ the right tools and invest the required amount of energy to gather them.

Adding eye-tracking technology to the available suite of methodological tools was a logical next step. Smith's main interest is the intersection between SLA theory and CALL, so questions about whether the SCMC environment afforded learners more opportunities to notice certain features in the input, including corrective feedback from their interlocutor, as well as noticing features in their own output, was very compelling. Smith's current approach is to combine multiple modalities of data collection from learner tracking with retrospective techniques, such as stimulated recall.

Lijing Shi and Ursula Stickler started researching online Chinese tutorials by recording tutorial interaction in a multimodal synchronous environment. This teaching/learning environment allowed students to interact with a tutor and with peers during scheduled online sessions. The tutor could upload images and text, so called "whiteboards," to prepare the lesson. The students could speak, use text chat to communicate in writing, move items around the whiteboard, and use emoticons to express feelings, agreement and disagreement, and raise their virtual hand to indicate a willingness to speak.

The initial analysis of online language tutorials was done on the basis of screen capture and video recording without recourse to any eye-tracking equipment. This method revealed the different modes used and combined, for example, linking expressions of emotion with verbal utterances, and identifying the multiple ways the tutor provided feedback to students in this rich environment. All of these aspects provided valuable information about the processes and the possibilities of online language teaching. In addition, we employed qualitative methods, such as field notes and stimulated recall, when gathering information from a tutor and a student. We wanted to find out whether the tutor's intention in conducting the teaching tasks matched with the students' perceptions. Our findings confirmed that sometimes they are, and sometimes they are not (Stickler & Shi, 2013).

A weakness of our method was that we could not capture the students' or the tutor's reflections immediately. Teachers' intentions in lesson planning might be easier to capture, as they are rational and planned events. Students' perceptions and expectations, on the other hand, can change, depending on circumstances in the tutorial. A tutor's instruction might be confusing or misinterpreted, and that can lead to students expecting a different task than was intended by the tutor.

These instances of confusion and puzzlement that we discovered quite indirectly in our previous research led us on the search for a better method to capture them "on the go," as it were, as close as possible to the event, and we came across Smith's paper and eye-tracking as a possible tool.

In 2011, we started our first eye-tracking project, investigating students' eye focus during Chinese online tutorials. In 2013, we followed this up with a study of the tutors' gaze focus during tutorials.

Before we explain more about how we carried out our own research with eye-tracking software, we will talk about eye-tracking research in different contexts: reading research, usability studies, accessibility and – more specifically – CALL research.

Eye-tracking as a research tool: Three areas and three approaches

Eye-tracking in reading research

Eye-tracking technology has been employed as a tool in psychological reading research for over 100 years. One of the first researchers to study eye movements was Émile Javal, who wrote a series of articles on the visual process during reading from 1878 to 1905. "Beyond mere visual observation, initial methods for tracking the location of eye fixation were quite invasive – involving direct mechanical contact with the cornea" (Jacob & Karn, 2003, p. 574). The first non-invasive eye-tracking technique was developed by Dodge and Cline around 1901, which could record the light reflected from the cornea (Wade & Tatler, 2011). The main eye-tracking techniques were various combinations of corneal reflection and motion pictures before the first head-mounted eye tracker was invented in the late 1940s (Hartridge & Thompson, 1948). Mackworth and Mackworth (1958) devised a system to record eye movement, superimposed on the changing visual scene viewed by the participant. "Eye movement research and eye-tracking flourished in the 1970s with great advances in both eye-tracking technology and psychological theory to link eye-tracking data to cognitive process" (Jacob & Karn, 2003, p. 574).

Eye movements during reading can be used to infer moment-by-moment cognitive processing of a text by the reader without significantly altering the normal characteristics of either the task or the presentation of the stimuli (Dussias, 2010). These movements are considered empirical correlates of processing complexity, which allows us to make inferences about perceptual and cognitive processes. As Rayner (1998), one of the most prominent eye-tracking researchers, explains, eye movement patterns can provide insights into a reader's cognitive

processes during things like pronoun resolution and co-reference and resolving lexical and syntactic ambiguity in both L1 and L2.

The most widely used measure in eye-tracking research is the eye fixation. Eye fixations reflect when information is being encoded, allowing readers to extract important and useful information about the text (Dussias, 2010). Though there is considerable within- and between-reader variability, which is brought about by differences in cognitive difficulty in processing a text, eye fixations during (L1) reading in English generally last approximately 200–250 milliseconds (Rayner, 2009). Reading research also shows that L1 readers do not fixate on every word in a text, but rather they fixate on about two-thirds of the total words (Just & Carpenter, 1980). Things that have been found to affect whether and for how long a target is fixated include word frequency, length, predictability, and function, as well as the syntactic and conceptual difficulty of the text (Dussias, 2010; Rayner, 2009; Rayner & McConkie, 1976; Rayner, Carlson, & Frazier, 1983; Rayner, Sereno, Morris, Schmauder et al., 1989).

The duration of a fixation is often argued to be linked to the processing-time applied to the object being fixated. Researchers assume that a longer fixation duration indicates either difficulty in extracting information, or that the object is more engaging in some way (Just & Carpenter, 1976). This reflects the so-called eye-mind assumption mentioned above, which holds that the reader's eyes remain fixed on a word as long as the word is being processed.

Eye-tracking in HCI research

The second area where eye-tracking is used and has been gaining popularity recently is human-computer interaction (HCI) and its two applications: usability research and assistive technology. Due to different research purposes, there is a noticeable difference in terms of what eye-tracking equipment is used and how eye-tracking data is collected, analysed, and interpreted. The two main options are reading research and usability research. Eye-trackers take samples of the corneal reflection at varying frequencies, measured in Hertz (Hz). For example, while a sampling rate of 60 Hz is considered good enough for usability studies, reading research requires sampling rates of around 500 Hz or more (Poole & Ball, 2006). In the context of usability evaluation, the following three metrics are mainly used: fixation-derived metrics (e.g., fixation duration, number of fixations overall), saccade-derived metrics (e.g., number, amplitude), and scanpath-derived metrics (Poole & Ball, 2006).

Researchers in HCI have deployed eye-tracking to improve interface design by, for example, investigating the nature and efficacy of information search

strategies on menu-based interfaces or evaluating the effective usability features of websites. Lately, market research has used eye-tracking to determine what type of advert design on websites attracts the greatest attention (Poole & Ball, 2006). A good overview of eye-tracking for usability research from its beginnings to the publication date of the chapter can be found in Jacob and Karn (2003).

Although not directly concerned with language learning, taking this perspective on the interaction between users and screen can provide valuable information about the influence of specific design features on learners' shifting attention and cognitive focus. Research on multimodal online environments can benefit from this research method and from widening the perspective beyond text-on-screen. Particularly, in researching beginner language learners, multimedia applications often involve images or a combination of visual and textual representations of a concept. How learners use visual information to supplement their language learning at this stage can be observed with eye-tracking.

In usability studies, eye-tracking has found applications in website design and virtual training. In this type of research, compared to reading research, eye-tracking measures are not as detailed. The main metrics used are fixation duration, rate and count, and scan path. Ideally, eye-tracking research is carried out to mimic the user environment closely, thus providing information about real users engaged in authentic tasks. It is applied research that needs to be fed back into the design and production of web-interfaces, digital displays, virtual environments, and other interfaces between human users and computers.

Recent research in this area also suggests the use of supplemental data, such as questionnaires and interviews, to extract deeper understanding of the user gaze data (Nielsen & Pernice, 2010). Using a mixed methods approach can extend the data provided by eye-tracking to a fuller understanding of the complexity of the user's thought process or intention. Gidlöf, Holmberg, and Sandberg (2012), for example, used retrospective interviews to supplement the quantitative data collected through eye-tracking teenagers' perusal of online advertising. The qualitative data revealed "advertisement avoidance strategies" that changed the researchers' estimate of how much online advertising is actually taken in by the adolescent reader.

To access data that is even more closely linked to the "real world" experience of authentic users, researchers have tried to employ mobile eye-tracking devices for over a decade. As recent as 13 years ago, Jacob and Karn (2003) were still quite pessimistic about the feasibility of mobile eye-tracking, due to the high interference, necessary restriction of users' movements, or unreliable data. With the advent of head-mounted, easy-to-wear eye-tracking glasses, and new mobile eye-tracking devices, a further push towards real-life studies has gained momentum.

Another aspect of eye-tracking research is its potential use for assistive technologies. The Open University, UK, is a distance teaching institution attracting a high number of disabled learners. To facilitate their learning, research is being conducted into accessibility of digital information and assistive technologies. In our labs, we test course websites for visual complexity, visual material for impaired users, and alternative input tools for people with mobility issues. Eye-tracking has been one of the most promising tools as a device for computer input (Levine, 1981; MacKenzie, 2012) and computer interaction for disabled users using their eyes for input (Donegan et al., 2012; Hutchinson, White, Martin, Reichert, & Frey, 1989). The latest accessibility research has used eye-tracking to create a gaze-based control system for interacting in a virtual environment (Jimenez, Gutierrez, & Latorre, 2008) and to control in-car functions, like audio and comfort modules via line of sight, in automobile head up display (Fang, Kong, & Xu, 2013).

Eye-tracking in SCMC research

Eye-tracking in SCMC research has been shown to be a useful and effective tool for identifying what learners attend to during chat interaction. O'Rourke (2008, 2012) used eye-tracking as one measure to illustrate the insufficiency of relying on output logs. He also employed this technology to show learner reading patterns during SCMC, specifically the nature of learner self-monitoring of output during chat. Smith's (2010, 2012) work has explored the effectiveness of corrective feedback on learners during chat interaction. Smith (2010) showed that learners noticed about 60% of the intensive recasts they received with lexical recasts being much easier than grammatical recasts for students to notice, retain, and produce more accurately on a written post-test. Students were also better able to use these targets more productively in subsequent chat interactions. Smith (2012) compared the effectiveness of using stimulated recall and eye-tracking as measures of learner noticing of corrective feedback. He confirmed the strength of both measures in this regard. Further, the eye-tracking and stimulated recall data also suggest that although learners engage in similar amounts of viewing activity across recasts targeting various linguistic categories, they are able to notice semantic and syntactic targets more easily than morphological targets.

Smith and Renaud (2013) employed eye-tracking technology to explore the relationship between second language (teacher) recasts, noticing, and learning during task-based SCMC. Using occurrence, number, and duration of fixations as independent variables, they showed a positive relationship between noticing of lexical and grammatical form and post-test success one week later. Specifically, learners focused on close to 75% of teacher recasts, with between 20% and

33% of these resulting in post-test gains. Suggestive (but not significant) effects were found for number of fixations and post-test success. Stickler and Shi (2015) combined eye-tracking with stimulated recall interviews to investigate online language tutorials, looking not only at the online reading process of L2 learners but also at their speaking interactions with other learners and the teacher.

Three approaches

After exploring *how* eye-tracking has been used in different disciplines, we are now going to look at the *why* and try to link the research areas to underpinning philosophies. Fundamental to the existing strands of eye-tracking research are three different approaches: simplified, they can be called empiricist, socio-constructivist, and participatory.

The empiricist, or neo-empiricist, approach is based on the idea that observational or sensory evidence is indispensable for knowledge of the world. Behaviourist research, and much of psychological research, will most likely fall into this category (see for example Rayner, 1998). This type of approach is suitable for a cognitive perspective. The second strand bases its quest for knowledge on a socio-constructivist understanding of the world (Glasersfeld, 2001; Prawat & Floden, 1994; Vygotsky, 1978; Zuengler & Miller, 2006); facts are determined by the relationship between people and their environment. Researchers take part in the process of finding out the same as their "subjects," and findings can never be determined by simply distancing the research instruments from the research subject. A reflection of the researcher's own thinking is a fundamental part of the research process and the findings. Some of the psychological, and much of the sociological and educational research, will fall into that category, particularly those areas focusing on the social aspects of learning and behaviour (see for example Gidlöf et al., 2012; or Smith & Renaud, 2013). This research places human action in a social context, seeing the tools used (e.g., language, computers) as mediating interaction with the world (Wertsch, 2007). And finally, research can also be seen as fundamentally an interested engagement for the benefit of both participants and researchers. Action research (Lewin, 1946) is a prime example of this type of engagement; the quest for knowledge here is overtaken by a quest for change or improvement of the human condition. Supporting accessibility for disabled users by using ICT is a clear case in question, as are participatory action research projects, particularly in teaching or training (for example Fang et al., 2013; or Stickler & Shi, 2015).

Two examples of eye-tracking studies in SCMC

Case one: Chinese learners in an online language tutorial (Lijing Shi and Ursula Stickler)

As mentioned earlier, Shi and Stickler started their eye-tracking research looking at Chinese online tutorials. In order to better understand why learners are puzzled or fail to grasp a tutor's instructions exactly, we decided to find out precisely what students' attention is focused on during online tutorials. Teaching at a distance learning institution, we are very aware of the necessities of a clear and unambiguous interface our students can use without direct instruction or intervention from a teacher. Online tutorials are an integral part of our students' language learning, and one of the few opportunities they have for practising speaking in their L2. For this reason, every step in improving the online learning experience must be well planned and, ideally, grounded in principles drawn from research. Our university is well equipped for this type of research, placing great emphasis on usability and accessibility of all learning materials for all students.

We chose eye-tracking to capture "exact" information about language learners' attention focus, e.g., areas of interest, frequency and duration of gaze, during an online tutorial. As online tutorials consist of different tasks and activities, we investigated two tasks: learners' attention focus during reading tasks, as well as when they are engaged in interactive tasks.

Eye-tracking is a powerful tool for identifying *what* learners fixate on, as well as *when* and *for how long* they fixate on a given point of text or an image. The technology tells us nothing about *why* learners fixate their eye gaze on a specific point. Hence, the way we employ eye-tracking does not rely solely on quantitative measures. Bearing in mind limitations of the eye-tracking method, we combined it with stimulated recall interviews to understand the reasons behind learners' attention. To identify the "instances of puzzlement," it would have been enough to record gaze focus and find recordings where the gaze flickers or fixation points are more disparate than usual. This can be interpreted as a sign of difficulty, confusion, or puzzlement. However, to say with some confidence that the learner is just at this moment confused by the instructions, does not really know where to find the answer to a given question, or is overwhelmed by the task, we still rely on the recollection of the learner. And using, once again, stimulated recall to gather this information leads us to a mixed-methods approach.

Our participants were ten adult learners of Chinese. Most were in the early stages of their study of the Chinese language and were classified as belonging to the category of beginners to lower intermediate students. All learners were computer literate adults in full-time or part-time employment and had taken Chinese

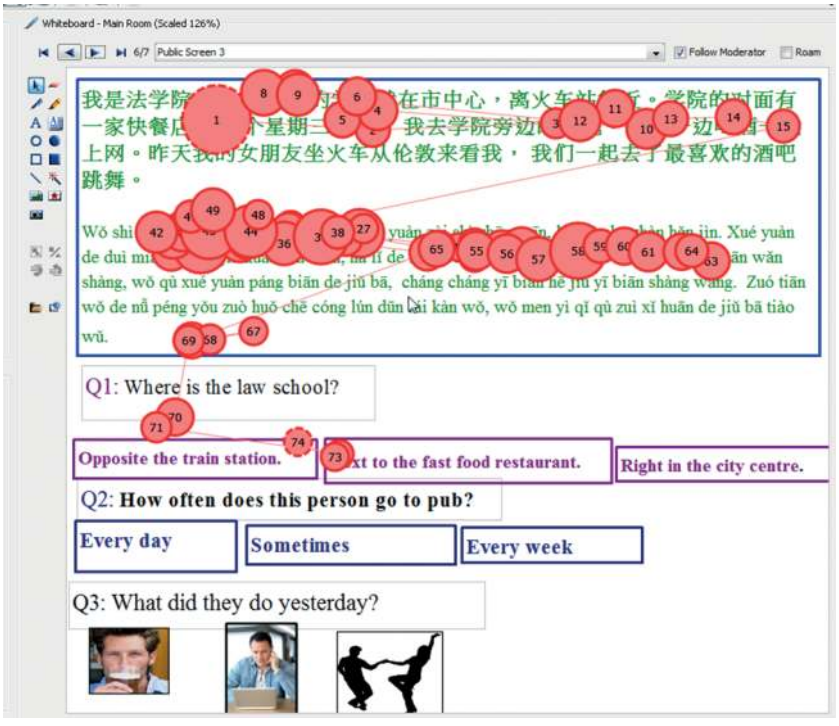


Figure 8.1 Gaze-plot of eye-tracking Chinese reading

as an optional course. For this study, the learners took part in one reading and one interactive online activity, both of which were recorded in the eye-tracking lab at the Open University, UK. First, their gaze focus was tracked and recorded (see Figure 8.1), and in subsequent stimulated recall interviews, the learners watched the recording of their gaze focus and simultaneously reflected on their engagement with the screen and recalled their intentions during the reading or speaking tasks.

Using eye-tracking data helped us to demonstrate that during reading tasks, when Pinyin[1] transcriptions as well as Chinese characters were presented, all beginner and lower intermediate participants focused to some degree on the Pinyin. Our stimulated recall interviews revealed some key motives influencing learners' attention on Pinyin and character reading: for comprehension, confirmation, and consolidation. Weaker learners relied on Pinyin for comprehension, as they had limited knowledge in characters, whereas those with more knowledge in

1. Pinyin is a method of representing Chinese characters with Western script, making it easier for novice learners of Chinese to read and pronounce the words.

characters used Pinyin to confirm as well as to consolidate their understanding of characters, and vice versa.

The interactive task we used was a speaking practice during an online multimodal tutorial involving a tutor and three to four other students, apart from our participant in the lab. In analysing the gaze focus of our participants, we decided to concentrate on different areas of interest on the screen: the whiteboard containing information for the task, vocabulary help, and images; the participants' window, displaying the names of everyone present in the tutorial and the emoticons they displayed; and the technical areas, where learners use tools, e.g., the microphone or text chat, to interact with others in the online environment. We clustered together Areas of Interest (AoIs) of the same type so all the sections with technical functionality, all the social interactive areas, and all the content sections were clustered and added together for numerical analysis. Fixation duration on the same type of AoIs shows that learners' gazes were drawn to content AoIs approximately 70% of the overall fixation duration, to social AoIs for approximately 20%, and to the technical AoIs for approximately 10%.

Experienced online teachers usually state that they expected to find that participants spent about one-fifth of their attention on social AoIs. However, experience and anecdotal evidence is one thing, to prove and quantify this finding using eye-tracking is going beyond that and still a valid research endeavour. Thus, eye-tracking data has helped us to verify what some experienced teachers might intuitively already know. During their stimulated recall interviews, participants explained their needs for spending time on social presence. They liked to know who the other participants were and see peers' responses to their performance. Online language learning is not just a cognitive activity but also an interactive and social one. Combining both these methods, eye-tracking for numerical data and stimulated recall interviews for participants' views, our study confirmed the importance of social presence in synchronous online tutorials, and the role of Pinyin for both reading comprehension and speaking production. For full details of our study, see Stickler and Shi (2015).

Case 2: The effectiveness of written recasts in teacher-student online conferences (Bryan Smith)

Smith's research strives to make whatever findings may emerge to be transparently relevant to classroom teachers. For this reason, common and authentic tasks as well as freely available software and websites are used wherever possible. In Smith and Renaud (2013), we capitalized on planned teacher-student conferences on writing assignments as our instructional context. Sixteen volunteers (8 from a

Spanish class and 8 from a German class) agreed to conduct one of their planned three teacher-student conferences online in a synchronous environment. In consultation with the teachers, we decided to use Google Talk as the chat interface for the study. The treatment consisted of one fifteen-minute online conference (text chat) about a first draft of an essay, due the next week.

As the research questions for this particular study concerned the effectiveness of written recasts by the teacher, the instructors were asked to provide full recasts to learners when it seemed natural to do so. They were also asked to provide corrective feedback on whatever they chose, but to pay special attention to errors of morpho-syntax, such as grammatical gender, as well as those having more to do with word choice or spelling in order to get a variety of recast targets. Since previous research suggests that learners vary widely in their production of immediate and delayed uptake, it was decided to not use uptake as a measure of noticing or learning. Rather, noticing was based on the occurrence and duration of eye fixations on a recast target, as well as the number of fixations on that same target. In terms of learning, we decided that individually sculpted post tests were in order, since it is impossible to create an immediate post-test based on learner interaction that just occurred seconds before. The delayed post-test was constructed by taking each of the problematic utterances that elicited a recast from the teacher and isolated that line as an individual post-test item. That is to say, learners' own chat transcripts were used as the basis for their post-tests. In all cases, there was at least one error in each of the utterances. An equal number of distractor items were developed by the researchers, as well for inclusion on the post-test. Learners were asked to identify whether each line on the post-test was correct as presented or if it needed to be corrected. If the latter, then they were required to rewrite it in a target-like fashion.

Teacher recasts were coded for number of targets within each recast (the number of errors corrected in the recast), the specific focus of each target (lexical, agreement, tense, spelling, and other), and perceived difficulty (agreement and tense, for example, were coded as difficult, whereas lexical items were not). Eye fixations, where they occurred, were coded for number (number of different fixations on a given target) and total duration. Only fixations over 200 ms were considered viable.

Through this rigorous coding and tracking, we were able to come to the following conclusions:

1. Learners focused on recasts of their non-target-like utterances about 72% of the time, and they often looked at the salient features in the recast more than once.

2. Between 20% and 33% of the targets were scored as correct on the post-test one week later (with no pedagogical intervention in the meantime).
3. The strongest predictor of post-test success was whether or not the learner fixated on the recast target for at least 200 ms.
4. None of the following variables seemed to affect post-test score: fixation *duration* on the target, linguistic focus, number of targets within a given recast, complexity, and difficulty.
5. There was a strong suggestive effect (not statistically significant) for the number of fixations on a target and the likelihood that the learner would get that target correct on the post-test, with three fixations being the best.

For the full details of the study, see Smith and Renaud (2013).

Findings and possibilities for other researchers

Using eye-tracking in our two studies proved fruitful, as it showed us elements of student learning that would have remained hidden in traditional retrospective methods or even in video screen capture. In Smith's study, support for learning measured in a post-test could be linked to eye fixation data, thus showing that recasts by the teacher have a measurable influence on student learning. In Shi's and Stickler's research, the focus of students' attention on the social areas of the screen could be measured exactly, quantified, and correlated with other data. Visualizations produced by good eye-tracking software also proved extremely useful as a stimulus for the recall interviews. Thus, combining two methods and two methodologies (quantitative with qualitative) created a deeper understanding.

As eye-tracking is a relatively new research method and has only recently become available to a higher number of researchers, there are still many questions and areas that new researchers can investigate. By looking at traditional questions (Do recasts work? Is noticing linked to uptake and learning? Are some things more difficult to notice than others in the input?) with new methods, or by asking different questions (Why are social areas useful in online learning spaces?), the field of language learning research can be extended. SCMC is an area that is developing fast, and while there are still numerous questions waiting for answers, new questions arise constantly with new devices, new software, and new online language teaching contexts.

Reflection and challenges

Cost and set-up

Overall, we can say that conducting eye-tracking research in our fields has been successful; it helped us gain new knowledge and experience. On the other hand, there are clearly challenges and difficulties. The challenges, which are detailed in the following sections, can relate to equipment, set-up, data recording, and data analysis.

The challenge of eye-tracking starts by getting access to equipment. After all, the cost of hardware and software of suitable quality can still be quite prohibitive. Technical limitations of the equipment mean that often there is only one eye-tracker available, and only one person's eye movements can be recorded at a given time. Normally, eye-trackers are placed in a laboratory, which means space is limited and the environment is not authentic. Potential negative influences of the lab-environment need to be taken into account. Another challenge is making sure that the text size used for the chat is large enough – about 36 pt font. Only this way can one be sure that the fixation ball that the software produces in the output will sufficiently discriminate between target words and parts of words.

Dealing with more complex interactions, as for example in synchronous online tutorials or text chat interactions, is more complicated than just recording the eye movements of a reader engaged in reading a static text. Simply on a practical level, to get learners to come online at the same time for a synchronous task can prove difficult at a distance. In the laboratory setting, getting the eye-tracker to record the appropriate window of the screen can be problematic. Videoconferencing with its multimodality places additional difficulties, not only on the set-up, but also later in the analysis phase of the project.

Data analysis

Data collected by eye-tracking can be massive and complicated. To be able to analyse it, choices have to be made and complex features simplified. For example, areas of interest can be selected on a screen if the whole screen is flooded with too much information. But it is necessary to take into account that making a selection already reduces the complexity of data, potentially losing interesting information.

Finally, one should be aware of different variables participants bring to the research; they might have different levels of language skills and information and communication technology (ICT) skills, be more or less familiar with the particular software used, and more or less inhibited by a lab setting. For the specific

technology, there are additional challenges. For instance, some people's eyes are more difficult to track than others'. Eye-tracking data can be influenced by participants' physical features, such as the size of their pupils, the kinds of spectacles they wear, and by their movements vis-à-vis the eye-tracker. These variables can, of course, be what the researcher is studying, but if she/he is not aware of them from the outset, they can also hinder the research.

The chief challenge in analysing eye-tracking data for SCMC interaction is that the screen is constantly shifting in ways that the researcher cannot predict in advance. This means that one cannot assign AoIs (as on a static screen, advertisement, etc.) in advance. It is possible to do this after the fact, however, once these areas have been determined (as in the case described above). Even so, one would need to ascribe several different AoIs in complex recasts that have several corrections embedded in them – one for each error being corrected. Likewise, chat screens shift upward each time the return key is pressed. This means that the AoIs will soon be off the screen. To get around this challenge, Smith and Renaud (2013) needed to examine the eye gaze record for each recast several times by segmenting the video file that showed the recast on the learners' screens into as many individual files as there were shifts of that recast on the learner's screen. That is to say, if a recast remained visible on a learner's screen for 90 seconds before scrolling off the top, there may be three or four individual video clips of segments of those 90 seconds. Each of these segments must be evaluated independently of one another, with the total number of fixations, fixation duration, etc., being combined, later to reflect the true eye gaze pattern for that specific recast. Such a requirement makes for a quite lengthy data analysis phase of the research.

In interpreting the data from eye-tracking, novices will come across a whole new and specialised vocabulary used by eye-tracking researchers. Literature lists in handbooks on eye-tracking, for example Duchowski (2003), provide a good overview of the technical aspects.

Recommendations

Apart from a few researchers using "pure" eye-tracking as their source of data, many studies combine eye-tracking with other methods. To increase the validity and reliability of eye-tracking data, usability researchers, for example, suggest combining eye-tracking with stimulated recall, questionnaires, interviews, or observation (Nielsen & Pernice, 2010). Some SCMC researchers add key-log data for triangulation of the findings. It is worth considering, however, that mixed-methods research, although fundamentally more reliable, is often more challenging to design and carry out.

We have encountered numerous challenges in our own research, as we were and still are working in a novel application of eye-tracking technology and are exploring just how far we can push the boundaries of this research technique. Using this technology is a learning process that often involves trial and error. As in all research, complications will arise, and researchers must be ready to deal with these.

Conclusions

Reasons for and against eye-tracking as method choice

Among the most important things to decide before embarking on an eye-tracking study in CALL is whether or not eye-tracking data is really required to answer specific research questions. We say this because of the expense, time commitment, and complex data-collection and analysis procedures an eye-tracking study demands. We strongly believe that researchers should strive to build the richest record possible; however, one should not ignore issues of practicality. It is also our opinion that eye-tracking should be used in conjunction with other, more established data-collection techniques. The only information eye-tracking data provides is where, when, and for how long a participant was looking at a specific point on the computer screen. Making inferences about why participants' eye gaze was fixated on a specific location requires an uncomfortable leap of faith without other independent measures. One also needs to determine in advance which of the eye-tracking outputs (fixation duration, heat maps, etc.) will be used as measures and why. Does a research question simply require a binomial variable, such as eye fixation equals *yes* or *no*, or can a case be made for counting the duration of each eye fixation as a more continuous variable? Finally, as in all research, things will not go perfectly! Researchers need to be prepared and take preemptive measures to handle calibration problems, due to influencing factors, like make-up, piercings, eyeglasses, and eye shape. These factors often confuse eye-trackers as they attempt to lock onto a participant's pupils.

Guidelines

Anyone new to CALL research and who is interested in employing eye-tracking technology in their research will need to clearly establish the nature of their research and explore whether and how eye-tracking techniques might help to answer a specific research question.

Although we have all admitted to finding eye-tracking a fascinating and insightful research method, one should not be taken in by fancy technology or the "latest gadget" approach when designing a research study. As it is commonly stated, at the start of every research is the researcher's own question, or even before that: her or his own interest in engaging with research.

To determine whether and how someone wants to use eye-tracking, it is important to first think about one's own position with regards to the fundamental questions of knowledge and understanding raised in the sections above. Then, familiarizing oneself with an overview of eye-tracking research (Jacob & Karn, 2003; Lai et al., 2013) to determine where to position this future research. All three approaches to research mentioned above are equally justified, but any choice will depend on the purpose of the project.

If the intention is to increase specialist knowledge, e.g., about attention focus during second language reading (cognitive aspects), a neo-empiricist approach will probably be chosen, setting up an experiment that measures different users' eye movements during an on-screen reading task with as little distraction as possible. This approach links to empiricist principles detailed above. Methodologically, researchers would normally have clearly defined questions and measurements, which will help narrow down the possible answers, and a straightforward, tightly planned, well-designed research set-up.

If the main interest is in understanding the changing behaviour of learners when engaged with either a human or a machine interaction, an activity theory approach (Engeström, 2001) can be used and the influence of mediating tools can be observed by tracking the gaze focus during online interactions. The set-up will be less tightly controlled, and rather aim for a more naturalistic setting to observe learners' behaviour as it occurs in an authentic environment. Although the research might take place in a laboratory, it will include external factors, such as participants' objectives for learning and attitudes towards ICT.

Finally, if the main aim is to improve the opportunity for learners to interact with each other in an online tutorial or to increase the awareness and range of strategies for teaching online available to language tutors, a set-up that combines eye-tracking with reflective and awareness-raising methods, for example, stimulated recall interviews might be most useful. In this type of research, the participants' experience and growth will be important outcomes on a par with an increase in knowledge and understanding of the online interactions.

Researchers who want to implement change often have a vested, even passionate interest in the process, similar to what we found in examples of action research above. The methodological approach is different here: researchers usually understand the given situation to some extent and can express this quite well, but to raise this knowledge to the level of academic research, they engage in systematic

interaction with the situation often based on close observation, participation, and intervention. Although research design is a part of this type of research, and data will be collected and measured, the researchers' engagement with participants will not stop at this point, as the goal is to engender change. Reflection plays an important role in this type of research.

The (neo)empiricist, sociocultural, and participatory approaches are just a rough division of what researchers do in real life. A lot of our work is actually placed in between disciplines, approaches, and epistemological stances. Eye-tracking can be used as one of the tools that delivers information or allows the researcher to engage more deeply with the participants. Researchers can also combine two approaches in an attempt to provide data ("evidence") to convince stakeholders that something is in need of change. Or they might start off with an action research approach to online learning, only to find out that the eye-tracking data in itself has given them information about learning processes.

Making a decision

Any researcher who, after reading this far, thinks that eye-tracking may be a worthwhile method for investigating learners' online behaviour, one that may help learners make the most of SCMC for language learning, may ask themselves some questions that are specific to their own situation. For example, apart from the challenges we have listed, what other difficulties can arise? For instance, researchers working for a small institution might find it more difficult to get access to eye-tracking equipment or the technical support for their study. Researchers might be worried that there are not enough suitable journals to publish their findings. To help with these final considerations, we list a few questions to guide through the decision process.

- What are the pros and cons of a decision? Draw up a list and allocate specific weight to different items (e.g., the "pro" of publishing a good article vs. the "con" of having to learn how to operate an eye-tracker).
- What are the costs and benefits for the various parties concerned? (e.g., the cost to the institution, the benefit to future students, etc.).
- What is the worst that can happen? Would it endanger students or the broader research agenda? Conduct an informal risk assessment.
- Is it worth the effort? The cost? The time? Draw up an accountancy sheet.
- How will the masses of data be dealt with? Is it worth collecting so much?
- How much data will be needed as a minimum?
- Can a little "trial run" be conducted first? If full commitment is difficult straight away, why not do a pilot study with just one participant?

Next steps

Once a researcher has decided that eye-tracking is well worth pursuing as a research method, the next steps are most likely practical in nature: finding out whether a university or research institution has eye-tracking technology, getting training in how to use the equipment, and book the labs. We suggest that making contact with researchers in an affiliated institution's psychology department is a good next step, as they are most likely the ones to have this type of equipment. Alternatively, we highly recommend renting an appropriate system before purchasing one. In many cases, initial training is included in this rental. More important than these practical aspects, however, are the conceptual challenges of designing a robust research study using eye-tracking: aligning the methodology with any underlying research interest, selecting suitable methods of data collection and analysis for every step of the project.

By ensuring that the findings are relevant, reliable, and innovative, eye-tracking can contribute significantly to an investigation of learner interactions in online language learning.

References

Anderson, R., Ferreira, F., & Henderson, J. M. (2011). I see what you're saying: The integration of complex speech and scenes during language comprehension. *Acta Psychologica*, 137(2), 208–216. doi:10.1016/j.actpsy.2011.01.007

Bax, S. (2003). CALL – past, present and future. *System*, 31(1), 13–28. doi:10.1016/s0346-251x(02)00071-4

Beauvois, M. H. (1992). Computer assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals*, 25, 455–464. doi:10.1111/j.1944-9720.1992.tb01128.x

Caffrey, C. (2008). Viewer perception of visual nonverbal cues in subtitled TV Anime. *European Journal of English Studies*, 12(2), 163–178. doi:10.1080/13825570802151439

Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behaviour. *System*, 32, 481–503. doi:10.1016/j.system.2004.09.008

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183–198. doi:10.1111/j.1540-4781.1996.tb01159.x

Collentine, J. G. (2000). Insights into the construction of grammatical knowledge provided by user-behaviour tracking technologies. *Language Learning Technology*, 3, 44–57.

Donegan, M., Majaranta, P., Hansen, J. P., Hyrskykari, A., Aoki, H., Hansen, D. W., & Räihä, K.-J. (2012). Conclusion and a look to the future. In P. Majaranta, H. Aoki, M. Donegan, D. Hansen, P. J. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye-tracking: Advances in assistive technologies* (pp. 365–385). IGI Global.

Duchowski, A. T. (2003). *Eye-tracking methodology. Theory and practice*. London: Springer. doi:10.1007/978-1-4471-3750-4

Dussias, P. (2010). Uses of eye-tracking data in second language sentence processing research. *Annual Review of Applied Linguistics*, 30, 149–166. doi:10.1017/S0267190510000005X

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156. doi:10.1080/13639080020028747

Fang, Z. G., Kong, X. Z., & Xu, J. (2013). Design of eye movement interactive interface and example development. *Information Technology Journal*, 12(10), 1981–1987. doi:10.3923/itj.2013.1981.1987

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behaviour in CALL. *Computer Assisted Language Learning*, 20(5), 409–442. doi:10.1080/09588220701746013

Fischer, R. (2012). Diversity in learner usage patterns. In G. Stockwell (Ed.), *Computer assisted language learning diversity in research and practice* (pp. 14–32). Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9781139060981.002

Gidlöf, K., Holmberg, N., & Sandberg, H. (2012). The use of eye-tracking and retrospective interviews to study teenagers' exposure to online advertising. *Visual Communication*, 11(3), 329–345. doi:10.1177/1470357212446412

Glasersfeld, E. (2001). Radical constructivism and teaching. *Prospects*, 31(2), 161–173. doi:10.1007/BF03220058

Hartridge, H., & Thompson, L. C. (1948). Methods of investigating eye movements. *British Journal of Ophthalmology*, 32, 581–591. doi:10.1136/bjo.32.9.581

Heift, T. (2007). Learner personas in CALL. *CALICO Journal*, 25(1), 1–10.

Hutchinson, T., White, K., Martin, W., Reichert, K. & Frey, L. (1989). Human-computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6), 1527–1534. doi:10.1109/21.44068

Jacob, R. J., & Karn, K. S. (2003). Eye-tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, 2(3), 4.

Jimenez, J., Gutierrez, D., & Latorre, P. (2008). Gaze-based interaction for virtual environments. *Journal of Universal Computer Science*, 14(19), 3085–3098.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441–480. doi:10.1016/0010-0285(76)90015-3

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixation to comprehension *Psychological Review*, 87(4), 329–354. doi:10.1037/0033-295X.87.4.329

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10(0), 90–115. doi:10.1016/j.edurev.2013.10.001

Levine, J. L. (1981). Any eye-controlled computer. Research Report RC-8857. New York, NY: IBM Thomas J. Watson Research Center, Yorktown Heights.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2(4), 34–46. doi:10.1111/j.1540-4560.1946.tb02295.x

MacKenzie, I. S. (2012). Evaluating eye-tracking systems for computer input. In P. Majaranta, H. Aoki, M. Donegan, D. Hansen, P. J. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye-tracking: Advances in assistive technologies* (pp. 205–225). Hershey, PA: IGI Global. doi:10.4018/978-1-61350-098-9.ch015

Mackworth, J. F., & Mackworth, N. H. (1958). Eye fixations recorded on changing visual scenes by television eye-marker. *Journal of Optical Society of America*, 52, 713–716. doi:10.1364/JOSA.52.000713

Nielsen, J., & Pernice, K. (2010). *Eye-tracking web usability*. Berkeley, CA: New Riders.

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer-mediated communication and language learning. *Computer Assisted Language Learning*, 21(3), 227–251. doi:10.1080/09588220802090253

O'Rourke, B. (2012). Using eye-tracking to investigate gaze behaviour in synchronous computer-mediated communication for language learning. In M. Dooley & R. O'Dowd (Eds.), *Researching online interaction and exchange in foreign language education: Theories, methods and challenges* (pp. 305–342). Frankfurt, Germany: Peter Lang.

Poole, A., & Ball, L. J. (2006). *Eye-tracking in HCI and usability research. Encyclopedia of human-computer interaction* (pp. 211–219). Hershey, PA: IGI Global. doi:10.4018/978-1-59140-562-7.ch034

Prawat, R. S., & Floden, R. E. (1994). Philosophical perspectives on constructivist views of learning. *Educational Psychology*, 29(1), 37–48. doi:10.1207/s15326985ep2901_4

Pujola, J.-T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, 14, 235–262. doi:10.1017/S0958344002000423

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372. doi:10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. doi:10.1080/17470210902816461

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behaviour*, 22, 358–374. doi:10.1016/S0022-5371(83)90236-0

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, 16, 829–837. doi:10.1016/0042-6989(76)90143-7

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3–4), 21–49. doi:10.1080/01690968908406362

Sauro, S. & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics*, 31, 554–577. doi:10.1093/applin/amq007

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning & Technology*, 12(1), 85–103.

Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. Directions and prospects for educational linguistics. In F. M. Hult (Ed.), (Vol. 11, pp. 79–97). Rotterdam, Netherlands: Springer.

Smith, B. (2012). Eye-tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning & Technology*, 16(3), 53–81.

Smith, B. & Sauro, S. (2009). Interruptions in chat. *Computer Assisted Language Learning*, 22(3), 229–247. doi:10.1080/09588220902920219

Smith, B. & Renaud, C. (2013). Eye-tracking as a measure of noticing corrective feedback. In computer-mediated instructor-student foreign language conferences. In K. McDonough & A. Mackey (Eds.), *Interaction in diverse educational settings* (pp. 147–165). Amsterdam, Netherlands: John Benjamins. doi:10.1075/lllt.34.12ch8

Stickler, U., & Shi, L. (2013). Supporting chinese speaking skills online. *System*, 41(1), 50–69. doi:10.1016/j.system.2012.12.001

Stickler, U., & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO Journal*, 32(1), 52–81. doi:10.1558/cj.v32i3.27737

Suzuki, S. (2013). Private turns: A student's off-screen behaviours during synchronous online Japanese instruction. *CALICO Journal*, 30(3), 371–392. doi:10.11139/cj.30.3.371-392

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA.: Harvard University Press.

Wade, N., & Tatler, B. (2011). Origins and applications of eye movement research. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 17–45). Oxford, United Kingdom: Oxford University Press.

Wertsch, J. V. (2007). Mediation. In H. Daniels, M. Cole, & J. V. Wertsch (Eds.), *The cambridge companion to Vygotsky* (pp. 178–192). Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CCOL0521831040.008

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, 97(1), 254–275. doi:10.1111/j.1540-4781.2013.01432.x

Zuengler, J., & Miller, E. R. (2006). Cognitive and sociocultural perspectives: Two parallel SLA worlds? *TESOL Quarterly*, 40(1), 35–58. doi:10.2307/40264510

CHAPTER 9

Analysing multimodal resources in pedagogical online exchanges

Methodological issues and challenges

Cathy Cohen and Nicolas Guichon
Université Claude Bernard Lyon 1, France /
Université Lumière Lyon 2, France

This chapter focuses on the contribution to web conferencing-based pedagogical synchronous interactions of meaning-making multimodal resources (spoken language as well as gesture, gaze, body posture and movement). The first part of the chapter explores different methodological approaches to the analysis of multimodal semiotic resources in online pedagogical interactions. Having presented an overview of what research into synchronous web-mediated online interaction can bring to the field of CALL, we discuss the importance of determining the relevant units of analysis which will impact the granularity of transcription and orient the ensuing analyses. With reference to three of our own studies, we then explore different methods for studying multimodal online exchanges depending on the research questions and units of analysis under investigation. To illustrate the various ethical, epistemological and methodological issues at play in the qualitative examination of multimodal corpora, the second part of the chapter presents a case study that identifies the different steps involved when studying online pedagogical exchanges, from the initial data-collection phase to the transcription of extracts of the corpus for publication.

Keywords: multimodal resources, web-mediated pedagogical interaction, units of analysis, webcam, transcription, multimodal corpora

Introduction

As a result of globalization and easy Internet access, opportunities for exposure to foreign languages have greatly increased over the past two decades (Kern, 2014). Language learners not only can access all types of documents (e.g., films, audio and video documents, written texts, and images) quickly and simply but also can

DOI 10.1075/lsse.2.09coh

exchange synchronously or asynchronously with speakers of the target language, opening up seemingly unlimited possibilities for foreign language contact and potential learning. These might be informal social interactions as learners seek out opportunities to use the target language with their peers, but they may also be specifically designed as pedagogical exchanges between a language teacher and learner, or between two learners under the coordination of a language teacher. Indeed, more and more language learning courses take place online. Such courses may involve both asynchronous (e.g., email or blogging) and synchronous (e.g., text chat or videoconferencing) tools. As a result, new interaction patterns and norms are constantly developing, and these combine a broad range of semiotic modes (Sindoni, 2013), which potentially offer new and diverse opportunities for learning.

The current chapter focuses on pedagogical synchronous interactions which use desktop videoconferencing (henceforth DVC), described by Kern (2014) as "a quintessential technological support for providing communicative practice with speakers at a distance, since it is the closest approximation to face-to-face conversation" (p. 344).[1] This powerful instrument to learn languages is an Internet-based system enabling two or more people located in different places to communicate online with simultaneous two-way audio and video transmission (Sindoni, 2013). The video transmission, made possible thanks to a webcam attached to each participant's computer, gives access to several meaning-making modes, including spoken language, but also other multimodal elements, such as gesture, gaze, body posture, and movement. With the growing number of online language courses and telecollaboration projects, it is clearly important for CALL practitioners to gain a better understanding of how these multimodal resources contribute to the pedagogical setting and to learning contexts, and also how the different semiotic resources are orchestrated in interactive technology-mediated situations (Stockwell, 2010).

This chapter will analyse the contribution of multimodal resources to pedagogical online exchanges. The first part explores the different methodological approaches to the analysis of multimodal semiotic resources in online pedagogical interactions. We begin by briefly reviewing recent literature in order to present an overview of what research into synchronous web-mediated online interaction can bring to the field of CALL. The issues of determining the relevant units of analysis will be discussed, as the latter have a clear impact on the granularity (i.e., the amount of detail provided by researchers) of transcription and orient the ensuing analyses (Ellis & Barkhuizen, 2005). Then, with reference to three of our own

1. Other technical arrangements are, of course, possible for videoconferencing, using tablets, or smartphones, for example.

studies, we explore different methods that can be employed to study multimodal pedagogical exchanges, depending on the research questions and the units of analysis under investigation. In the three studies, our focus is on the role played by technological mediation in online pedagogical exchanges and, in particular, on the affordances provided by the webcam (see also Chapter 3, this volume).

To illustrate the different ethical, epistemological, and methodological issues at play in the qualitative examination of multimodal corpora, the second part of the chapter will present a case study that identifies the different steps involved in the study of online pedagogical exchanges, from the initial data-collection phase to the transcription of extracts of the corpus for publication. The case study is an extract from Study 2, which is presented in the first part of this chapter.

Methodological approaches to the study of multimodal pedagogical interactions

In this section, we focus on different methodological approaches that can be employed to analyse how multimodal semiotic resources function in online pedagogical interactions. Studies exploring how these interactions are mediated and organized by the webcam are still quite limited, and different units of analysis have been the focus of recent research. It is important to determine the relevant units of analysis, as they have a clear impact on the type of data collected (quantitative or qualitative, see Table 9.1), on the granularity of transcription, and they will orient the ensuing analyses (Ellis & Barkhuizen, 2005).

We use the term *unit of analysis* to refer to the general phenomenon under investigation. Once the unit of analysis has been identified, it has to be operationalized by researchers who must then select the variable(s) that they are going to investigate. These are the features that the researchers believe constitute the unit of analysis (see Table 9.1). Several examples taken from the field of pedagogical DVC interactions are provided here to illustrate this. Design principles for videoconferencing tasks were used as units of analysis by Wang (2007). One of the components she explored was the role played by the webcam image in task completion. Using personal observation and post-session interviews with a small group of learners who participated in the study, she concluded that facial expressions and gestures visible via the webcam were key features that facilitated task completion. Satar (2013) focused on how social presence was established in online pedagogical DVC interactions. She explored how the trainee teachers interacting with one another used gaze, and how they compensated for the impossibility of direct eye contact. She identified a range of different uses of the webcam and highlighted the importance of eye contact for the establishment of

Table 9.1 Overview of studies on affordances of the webcam

	Duration	Type of data	Task	Design	Number of participants	Unit of analysis	Features/variables studied
Study 1	One interaction lasting around ten minutes per student	Quantitative	Describe four pictures	Experimental	Forty	Learner perceptions of online interaction	Feeling of psychological and physical presence; understanding of and by teacher; quality, naturalness, and enjoyment of interaction
						Rhythm of interaction	Silences, overlaps, turn duration, number of words
						Word search episodes	Frequency, duration
Study 2	One interaction lasting around ten minutes per student	Qualitative	Describe four pictures	Experimental	Three	Word search episodes	Multimodal orchestration of speech and non-verbal features (e.g., gaze, nods, gestures, facial expressions)
Study 3	One weekly interaction lasting around 40 minutes over a six-week period	Quantitative and qualitative	Range of different tasks and open-ended conversation	Ecological	Twelve	Framing choices	Teachers' semiotic self-awareness
					Three	Visibility of gestures in and out of the webcam	

social presence in online multimodal interactions. Guichon and Wigham (2016) explored the potential of the webcam for language teaching, focusing particularly on the unit of analysis of framing, in other words, how trainee teachers framed themselves in front of the webcam and as a result what information was made visible to their learners within the frame of the video shot. So, they investigated how trainee teachers made use of the affordances of the webcam to produce non-verbal cues that could be beneficial for mutual comprehension (see Study 3 below for more details). Their results emphasized the need for trainee teachers to enhance their critical semiotic awareness, including paying closer attention to framing, thus enabling them to gain a finer perception of the image they projected of themselves. In so doing, it was hypothesized that they should be able to take greater advantage of the potential of the webcam and, as a consequence, increase their online teacher presence.

Different methods can be employed to study pedagogical online exchanges, and researchers' choices of method will depend on the research questions they wish to investigate and the objectives of their study. We will take three examples from our own work to illustrate different approaches. In all three, we are interested in the role played by technological mediation in online pedagogical exchanges, and our particular focus is on the affordances (see Chapter 3, this volume) provided by the webcam. There are two common webcam setups. In the first, the webcam is integrated into the computer screen, where it is located in the centre just above the visible screen image and is not adjustable, except by moving the computer screen. In the second case, the webcam is a separate unit attached to the top of the computer screen or to another object, such as a shelf, or set beside the screen on a desk, and is thus more mobile. Table 9.1 provides an overview of these studies, which will be discussed in turn below.

Study 1: Quantitative approach on experimental data

The first study, reported fully in Guichon and Cohen (2014), adopted a quantitative methodology and had an experimental design. In this study, we explored the impact of the webcam on an online interaction by comparing several dependent variables between an audio-conferencing and a videoconferencing condition, using Skype. In the audio-conferencing condition, the webcam was switched off, whereas it was on in the videoconferencing condition. Our objective was to assess the webcam's contribution to the interaction. There were three research questions, each of which explored different units of analysis which we felt might operate differentially in the two experimental conditions. The first was learner perceptions, which were probed using a short post-task Likert scale questionnaire to

gauge learners' feelings of (a) the teacher's psychological and physical presence, (b) understanding of and by the teacher, and (c) the quality, naturalness, and enjoyment of the conversation. The second explored the rhythm of the interactions by measuring silences, overlaps, turn duration, and number of words. The third focused on frequency and duration of word-search episodes, which occur when "a speaker in interaction displays trouble with the production of an item in an ongoing turn at talk" (Brouwer, 2003, p. 535) and deploys an array of strategies (use of context, production of synonyms, solicitation of interlocutor's help, etc.) to avoid a communication breakdown. Before the experiment began, we had clear hypotheses, which stated that being able to see one's interlocutor would have an effect on the online pedagogical interaction. In other words, we stated that we expected to find a statistically significant difference between all the dependent measures under investigation in the audio-conferencing and videoconferencing conditions. Furthermore, for the dependent measures relating to learner perceptions, we predicted that the videoconferencing condition would be received more favourably than the audio-conferencing condition.

The independent variables were strictly controlled before the experiment began. Forty French students with a B2 level in English (according to the Common European Framework of Reference for Languages), the foreign language they were learning at university, took part in the experiment. Twenty of them were put in the videoconferencing condition and twenty in the audio-conferencing condition. Indeed, in order to be able to carry out certain statistical tests, it was necessary to have at least twenty participants in each condition. Statistical tests were used to verify that there were no significant differences between the two groups in terms of sex, age, English level, familiarity with online communication tools, and attitudes towards speaking English. Had there been differences between the two groups at this stage, we could not have been sure whether our results were due to initial group differences or, rather, to differences resulting from the testing conditions. In the experiment, each student interacted individually with the same unknown native English-speaking teacher who was always in the same setting. Furthermore, they all did exactly the same task, which consisted of describing four previously unseen photographs. This task was selected for two main reasons. First, it was not open-ended, and therefore enabled us to gather data that were comparable across the two conditions. Secondly, as observed by White and Ranta (2002), learners have to be "very precise in both vocabulary and structure, thus making demands on the learner's ability to quickly access specific linguistic knowledge" (p. 264). The four photographs showed individuals in simple situations (a group of young people at an outdoor concert; an old lady in a hospital; an intimate funeral procession; a sad child holding a teddy bear). Because lexical items carry a heavy communicative load, the meaning of such items must be negotiated if they are

unknown to learners in order to avoid communication breakdowns that would prevent the conversation from advancing (Blake & Zysik, 2003). Each of the four photographs contained what were considered to be problematic lexical items (e.g., loudspeakers, earring, wheelchair) likely to lead to word-search episodes, one of the units of analysis of the study. If students failed to give sufficient details or justify what they were saying in their descriptions with specific references to elements in the photographs, or if their descriptions were considered to be unclear or not sufficiently precise, the teacher was instructed to incite participants to elaborate, asking further open questions, such as, "How can you tell?" and "What makes you say that?" The aim of these questions was to provoke word-search episodes. When the interaction came to a halt because students lacked a key lexical item, they were encouraged to reformulate or describe the item in question. If they gave a word in French, the teacher feigned a lack of understanding, prompting students to find another way of communicating their idea. The duration of the interaction for all participants was set at around ten minutes.

In order to compare the different dependent variables between the two experimental conditions and assess the contribution of the webcam, it was necessary to carry out a quantitative study. In other words, we had to be able to measure the different variables in the two experimental conditions to see how they compared. So, for example, the number of silences and word-search episodes were counted, and turn durations were measured (see *Annotation* below for more details as to how this was achieved). All the data were then imported into SPSS, allowing statistical comparisons to be made between participants in the two conditions. Our results showed that, contrary to our predictions, there were fewer differences than we had anticipated between the videoconferencing and audio-conferencing conditions on the dependent measures, with few comparisons reaching statistical significance. The main difference was the greater number of student silences in the audio-conferencing condition.

This first study was clearly time consuming in terms of data collection and analysis. It also involved many people: forty students, a teacher, an assistant who helped organize the data-collection sessions, four research assistants to transcribe and annotate the data (see *Annotation* below), and two researchers who analysed the data and wrote up the research for publication. Although the differences between the results obtained from the two experimental conditions were far less clear-cut than we had expected, the results were nevertheless thought-provoking. We considered that, although from a quantitative point of view the presence of the webcam did not seem to have a great impact on the pedagogical interactions with regard to the units of analysis which were investigated, the webcam image could nevertheless be facilitative and modify the quality of the mediated interaction. The reality was in fact considerably more complex than our findings seemed

to show. Hence, these results also highlighted the limitations of using quantitative data to grasp the more subtle interactional aspects in a multimodal learner corpus.[2] Furthermore, our results provided a good example of the iterative process of research, with the first more generic experiment being a necessary step to reveal the need to explore particular parts of our corpus using a much finer-grained analysis. This led us to conduct our second study.

Study 2: Qualitative approach on experimental data

In this study (Cohen & Guichon, 2014), we carried out a qualitative and descriptive analysis on small sections of the videoconferencing data taken from the first experimental study. In other words, we used part of the same corpus used in Study 1, but this time to conduct a microanalysis. The analysis focused on short sections of just three of the twenty videoconferencing interactions, in order to examine how the learners and the teacher used the webcam strategically at different times during their exchanges.

Since we were particularly interested in training language teachers to utilize the affordances of the webcam during pedagogical online interactions and to develop their critical semiotic awareness, we considered that only a fine-grained analysis of non-verbal behaviour in the videoconferencing condition would enable us to identify when and how the interaction was facilitated by the appropriate use of the webcam by participants.

The methodology employed in Study 2 was quite different from the first. This time, we worked within the Conversation Analysis (CA) paradigm, as articulated in work initially conducted by gesture specialists (e.g., McNeill, 1992) and more recently pursued by researchers working on gesture in the field of Second Language Acquisition, such as McCafferty and Stam (2008) and Tellier and Stam (2010). We adapted the methodology of these authors who focus on face-to-face pedagogical interactions in order to investigate pedagogical computer-mediated interactions. We also integrated an approach from the broader domain of multimodal discourse analysis, as applied by Norris (2004) and Baldry and Thibault (2006), whose work is not conducted in the pedagogical field. Finally, our approach was influenced by recent work carried out by Sindoni (2013), who has explored non-pedagogical online interactions using a multimodal approach. In other words, the methodological approach we adopted was influenced by work

2. Perhaps there would have been a greater difference between the two conditions if a different, more interactive task had been used, such as one requiring the learners to describe the layout of a room to the teacher while she produced a drawing according to their instructions.

conducted in several domains of scientific research. By combining and adapting elements from these different areas, we created a method suitable for analysis in our own field of investigation, i.e., the study of multimodal resources in pedagogical online exchanges.

In this second study, we explored the contribution to meaning making of several non-verbal semiotic resources other than speech and investigated how they helped the teacher to manage the online pedagogical interaction and how they were orchestrated. The modes studied were proxemics, gesture, head movements, eye contact, gaze, and facial expressions. Each of these modes will now be presented briefly, with specific reference as to how they function in a videoconferencing interaction.

We considered *proxemics*, that is to say the physical distance individuals take up in relation to one another and to objects in their environment. Proxemics functions quite differently when interacting online using videoconferencing, since participants are not in the same location. Sindoni (2013) has observed that "distance is not established by those who interact, but between one participant and one machine. This distance foregrounds the *representation* of distance among users" (p. 56). Therefore, since participants are not in the same place during a mediated interaction, they must position themselves at an appropriate distance from their computer screen, framing their head and upper torso, to create just the right feeling of proximity. Being too far away may create a feeling of remoteness, while being too near, with just the head taking up the whole computer screen, may lead to a feeling of excessive closeness. Added to this, whatever position the user chooses, because he has constant access to his own image in the smaller frame on his computer screen, he is able to monitor and manipulate the image he wishes to project to his interlocutor (Sindoni, 2013). This affordance provided by webcam-mediated communication also gives the user greater control over the construction and negotiation of social space.

We examined different types of *gesture*, defined as the use of the arms and hands for communicative purposes (McNeill, 1992). We focused in particular on those gestures which were visible in the webcam: iconic gestures representing an action or an object; metaphoric gestures illustrating an abstract concept or idea; and deictic gestures used to point towards concrete or abstract spaces. Our objective here was to assess what type of information was communicated by these gestures and to what extent they appeared to facilitate (or not) the online exchange. For instance, were they transmitting some information to the interlocutor to complement or accompany what was said in the verbal channel (co-verbal gestures)? Or were they self-regulatory gestures, produced unintentionally to help speakers think, thereby allowing them to maintain a sense of coherence for themselves (McCafferty, 2008)? To what extent were they visible in the webcam?

Head movements, which may convey meaning between interlocutors (e.g., nodding in agreement; shaking one's head from side to side to convey disagreement; holding one's head quite still while fixing one's gaze on someone to indicate concentration and focus), were also considered.

Finally, we were interested in *eye contact*, *gaze*, and *facial expressions*. Compared to face-to-face conversation, gaze management is very different in online video interactions. With the current state of technology used in videoconferencing systems, it is impossible for speakers to make direct eye contact with one another (see De Chanay, 2011). When speakers direct their eyes to their interlocutor's image on their computer screen, assuming that the webcam is placed on or at the top of the screen somewhere, their eyes are slightly lowered, so not aimed directly at their interlocutor's eyes. They can choose to look directly at the webcam, which gives the interlocutor the impression that he is being looked at straight in the eyes, but in so doing, paradoxically, the speaker can no longer focus on the interlocutor's image on the screen (De Chanay, 2011). So, not only are there fewer visible gestures to facilitate communication and intercomprehension in videoconferencing interactions, but there is also the impossibility of mutual gaze. Cosnier and Develotte (2011) hypothesize that speakers compensate for this through facial expressions, which become more important and seem to be more numerous and perhaps over-exaggerated in videoconferencing interactions compared to face-to-face conversations, precisely to compensate for the lack of visible hand and arm gestures.

The different non-verbal semiotic modes have been discussed separately here, but of course during any chosen communicative event, they are operating simultaneously, and, as Sindoni (2013) has argued, "Ensembles of semiotic resources […] produce effects that differ from those produced by a single semiotic resource *and* from the mere *sum* of semiotic resources" (p. 69). A transcript and microanalysis taken from this study corpus is provided below (see *Multimodal transcript and textual analysis*) as an illustration of our approach. Since the study was exploratory, our hypotheses emerged progressively as the data were explored. Three angles of analysis became apparent with regard to gesture: (a) self-regulatory versus co-verbal gestures, (b) gestures which contribute something to the construction of the message versus gestures which potentially cause interference and are distracting, and (c) redundant gestures which duplicate what is said in the verbal channel versus complementary gestures which add some new information. The other modes under investigation will be exemplified in the detailed transcription of a small extract of the data below. Overall, the main results of this study indicated that the online teacher was better able to monitor the interaction if she was attentive to subtle visual and verbal cues (e.g., gesture, gaze, and facial expressions) and was able to deal with the needs of the learner in a timely fashion.

As Doughty and Long (2003) have pointed out, there is a short window during which feedback given by teachers is especially relevant and more likely to have an impact on learning, as will be illustrated in the extract analysed below.

This qualitative study provided us with rich and complex data, enabling us to gain insights into the multimodal orchestration of the different semiotic resources in an online pedagogical interaction. However, we were using data collected for a study carried out in experimental conditions – the interaction duration was fixed; it was the first time that both the teacher and the learners had met and taken part in an online pedagogical interaction. So, the findings may have been attributable, to some degree at least, either to the novelty of the learning situation and/or to the task learners were asked to carry out. In other words, the conditions of this second study, and indeed the first, lacked ecological validity. Thus, in our third study, we tried to address this methodological shortcoming.

Study 3: Quantitative and qualitative approach on ecological data

As shown in Table 9.1, the corpus for the third study was collected in natural conditions in order to provide a more ecological perspective. The context was a telecollaborative project in which twelve trainee teachers of French as a foreign language met for online sessions in French with undergraduate business students at an Irish university.[3] Each trainee teacher met with the same learner (or pair of learners) once a week for approximately forty minutes over a six-week period. Over this period, the trainee teachers proposed a range of different interactional tasks to their learners. So, unlike Study 2, which was conducted in experimental conditions, i.e., it was set up with the sole purpose of conducting an experiment to test our different hypotheses, Study 3 used data collected from an online course that was set up between two universities with learner training in mind: helping Irish learners to develop their interactional skills in French, and helping students training to be French teachers to develop their online teaching skills. Thus, this teaching and learning situation was not set up initially for research purposes, but the data collected from the online sessions were used subsequently to conduct research.

The research carried out in this study (Guichon & Wigham, 2016) focused on very specific elements taken from the sizeable corpus that was collected. As in the previous two studies, we were interested in how participants used the affordances of the webcam, but this time, the particular focus was on framing, i.e., how the trainee teachers framed themselves in front of the webcam and, as a result, what information was made visible to their learners within the frame of the video shot.

---

3. ISMAEL projet: <http://icar.univ-lyon2.fr/projets/ismael/index.htm>

For the qualitative part of the study, the same method of analysis was used as in Study 2. Two questions were explored here. Firstly, in order to study teachers' framing choices, screenshot images were taken of the twelve trainees each week over six weeks, at around minute seventeen of their online interaction. A quantitative approach was adopted to provide an indication of the frequency of the trainees' different framing choices along a continuum, from an extreme close-up shot, to a close-up, to a head-and-shoulder shot, and to a head-and-torso shot. In parallel, a qualitative approach was used to conduct a fine-grained analysis on the same data and, in particular, how the trainees positioned their gestures in relation to the webcam over the six-week course.

The findings revealed that head-and-shoulder shots, followed by close-up shots of themselves, were those most favoured by the trainee teachers. Furthermore, qualitative analysis of the data showed that certain trainee teachers adjusted the position of some of their gestures, in particular highly communicative iconic and deictic gestures, so that they were framed and therefore more likely to be visible to learners and, therefore, potentially helpful for learner comprehension. For example, a thumbs-up gesture, to compliment a student on something she said, was positioned right in front of the webcam in order for it to be seen, rather than in the more natural gesture space, which would fall below the level of the webcam. Furthermore, quantitative analyses revealed that these gestures were held longer in front of the webcam. So, such teaching gestures, which clearly had a communicative purpose, appeared to be produced by these trainee teachers quite intentionally, and consequently were aimed at the webcam and remained visible to the language learners for some time.

The second question investigated in this study explored the communicative functions of gestures that were visible or invisible in the frame. For technical and practical reasons explained fully in the study, data were collected for just three participants for just one session each. The teacher trainees were filmed using DVC with their learners with two distinct recordings. A screen recorder captured all onscreen activity, including what was visible and audible through the webcam, and an external camera, oriented towards the trainee teacher, was used to film what lay outside the webcam's view (the *hors champ*). When the two sets of recordings were compared, it became clear that the trainee teachers continued to perform many potentially co-verbal gestures which were either invisible or only partially visible in the webcam recordings, which only captured a close-up of the head and upper torso area. In contrast, extra-communicative gestures, such as touching their hair or scratching their ears, became much more visible because of the magnifying effect provided by the restricted view offered through the webcam. Such gestures, which may have gone unnoticed in a face-to-face interaction because of the presence of other broader contextual elements, were more difficult

to miss when communicating using DVC. Indeed, if numerous, they could become rather distracting and interfere with communication.

So, the findings of this study highlighted the need to train teachers "to become critically aware of the semiotic effect each type of framing could have on the pedagogical interaction so that they made informed choices to monitor the image they transmit to their distant learners according to an array of professional preoccupations" (Guichon & Wigham, 2016, p. 73). This ecological study provided valuable information that could be reinvested in future teacher-training courses.

Synthesis

We have explored three different studies, each of which investigates the role of HCI (human-computer interaction) in online pedagogical exchanges, with a particular focus on the affordances provided by the webcam. Both quantitative and qualitative analyses are valid means to explore the data collected, as long as the method is sound and the objective clearly stated. The qualitative microanalysis of a much broader range of units of analysis investigated within the field of web conferencing-supported teaching is certainly to be encouraged in order to further enhance our knowledge of HCI in a pedagogical setting. By putting certain elements of the interaction into the spotlight, we may progressively untangle the complexity of these online pedagogical exchanges.

The three studies discussed above highlight the complexity of designing research in a domain in which technologies for language and learning are continuously evolving (e.g., from communicating using DVC to more recent communication tools, such as tablets and smartphones). Furthermore, as these tools become more commonplace both in private and professional spheres, teachers should become increasingly aware of the semiotic affordances they offer, and teachers and learners should be more comfortable and accustomed to interacting with them. So, while the same questions related to language acquisition remain, researchers working in the field of computer assisted language learning have to adapt their research designs constantly in order to take these changes into account.

In the first part of this chapter, we have explored different methods for studying multimodal resources in pedagogical online exchanges. However, in order to be able to conduct the type of analyses presented above, researchers have to ensure that their data are collected and stored in such a way that they can be later transcribed and annotated. Whether the study is quantitative and experimental or qualitative and ecological, numerous transformations are required to progress from the initial data-collection stage to the creation of a corpus that can be presented in academic publications or at conferences, and also perhaps be shared among researchers (see Chapter 10, this volume).

In the next section of this chapter, we examine these different stages and investigate the opportunities and challenges concerning the study of data relating to synchronous mediated language learning and teaching.

Reflections on a multimodal approach to synchronous pedagogical online interactions

From the traces of mediated activity to a corpus that can be studied from different perspectives

Any mediated learning activity produces traces; digital traces, currently much used in the field of Learning Analytics, can be computer logs that provide quantitative information (frequency of access, time spent on a task, number of times a given functionality is used, etc.). The aim of these digital traces is to understand and optimize learning and learning environments (Siemens & Baker, 2012). Digital traces can also be comprised of "rich histories of interaction" (Bétrancourt, Guichon & Prié, 2011, p. 479) that provide multimodal data and time stamps that can be gathered from digital environments in order to gain an insight into certain teaching and learning phenomena. This second form of traces has been studied by researchers in the field of computer-mediated communication (CMC) for the last twenty years (see for instance Kern, 1995; Kost, 2008; Pelletieri, 2000). Thus, traces collected in forums, blogs, emails, audio graphic platforms, and videoconferencing have been built into corpora to study the specificities of mediated language learning, usually by using conversation and/or interaction analytic tools.

The present section focuses on mediated learning interactions to illustrate how technology helps fashion methodological and scientific research agendas in the field of mediated interactions. Several operations are at play when researchers deal with a data-driven study of multimodal learning and teaching, when they strive to create a corpus that can offer different types of analyses, as was illustrated in the first part of this chapter.

If we take the example of a corpus composed of recordings of online learning interactions mediated by a DVC facility, three main operations can be identified: corpus fabrication, annotation, visual and textual representation. Each of these operations will be explained and illustrated by a case study using data that were initially collected for a larger research project (Guichon & Cohen, 2014, discussed in Study 1 above). However, before we do this, it is important to underline the ethical aspects that researchers must respect when dealing with data that include images of participants.

Ethical considerations

Ethical issues are relevant to all research involving humans (see Chapter 10, this volume). In the case of the type of studies described above, which may involve the publication of participants' images, certain issues should be considered very carefully.

Before recording begins, researchers must obtain written informed consent from participants: first, that they agree to be recorded; second, that they agree to be recorded for research purposes; and third, that they agree that recordings (or screenshots) may be displayed publicly or published (ten Have, 1999). If participants consent to all three, they must understand fully what is at stake. For example, will they be recognizable from the recordings (visual, auditory)? Will their faces be blurred/pixelated to avoid recognition? Where will the recordings be shown, and where will they be published? Will they be available freely online to anyone (for a limited period of time)? Will participants have access to the recordings before they are used, in order to confirm or cancel their informed consent? (See Yakura, 2004, for an excellent discussion of the issues at stake here.)

The above questions present real challenges for researchers. First and foremost, if recordings or screenshots are to be used publicly, anonymity cannot be ensured at every stage (Yakura, 2004). Secondly, depending on what participants have consented to, researchers may be more restricted in what they can present and/or publish. If, for instance, researchers wish to provide a fine-grained analysis of the different non-verbal semiotic modes employed by participants, but are only authorized to publish faces which have been blurred, displaying eye contact, gaze, and facial expressions becomes impossible, thus "rendering the data unusable for certain lines of linguistic inquiry" (Adolphs & Varter, 2013, p. 149).

How can researchers circumvent this problem in order to preserve and communicate to others some of the richness of their data? To compensate to some extent for the loss of visual information, researchers could provide very detailed written descriptions (Lamy & Hampel, 2007). In a recent study by Sindoni (2013), because of reservations expressed by certain participants about the publication of screenshots, she opted to use drawings instead. However, she recognizes the drawbacks of this, stating, "they are time-consuming and require specific expertise, so that they can be used selectively, only for very brief and fine-grained analyses. Furthermore, drawings incorporate the researcher and artist's bias that represent participants in their interactions" (Sindoni, 2013, p. 71).

Multimodal data collection

Several applications, for instance Camtasia or Screen video recorder, can be used to capture on-screen activity in an online interaction, and this can be converted into a video file (see Chapter 7, this volume). The advantage of such applications is that they can be installed beforehand on each participant's computer, and, once switched on, they capture everything that is visible on the screen and is audible around the screen, thus providing researchers with access to all the actions and utterances produced by the participants during the online interaction. Hence, whether the study is experimental or ecological (see above), traces of the mediated activity can be collected with little or no interference on the ecology of the learning situation, even though it must be underlined that screen-recording software can slow down the computer. This is quite different from classroom-based research that requires more intrusive devices (i.e., video cameras) to collect traces of the observable teaching and learning activities.

While the traces of the mediated learning activity constitute the main material of the study, complementary data must be collected via consent forms, researchers' field notes, pre- and post- interviews, or questionnaires with the participants to gather crucial information about:

- Ethical dimensions (as discussed above);
- Socio-demographics and learner profiles: age of the participants, gender, relations to one another (in case of an interaction), familiarity with the given program or application, level in target language and motivations, experience in learning or teaching online;
- Pedagogical dimensions: nature of the interaction, tasks, themes, documents used, instructions, place within the curriculum;
- Temporal dimensions: length of each interaction, frequency of interactions (e.g., once a week), duration of module (e.g., a semester);
- Methodological dimensions: how participants were recruited for the study, how their level was assessed, how they were divided (in case of an experimental study that compares two or several groups), what they were told of the aim(s) of the study, precisely how the data collection was organized, how ethical considerations were taken into account (see above);
- Technological dimensions: type of software and hardware used (e.g., desktop or laptop, devices used for recording, etc.).

The conjunction of field notes, questionnaires, interviews, and consent forms with the main data thus helps create "a dynamic constellation of resources, where meanings are produced through inter-relationships between and within the data sets, permitting the researcher literally to 'zoom in' on fine-grained detail and pan

out to gain a broader, socially and culturally, situated perspective" (Flewitt et al., 2009, p. 44).

The data that serve as the illustration for this chapter come from Study 2, discussed above. The reader will need to know some of the attributes of the two participants who took part in the larger study (Study 1 discussed above, see Guichon & Cohen, 2014). The learner concerned was twenty years old at the time of the study. His level had been assessed as B2 (according to the Common European Framework of Reference for Languages), and he described himself as a keen language learner. He used Skype for social purposes but had never used it for language learning. It was the first time he had interacted with the twenty-eight-year-old female native teacher, and this interaction was not part of his usual class. The teacher had several years of experience teaching non-specialist university students in a classroom setting and was a regular user of Skype, mainly for personal communication. However, this was the first time that she had taken part in an online pedagogical interaction. Neither of the two participants was informed of the study's purpose or hypotheses before the experiment. The task consisted of getting the student to describe four previously unseen photographs, as discussed above. These photographs were chosen because each one contained what were considered to be problematic lexical items likely to trigger word-search episodes (see above for definition), chosen as the unit of analysis for this research. The interaction via Skype lasted for about ten minutes, and participants were asked to concentrate only on oral communication and exclude the use of the keyboard and mouse.

All the secondary data (field notes, questionnaires, and interviews) had to be digitized and grouped together with the data comprising the traces of the mediated interaction "to reconstitute for researchers, in as many ways as desired, information about the original experience" (Lamy & Hampel, 2007, p. 184) and to enrich subsequent analyses.

Annotation

There are several computer software tools that researchers can use to code audio and video data. Among these, ELAN <http://tla.mpi.nl/tools/tla-tools/elan/> is a linguistic annotation tool devised by researchers at the Max Planck Institute (Sloetjes & Wittenburg, 2008). Figure 9.1 below shows a sample of the data that were annotated with ELAN, with which the researchers can:

1. Access the video stream of one or up to four participants;
2. Play the film of the interaction at will with the usual functionalities to navigate it;

3. View a timeline aligned with media time;
4. Transcribe, on the horizontal axis, the utterances of the participants (one layer per participant);
5. Add a new layer for each element they wish to investigate (indicating, for instance, the onset and the end of a gesture and its description);
6. View annotations of one layer in a tabular form to facilitate reading.

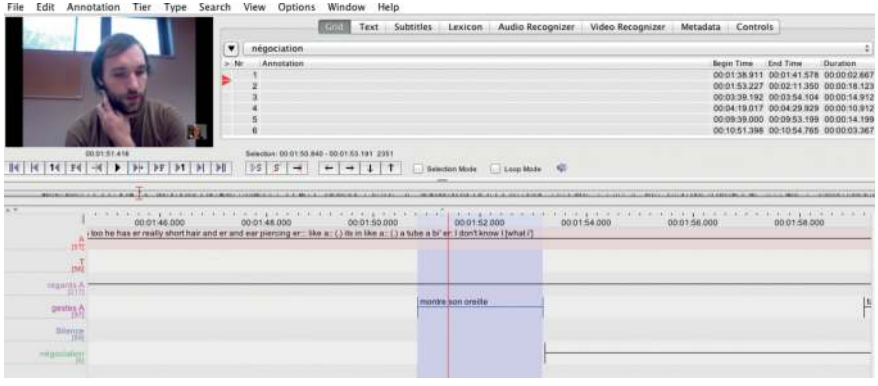


Figure 9.1 Example of a sample of data annotated with ELAN

With ELAN, there can be as many layers (called tiers) as is deemed useful for a given study (i.e., words, descriptions, events, translations, etc.). As the case study presented here focuses on the verbal and co-verbal behaviour of the learner who has to describe four photographs to a distant teacher via Skype, the elements annotated were as follows: the direction of the eyes (gaze towards the webcam, towards the screen, towards the documents on the table[4]), the gestures that were produced (e.g., *points to his ear*), and the silences between and within turns, because these are crucial during L2 oral production, especially during word-search episodes. Researchers working on multimodal data can thus align different features of the interaction, accurately transcribe data across modes, and then obtain a variety of views of the annotations that can be connected and synchronized.

The data from the three studies described in the first part of this article were all transcribed using ELAN. Hence, although the first study was quantitative and the second qualitative, the same annotation tool was used for both even though the tiers differed according to the focus of each study.

---

4. Eye-tracking was not employed in this study, but it could have been used profitably as a complement to provide more precise information about gaze direction (see Chapter 8, this volume).

Annotation corresponds to a necessary transformation of the data in view of further analysis. It is a time-consuming and demanding task that requires devising a coding scheme so that all annotations are consistent across different annotators. As noted by Adolphs and Varter (2013), coding schemes have to be carefully explained and recorded so that "they can be shared across different research communities and with different community cultures and different representational and analytical needs" (p. 155). It is methodologically sound to get two different researchers to annotate a sample of the same data in order to ensure the integrity of the coding scheme. This can be verified by calculating the inter-rater reliability to determine, for instance, whether two researchers interpret and code gestures consistently and reach a satisfactory level of agreement. If they fail do to so, the annotation scheme needs to be refined and re-tested in the same way until satisfactory inter-rater reliability is achieved (Allwood et al., 2007). Yet, as noted by Calbris (2011), "achieving the ideal of scientific objectivity when coding a corpus is a delusion, because coding depends on perception, an essentially pre-interpretative and therefore subjective activity" (p. 102).

Furthermore, priorities and research questions have to be carefully defined beforehand so that the granularity of the annotations does not evolve. Researchers such as Flewitt et al. (2009) have underlined that annotation already corresponds to a first level of analysis since it entails selecting certain features of the mediated interaction and leaving others out according to both a research rationale and agenda.

Once the data have been annotated, they can then be organized into a coherent and structured corpus (see Chapter 10, this volume, for a full account of corpus building and sharing). They may also be put on a server, allowing them to be shared with other researchers. In order to do this, close attention has to be paid to the formats of the data so that they are compatible with different computer programs. Providing researchers with clear information as to how to access the data, specifying all the contextual information (see above) and ethical dimensions (e.g., what can be used for analysis and what cannot be used for conferences or publications because participants have withdrawn their permission) are important steps to make the corpus *usable*, *searchable*, and *sharable*. The field of CMC would greatly benefit from having more researchers working on the same corpora; not only would it reduce the costs associated with corpus building, transcription and annotation, but also it would provide researchers with the opportunity of examining the same data using different tools, methods, and research questions and would therefore produce more significant and reliable results to the community at large (see Guichon & Tellier, forthcoming, for an example).

Multimodal transcript and textual analysis

Once the data have been organized into a coherent corpus, analyses can be made starting by the making of the transcript. The transcript can be defined as the representation of a sample of the corpus. Bezemer (2014) allocates two functions to the making of a multimodal transcript. The first function of transcription is *epistemological* and consists of a detailed analysis of a sample of an interaction in order to "gain a wealth of insights into the situated construction of social reality, including insights in the collaborative achievements of people, their formation of identities and power relations, and the socially and culturally shaped categories through which they see the world" (Bezemer, 2014, p. 155).

The second function is *rhetorical* in that the transcript is designed to provide a visual transformation of the trace of the interaction that can be shared with readers in a scientific publication. Transcripts chosen and prepared for an article are not illustrations of a given approach or theory but are both the starting point of the analysis and the empirical evidence that supports an interpretation and can be shown as such to readers. The researcher must therefore find an appropriate timescale (e.g., a few turns, an episode, a task, a series of tasks, a whole interaction) to study a phenomenon (for instance, negotiation of meaning in a mediated pedagogical interaction) and then define the boundaries of the focal episode. Making the transcript may also involve refining the initial research questions and determining what precise features will be attended to.

For our study on videoconference-based language teaching, it seemed crucial to understand how the distant teacher helped the learner during word-search episodes and used the semiotic resources (such as gestures, facial expressions, and speech) at her disposal. It was equally important to examine how the learner used different resources to signal a lack of lexical knowledge and how meaning was negotiated with the native teacher. Gestures, head and body movements, gaze, and facial expressions produced by both participants while the learner was trying to describe a photograph became features that were selected as especially important for the transcript (see Figure 9.2). Although conventions used for Conversation Analysis can be adjusted to multimodal transcription, new questions arise concerning the representation of co-verbal resources (gesture, gaze) with text, drawings or video stills and the alignment of these different representations so that the reader can capture how verbal and nonverbal resources interact (see Figure 9.2). Ochs (1979) underlined the theoretical importance of the transcript, arguing that "the mode of data presentation not only reflects subjectively established research aims, but also inevitably directs research findings" (as cited in Flewitt et al., 2009, p. 45). For instance, in Figure 9.2, the choice of presenting, when relevant, the images of the two interlocutors side by side (e.g., Images 5 and 6) was made

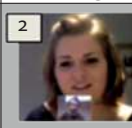
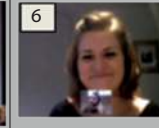
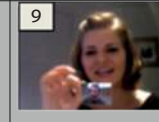
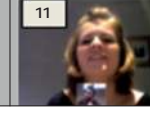
		LEARNER IMAGE	TEACHER IMAGE
1. Learner	The third young people is er a man too has er really short hair and er *(he looks down at the photographs)* *(she is focused on the screen and produces a slight smile)*		
	an ear piercing er::: like a:: (.) *(touches his ear while looking down)*		
	(looks up and looks at screen) it's in like a:: (.) a tube *(makes a gesture to represent a round hole)*		
	(looks up the screen) a bi:: er: *(points to his ear)*		
	(turns his face from the screen) I don't know [what I']		
	(looks at screen)		
2. Teacher	[xx] (.) it's big/ (mirrors learner's gesture (see 4) and looks at the screen with a smile)		
3. Learner	Yeah it's big (.) it makes a hole in his ear:: (touches his ear again and looks down)		
4. Teacher	OK (nods and smiles)		

Figure 9.2 Multimodal transcript of a word search episode

because we felt that the detail of their facial expressions, smiles, and micro gestures within the same turn was necessary to understand minutely the adjustments that occurred during such an interaction. Such a transcript allows a vertical linear representation of turns and makes it possible to unpack the different modes at play "via a zigzagged reading" (Sindoni, 2013, p. 82). Working iteratively on the

transcript and on the accompanying text (see Table 9.2) helps refine both because they force researchers to give saliency to certain features in the transcript (such as simultaneousness of different phenomena or interaction between different semiotic modes), while the text that they write has to deploy textual resources to recount them. Neither the transcript nor the text can stand alone; rather, they function as two faces of the proposed analysis.

Table 9.2 Textual analysis of the episode

In turn 1, certain marks of hesitation, long pauses, and self-admonishments ("I don't know") signal a communication breakdown while the learner is trying to find a way to describe the unknown lexical item. By touching his own ear repeatedly and miming a hole with his fingers, the learner is not only making his search visible to the teacher but is negotiating the meaning with her and looking for signs of her understanding. Her smile in image 6 suggests that she seems to understand what he is trying to describe, although he pursues his description in an attempt to be even more precise. As is visible in image 7, the student has what Goodwin and Goodwin (1986) would describe as a "thinking face," indicating to the teacher that he is still searching for the exact term, before he looks directly at the screen in image 8 – suggesting he wants confirmation from the teacher that she understands precisely what he is trying to describe. This search triggers a smile from the teacher and the mirror gesture (image 9) of that of the learner, which indicates that the teacher ratifies the description to a certain extent and that the interaction can continue while she is giving him her full attention by looking directly at the screen. Once the association of the verbal and nonverbal messages seems to have reached their objective, the learner verbally adds an element ("a hole") and gives redundant information by prodding his index finger at his ear again, making sure that the teacher has understood the lexical item (she nods in image 11), even if the precise word has not been found.

There is no stable way of making multimodal transcripts although researchers have been increasingly devising astute ways of approaching this (see for instance Bezemer, 2014; Flewitt et al., 2009; Norris, 2004; Sindoni, 2013). Reading these authors, several considerations arise in relation to the units of analysis that can be selected, the ethical dimensions that have to be attended to, the readability, and the presentation of multimodal transcription.

First, turns of speech that constitute the conventional unit of analysis in Conversation Analysis may not be as pertinent for multimodal analysis because, as noted by Flewitt et al. (2009), "as soon as multiple modes are included, the notion of speech turns becomes problematic as other modes contribute meanings to exchanges during the silences between spoken turns" (p. 45). New units of analysis must therefore be devised to capture the specificity of multimodal interactions. For example, what is a speech turn when an individual uses written chat and speech simultaneously? Second, a multimodal transcript makes participants identifiable, which makes it even more crucial to be vigilant about ethical

considerations (as discussed above). Finally, researchers must establish a careful balance between the representation of all the features that are to be considered in a multimodal interaction and what a reader – even a seasoned one – is able to capture when confronted with a thick rendering of multimodality. As noted by Flewitt et al. (2009), "the perceptual difficulties for the audience of 'reading' genuinely multimodal transcription might outweigh the advantage of its descriptive 'purity'" (p. 47). Eventually, there will be new ways of presenting multimodal data along with more traditional paper-based publication that will better render the multimodal nature of such data. How to transform multimodal data in order to make them accessible with various degrees of complexity or presentational choices constitutes one direction for future research.

Drawing conclusions

Once transcriptions are completed, researchers can proceed to analyses, such as the one proposed in Table 9.2. If their approach is *quantitative*, all the annotations can be exported to statistical applications that can be used on the results (Wittenburg et al., 2006). Quantitative studies can thus give insight into a certain number of phenomena that can be relevant to understanding online learning and teaching. For instance, the number of pauses, the frequency of overlaps, and the length of turns can shed light on the rhythm of a given interaction, as shown in Study 1 above. The number of gestures and facial expressions produced by the participants could also give indications as to the communication potential of videoconferencing. The main outcome of quantitative studies concerns the identification of interactional patterns.

Although some examples of quantitative studies can be found, studies usually rely on qualitative approaches to data and focus on short episodes. At this point, it is worth mentioning Jewitt's (2009) caveat about CMC researchers, working solely from a qualitative perspective, who may solely produce "endless detailed descriptions" and fail to address broader questions that nevertheless need to be answered (p. 26).

Yet, Adolphs and Varter (2013) point out that the community of researchers interested in multimodal analysis might profit from adopting a mixed approach and combining, when possible and pertinent, the conversation analysis of small samples of data with a corpus linguistics-based methodological approach. Thus, with the inclusion of large-scale data sets, such an approach could extend "the potential for research into behavioural, gestural and linguistic features" (Adolphs & Varter, 2013, p. 145).

Conclusion

In this chapter, we have shown the importance of taking into account the array of technologies (in this case, screen video recording and annotation tools) that accompany the construction, analysis, and transformation of interactional data. With ever-refined software and transcription techniques, interactional linguistics has come to integrate into its agenda the intrinsically multimodal nature of interactions (Détienne & Traverso, 2009). This is even more apparent when the interactions under study are themselves mediated by technologies, as is the case with videoconferencing-based exchanges. Technologies thus facilitate the gathering of interactional data and allow researchers to explore them, replay them at will, annotate them with different degrees of granularity, visualize them from different perspectives, and structure them according to different scientific agendas (Erickson, 1999). Not only do these technologies change the way researchers approach data, they also require them to develop new technical and methodological skills. As we have seen with the various steps involved in the collection, transcription, and analysis of multimodal data, the different techniques at play mostly concern the representation of data. Each transformation of the data results in a new object that can be subject to yet another transformation, until the refinement is complete enough to yield a satisfactory comprehension of the phenomena under study. This points to the essential work of representations that "serve as resources for communicating and meaning-making" to the scientific community and beyond (Ivarsson, Linderoth, & Saljö, 2009, p. 201) and are "achieved by combining symbolic tools and physical resources" (Ivarsson, Linderoth, & Saljö, 2009, p. 202).

The kinds of studies we have conducted not only help us to uncover the interplay of the different multimodal semiotic resources in online teaching environments but ultimately serve to improve the design of teacher-training programmes. For researchers, this includes gaining valid information about how to sensitise teachers to the affordances of the webcam in online interactions by encouraging them to pay attention to learner needs, thanks to visual cues. In so doing, they should develop their semio-pedagogical competence (Guichon & Cohen, 2016), that is to say their awareness of the semiotic affordances of media and modes and their subsequent ability to teach online using videoconferencing.

References

Adolphs, S., & Varter, R. (2013). *Spoken corpus linguistics. From monomodal to multimodal*. New York, NY: Routledge.

Allwood, J., Ahlsén, E., Lund, J., & Sundqvist, J. (2007). Multimodality in own communication management. In J. Toivanen & P. Juel Henrichsen (Eds.), *Current trends in research on spoken language in the Nordic countries* (Vol. 2, pp. 10–19). Oulu, Finland: Oulu University Press.

Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. London, United Kingdom: Equinox.

Bétrancourt, M., Guichon, N., & Prié, Y. (2011). Assessing the use of a trace-based synchronous tool for distant language tutoring. In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 conference proceedings. Volume 1 – Long papers* (pp. 478–485). International Society of the Learning Sciences.

Bezemer, J. (2014). How to transcribe multimodal interaction? In C. D. Maier & S. Norris (Eds.), *Texts, images and interaction: A reader in multimodality* (pp. 155–169). Berlin, Germany: Mouton de Gruyter.

Blake, R., & Zysik, E. (2003). Who's helping whom?: Learner/heritage-speakers' networked discussions in Spanish. *Applied Linguistics*, 24(4), 519–544. doi:10.1093/applin/24.4.519

Brouwer, C. E. (2003). Word searches in NNS–NS interaction: Opportunities for language learning? *The Modern Language Journal*, 87, 534–545. doi:10.1111/1540-4781.00206

Calbris, G. (2011). *Elements of meaning in gesture* (Mary M. Copple, trans.). Amsterdam, Netherlands: John Benjamins. doi:10.1075/gs.5

Cohen, C., & Guichon, N. (2014). *Researching nonverbal dimensions in synchronous videoconferenced-based interactions*. Paper presented at *CALICO Conference*, University of Athens, OH, USA.

Cosnier, J., & Develotte, C. (2011). Le face à face en ligne, approche éthologique. In C. Develotte, R. Kern & M.-N. Lamy (Eds.), *Décrire la conversation en ligne: Le face à face distanciel* (pp. 27–50). Lyon, France: ENS Éditions.

De Chanay, H. (2011). La construction de l'éthos dans les conversations en ligne. In C. Develotte, R. Kern & M.-N. Lamy (Eds.), *Décrire la conversation en ligne: Le face à face distanciel* (pp. 145–172). Lyon, France: ENS Éditions.

Détienne, F., & Traverso, V. (Eds.). (2009). *Méthodologies d'analyse de situations coopératives de conception*. Nancy, France: Presses Universitaires de Nancy.

Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning and Technology*, 7(3), 50–80. Retrieved from <http://llt.msu.edu/vol7num3/pdf/doughty.pdf>

Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford, United Kingdom: Oxford University Press.

Erickson, T. (1999). Persistent conversation: An introduction. *Journal of Computer-Mediated Communication*, 4(4), article 1. doi:10.1111/j.1083-6101.1999.tb00105.x

Flewitt, R., Hampel, R., Hauck, M., & Lancaster L. (2009). What are multimodal transcription and data? In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 40–53). London, United Kingdom: Routledge.

Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62(1–2), 51–75.

Guichon, N., & Cohen, C. (2014). The impact of the webcam on an online L2 interaction. *Canadian Modern Language Journal*, 70(3), 331–354. doi:10.3138/cmlr.2102

Guichon, N., & Cohen, C. (2016). Multimodality and CALL. In L. Murray & F. Farr (Eds.), *Routledge handbook of language learning and technology* (pp. 509–521). London, United Kingdom: Routledge.

Guichon, N., & Tellier, M. (Eds.) (forthcoming). *Enseigner l'oral en ligne*. Paris: Didier.

Guichon, N., & Wigham, C. R. (2016). A semiotic perspective on webconferencing-supported language teaching. *ReCALL*, 28(1), 62–82. doi:10.1017/S0958344015000178

Ivarsson, J., Linderoth, J., & Saljö, R. (2009). Representation in practices. A socio-cultural approach to multimodality in reasoning. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 201–212). London, United Kingdom: Routledge.

Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis*. London, United Kingdom: Routledge.

Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and quality of language production. *The Modern Language Journal*, 79(4), 457–476. doi:10.1111/j.1540-4781.1995.tb05445.x

Kern, R. G. (2014). Technology as pharmakon: The promise and perils of the Internet for foreign language education. *The Modern Language Journal*, 98(1), 340–358. doi:10.1111/j.1540-4781.2014.12065.x

Kost, C. R. (2008). Use of communication strategies in a synchronous CMC environment. In S. Sieloff Magnan (Ed.), *Mediating discourse online* (pp. 153–189). Amsterdam, Netherlands: John Benjamins. doi:10.1075/aals.3.11kos

Lamy, M.-N., & Hampel, R. (2007). *Online communication in language learning and teaching*. Houndmills, United Kingdom: Palgrave Macmillan. doi:10.1057/9780230592681

McCafferty, S. G. (2008). Material foundations for second language acquisition: Gesture, metaphor and internalization. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 47–65). New York, NY: Routledge.

McCafferty, S. G., & Stam, G. (Eds.). (2008). *Gesture: Second language acquisition and classroom research*. New York, NY: Routledge.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: The University of Chicago Press.

Norris, S. (2004). *Analyzing multimodal interaction – A methodological framework*. New York, NY: Routledge.

Ochs, E. (1979). Transcription as Theory. In E. Ochs & B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43–72). New York: Academic Press.

Pellettieri, J. (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practices* (pp. 59–86). New York, NY: Cambridge University Press. doi:10.1017/CBO9781139524735.006

Satar, H. M. (2013). Multimodal language learner interactions via desktop videoconferencing within a framework of social presence: Gaze. *ReCall*, 25(1), 122–142. doi:10.1017/S0958344012000286

Siemens, G., & Baker, R. S. J. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, New York, NY, 252–254. doi:10.1145/2330601.2330661

Sindoni, M. G. (2013). *Spoken and written discourse in online interactions: A multimodal approach*. New York, NY: Routledge.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation*, 816–820.

Stockwell, G. (2010). Effects of multimodality in computer-mediated communication tasks. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 83–104). London, United Kingdom: Continuum.

Tellier, M., & Stam, G. (2010). Stratégies verbales et gestuelles dans l'explication lexicale d'un verbe d'action. In V. Rivière (Ed.), *Spécificités et diversité des interactions didactiques* (pp. 357–374). Paris, France: Rivenue Éditions.

ten Have, P. (1999). *Doing conversation analysis: A practical guide*. London, United Kingdom: Sage.

Wang, Y. (2007). Task design in videoconferencing-supported distance language learning. *CALICO Journal*, 24(3), 591–630. doi:10.1558/cj.v24i3.591-630

White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness*, 11(4), 259–290. doi:10.1080/09658410208667060

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of Language Resources and Evaluation Conference*, 1556–1559.

Yakura, E. K. (2004.) "Informed consent" and other ethical conundrums in videotaping interactions. In P. Levine & R. Scollon (Eds.), *Discourse and technology – Multimodal discourse analysis* (pp. 184–195). Washington, DC: Georgetown University Press.

CHAPTER 10

A scientific methodology for researching CALL interaction data

Multimodal LEarning and TEaching Corpora

Thierry Chanier and Ciara R. Wigham
Université Blaise Pascal Clermont-Ferrand, France /
Université Lumière Lyon 2, France

This chapter gives an overview of one possible staged methodology for structuring LCI data by presenting a new scientific object, LEarning and TEaching Corpora (LETEC). Firstly, the chapter clarifies the notion of *corpora*, used in so many different ways in language studies, and underlines how corpora differ from raw language data. Secondly, using examples taken from actual online learning situations, the chapter illustrates the methodology that is used to collect, transform and organize data from online learning situations in order to make them sharable through open-access repositories. The ethics and rights for releasing a corpus as OpenData are discussed. Thirdly, the authors suggest how the transcription of interactions may become more systematic, and what benefits may be expected from analysis tools, before opening the CALL research perspective applied to LCI towards its applications to teacher-training in Computer-Mediated Communication (CMC), and the common interests the CALL field shares with researchers in the field of Corpus Linguistics working on CMC.

Keywords: LEarning and TEaching Corpora (LETEC), staged methodology, multimodal transcription, OpenData

Introduction

In many disciplines, research is based on the availability of large research data sets, built collaboratively from the work of many different research teams. Data are shared and form the basis for new analyses, or counter-analyses. To meet this demand for data, other researchers develop tools for the research cycle (tools for capturing and analysing data). When studying learner-computer interactions

DOI 10.1075/lsse.2.10cha

(LCI), researchers are concerned by the extent of data collection and by the description of the context in which data were collected. Studying online learning, in order to understand this specific type of situated human learning and/or to evaluate pedagogical scenarios or technological environments, requires accessibility to interaction data collected from the learning situation.

The intention of this chapter is to give an overview of one possible staged methodology for structuring LCI data. It presents a new scientific object: LEarning and TEaching Corpora (LETEC). After having clarified the notion of corpora, used in so many different ways in language studies, the methodology used to collect, transform and organize data in order to make them sharable through open-access repositories is described. We suggest ways in which the transcription of interactions may become more systematic, and what benefits may be expected from analysis tools before opening the CALL research perspective applied to LCI towards its applications to teacher-training in Computer-Mediated Communication (CMC), and the common interests we share with researchers in the field of corpus linguistics working on CMC.

Differentiating raw language data and corpora

Corpora in linguistics

In many areas of general linguistics or even applied linguistics, building and using a corpus is a tradition. A first definition offered by Biber, Conrad and Reppen (1998), following the seminal work of Sinclair (1991) (see O'Keefe et al. [2007] for full references), could be as follows: a corpus is a principled collection of texts, written or spoken, available for qualitative or quantitative analysis. The word corpus, however, may be indistinctly used by a graduate student to refer to her/his compilation of a set of language examples or a set of texts, or by a researcher in corpus linguistics. A similar confusion exists in the Humanities around the word database. Any set of data included in a spreadsheet, or even database software, is often labelled a database, while the second indispensable component of a database, i.e., its conceptual model or semantic level, is ignored. This model, also developed by the data compiler, is often considered as the most valuable component because, firstly, it brings data up to the level at which it may be considered as information and, secondly, because it allows queries and computations to be executed on the basic level of data.

Coming back to language issues, Bernard Laks, a scholar in speech corpora, often underlines the amount of time (over thirty years) it took for linguists to shift from the exemplum paradigm to the datum paradigm (Laks, 2010). At the

end of the fifties, a number of linguists, influenced by Chomsky, rejected the idea of working on corpora (perceived as "limited" in nature) and based their analyses only on sets of language examples, which sometimes were even invented in order to include what they considered as interesting phenomena. Today, many linguists consider that language should be studied in contexts of real usage and, consequently, that corpora are the way to capture language usage.

The nature of corpora and the methodologies for building them have largely evolved from the seminal work of Kucera and Francis (1964) who designed the Brown Corpus as a reference corpus for American English. For example, the DWDS (Digitales Wörterbuch der Deutschen Sprache, 2013) corpus of modern German contains billions of tokens/words. Teams of linguists, who have patiently chosen the various genres that reflect the way German is currently used (including Internet genres), have solved issues concerning rights access and collected the data. Raw data are never compiled as such, but rather transferred into standard formats, based on the eXtensible Markup Language (XML). Researchers developed XML schemas, which play a similar role to the conceptual model of databases. XML is used on top of the texts, sentences and words to add annotations.

Corpora in CALL

The language-teaching domain is also directly concerned with corpora. Launched in the nineties, conferences including TALC (Teaching And Language Corpora) have become popular among applied linguists, and some language teachers are interested in the idea of using different kinds of language corpora in their teaching (O'Keefe et al., 2007). As an example, if German academic writing is considered, linguists may study this type of language for specific purposes (LSP) before updating pedagogical handbooks with language structures that are actually used, or teachers may use the same LSP corpora with learners of German. The latter situation is often referred to as Data-Driven Learning (DDL) (Boulton, 2011), and several special interest groups within the CALL community have developed in this area, as well as dedicated journal issues.

Whereas the previous corpora captured language used in formal or informal situations only by native speakers, a team of linguists gathered in Belgium around Sylviane Granger to launch a new type of corpora, namely Learner Corpora. Productions (mainly academic essays) of learners of English as a second language were collected (Granger, 2004). Here again, the team did not confuse the concept of a corpus with a simple set of essays in electronic formats. They developed a framework for learner corpus research where data were collected, structured and, from 2009 onwards, annotated in the same way. They included productions of learners with different mother tongues to allow interlanguage comparisons.

The corpus paradigm

Taking into consideration the aforementioned lengthy experiences coming from corpus linguistics (whether general or applied linguistics), as well as international recommendations for the management of research data in all scientific disciplines, the corpus paradigm can be developed as follows:

Systematic data collection
Even when an individual researcher has a specific research question in mind, such as a specific kind of interaction s/he wishes to consider, the whole data set, including interactions, productions, logfiles (data related to what is called learning analytics) should be collected. It is a prerequisite to allow other researchers to reuse the corpus. It also relates to quality criteria. Often a researcher selects a subset of data from the whole data set in order to analyse it and publish an article. Quality in the research procedure implies that the researcher is able to explain the extent to which a selected subset of data does not correspond to a simple disconnected episode, but really reflects what happened during the online course.

Detailed data description
The context of learning situations encompasses many facets, as detailed later in this section. In regards to language corpora, in general, the detailed descriptions are often referred to as metadata. In the metadata, the researcher not only gives a corpus title, date, list of credits, but also explains how the data have been collected, edited and organized. Sociolinguistic information about the participants is detailed. As an example, let us consider a SMS corpus. Metadata will explain how messages have been collected on the phone network(s) and anonymized. They will document participants who sent the messages, the structure of the messages assembled in the body/text of the corpus, whether the date of a message corresponds to its date of posting or of collection, and the way in which IDs have been attributed to participants to guarantee that messages sent by the same person can be linked, etc. This information is essential if a researcher wishes to carry out a discourse-analysis study.

Data conversion
Time spent on data collection and description will be valued during the analysis phase. It is now generally considered as a multiple-step process, where output of a first analysis tool will become input for a second tool. Young researchers working on language-related data, whether oral, textual or multimodal (optionally, incorporating non-/co-verbal data), will often have to manage this analysis flow before the publication, for example, of their thesis. This has two main implications: (a) the

use of analysis tools that accept open formats for data input and that do not produce output in proprietary formats, and (b) conversion, organization and structuring of the collected data into standard formats. Besides open-access formats for images, audio or video files, the format for textual data is now based on XML, not simply a basic XML level, but levels of higher standards that allow annotations and multi-level analyses, as detailed further on.

Data release and distribution

As previously explained, a language corpus and its related analysis can only become part of the scientific research cycle if it can be freely accessed and when this access is guaranteed as permanent. Although solutions and access to procedures that guarantee this openness are well known, available and fairly simple, the current situation is blurred by the misuse of the term OpenData (see a relevant definition in Open Knowledge, 2013, as well as Chanier, 2013). If a researcher tries to access language corpora which pretend to be open access, s/he may discover free access to only a limited part of the corpus, or that the corpus cannot be downloaded, or, when it is a speech corpus, s/he may have access only to the transcripts but not the accompanying audio files. Under such circumstances, research on the corresponding data is impossible. However, there currently exist more frustrating situations – for example, when a researcher adds an extra level of annotation and wants to publish this, but suddenly realizes that s/he is not allowed to do so because the licence attributed by the original collectors of the corpus forbids any derivative work. Securing open access intertwines several steps of a corpus' lifecycle. Before data are collected, the researcher will consider the question of ethics and rights related to participants and their productions, choose the licence under which to release the future corpus, and choose in which repository the corpus will be deposited for archiving, for example, at the European level, DARIAH (2013).

Clarifying some terms

Before considering corpora specific to LCI, definitions of terms used in many different ways across the field of linguistics, as well as in other disciplines, need to be elicited (see Chanier et al., 2014).

Firstly, the word text is interpreted here in its broad sense relating to its multimodal nature, with respect to Baldry and Thibault (2006) who defined texts as "meaning-making events whose functions are defined in particular social contexts" (p. 4), and Halliday (1989) who declared that "any instance of living language that is playing a role some part in a context of situation, we shall call it a text. It may be either spoken or written, or indeed in any other medium of expression that we like to think of" (p. 10). Simply stated, learners compose a text when they produce utterances, for example, in an audio chat.

Secondly, an online environment may be synchronous or asynchronous, mono- or multimodal. Modes (text, oral, icon, image, gesture, etc.) are semiotic resources that support the simultaneous genesis of discourse and interaction. Attached to this meaning of mode oriented towards communication, we use the term modality as a specific way of realizing communication as per the human-computer interaction field (Bellik & Teil, 1992). Within an environment, one mode may correspond to one modality, with its own grammar constraining interactions. For example, the icon modality within an audio graphic environment is composed of a finite set of icons (raise hand, clap hand, is talking, momentarily absent, etc.). In contrast, one mode may correspond to several modalities: Text chat has a specific textual modality that is different from the modality of a collective word processor, although both are based on the same textual mode. Consequently, an interaction may be multimodal because several modes are used and/or several modalities (see also Chapter 9, this volume).

After having considered criteria for general types of language corpora, the next section presents criteria specific to LCI illustrated by the LETEC approach.

An illustration of the staged methodology for building LETEC

The LETEC approach to data collection, structuring and analysis comprises successive phases (Figure 10.1). It has been developed from 2006 onwards by the Mulce project (Reffay et al., 2012). Using a case-study approach, this section describes these phases in turn, referring to the example of the online English for Specific Purposes course, Copéas, and its associated LETEC (see Chanier et al., 2009). This ten-week intensive course ran in 2005 and formed part of a Master's program in Distance Education in France. The course's aims were for students to be able to think critically about using the web for learning and to practise their oral and written English skills online. Each week, the students participated in online tutored discussions in the online platform Lyceum.

Lyceum is an audio-graphic conferencing environment that included communication modalities (audio chat, text chat, iconic system) and shared editing modalities (whiteboard, concept map, shared word processor). For the reasons already given, it was a multimodal environment, as shown in Figure 10.1, and explained in Ciekanski and Chanier (2008). Lyceum no longer exists. However, thanks to the availability of LETEC data, the environment's features, as well as how participants used it to work and communicate, can be studied and compared to other environments.

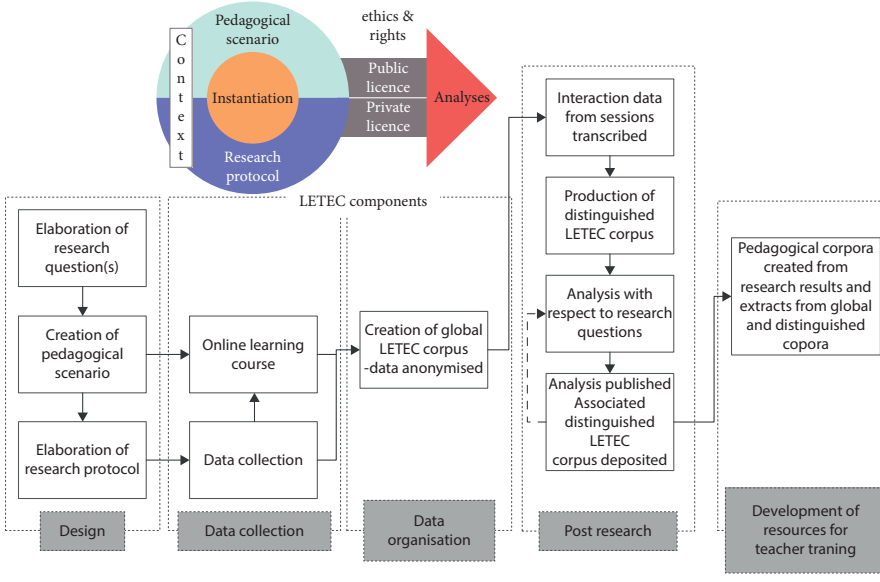


Figure 10.1 Successive phases of a LETEC approach to an online learning situation. LETEC components are illustrated in the top-left hand schema

Design: Pedagogical scenario and research protocol

The first stage of a LETEC methodological approach is to determine the research focus. That is to say the type of phenomenon concerned and the aspect that is of interest. At this stage, it is important to imagine the possible end product that is initially intended. The Open University (2001) has examined a range of general purposes for conducting educational research: to describe, explain, predict, evaluate, prescribe and theorize (p. 30). Identifying a clear research purpose will influence how the research questions are formulated, the type of data to be investigated and how the researcher can select these. Although the research focus will be determined at the beginning of the research process, it is important to note that research questions may not be formulated until later on, or, if formulated during the design phase, they may be modified in between the LETEC design stage and the post-research analysis stage and will most likely become more focused.

In parallel to determining the research focus and specifying the research questions, the online learning context in which it will be examined needs to be elaborated. The design of an online learning situation requires the creation of a pedagogical scenario. This describes (a) the whole online environment (such as a Learning Management System [LMS], a videoconferencing system and their different subcomponents); (b) the various roles the participants (teachers, learners,

or experts, such as native speakers) will undertake during the course; (c) each course activity and the role of each participant during this (e.g., one learner may act as a group animator/tutor) and the component of the online environment the activity is linked to; (d) how activities are sequenced (the workflow); (e) the resources that will be used and produced; and (f) the instructions that govern the learning activities. To avoid confusion between the role of the participants who are involved in supporting the learners and the learning tasks, the pedagogical scenario may consist of a learning scenario and a tutoring/supervision scenario, the latter detailing how different participants will aid learning and how teachers/tutors will intervene during the course in supervisory actions. Put simply, the pedagogical scenario will answer the question of *who does what, when, with what tools and for what results* (see IMS-Learning Design in IMS-Learning, 2004).

If the online learning situation is to be the focus of a research study, it is also necessary to elaborate a research protocol. This will take into account the variables to investigate, the participants in the study, human subject ethical protections, the methods and procedures to be used for data collection and any reliability or validity of collection methods. In relation to the pedagogical scenario, the research protocol details moments at which activities uniquely related to the research will be accomplished (e.g., consent form distribution, pre- and post-course questionnaires, post-course interviews). If observation is to occur, the role of the researcher(s) will also be determined.

The pedagogical scenario and the research protocol could be described as a simple text and assembled with all the documents (pedagogical guidelines, instructions given to teachers, learners, questionnaires forms, etc.); however, this description has to be detailed. It represents more than the usual context of interactions. Research in CALL studies the influence of the learning situations on the interactions and their outcomes. Hence, scientific investigation can commence only if the learning context is explained in a way that a researcher who did not participate in the course could understand the situation. This is why it is recommended to use standard[1] formats for describing these elements, particularly formats that allow visual presentations of the pedagogical scenario, the research protocol and that allow links to resources (IMS-Learning, 2004).

---

1. The word *standard* is frequently used in this chapter to refer to formats that are shared among academic communities to describe different levels of information within corpora. When large sets of communities agree upon a standard, it may become an international *norm* (such as those used by ISO – International Standard Organization). Useful standards generally need to be open (not attached to proprietary formats) and accepted by a wide range of software analysis tools (asset often called *interoperability*).

Data collection

After planning the online learning situation and the research design, the next phase is to systematically gather the data. Data collection focuses on acquiring information, in an ethical manner, to attempt to answer the research questions elaborated during phase one of the LETEC approach and with reference to the research protocol established.

This phase has to be carefully planned beforehand. Earlier on, we mentioned decisions that have to be made before collection and which may influence other phases: interaction data may be difficult to extract from some environments but easier from others that have the same affordances; data formats generated by the learning environments or from other recording devices (audio recorder, screen capture software, etc.) should be easy and not too time-consuming to handle in the next phase. They should have standard output formats or formats that are easy to convert to these; questions of ethics and rights should have been cleared, and consent forms which clearly indicate future corpus use (see the section hereafter) should be distributed and signed. Zourou (2013) provided a good example of obstacles which may be encountered when collecting data stemming from informal learning situations, such as: Who owns user data in these communities? How accessible is user data? What are the consequences of data ownership and accessibility for research purposes?

Data organization

In this section, we present one way to transform raw data into research data, how to organize them and how to document them in an exhaustive yet informative manner. Besides folders of data coming from the above-mentioned learning design and the research protocol, we detail those gathering participants' productions, ethics and rights information, and the overall organization of the corpus (entitled a global corpus). Later, another corpus type is presented (distinguished corpora), which can be derived from the global corpus after research and analyses have been performed.

Course instantiation

The pedagogical scenario could be perceived as a kind of model of a course, an "abstract class," as phrased in object-oriented languages. When the course takes place, participants (individuals, groups) bring to life this model, i.e., it becomes an

"instantiation" of the class. Of course, during this "live" course, events may differ from what was originally planned.

The instantiation component is at the heart of the corpus as this folder regroups all of the data elicitation (Mackey & Gass, 2005). These data are derived from the learning context: all of the participants' productions, including the interaction tracks recorded during the online course. For the Copéas course, this folder includes screen capture videos of the online sessions in Lyceum and the students' reflective reports about the course.

Before regrouping the produced data, a preliminary treatment phase is necessary. Firstly, each resource receives a unique identification code (ID) so that later, in the corpus structuration phase (see hereafter), they can easily be listed and described. A strategic policy is to define IDs which uniquely identify a resource among a set of corpora, e.g., a participant ID may contain the name of the student group to which s/he belongs, the corpus name and course session name – if it is a recording, its mode (video, speech, etc.).

Secondly, all produced data are anonymized through a systematic process. In the Copéas corpus, full names of participants were replaced by participant codes. It is preferable to create meaningful codes which will facilitate data investigation later on. A code can refer to such an aspect as the role of the participant in the course (tutor, student, researcher), his/her gender, or his/her group ID. One should provide a table that regroups the code, sociolinguistic information, language biography (foreign languages spoken, language level, number of years spent studying the language and context of study) for every participant. Anonymization also includes modifying any information in the produced data that could lead to the identification of a participant or skew a researcher's analysis of the data. While it is important to anonymize the data, researchers should replace it with meaningful information. It is useful to include the reasons for anonymization so as to allow interpretations of the interaction. For example, a participant's phone number in a text chat message could be replaced with a code and labelled to highlight that the original information corresponded to a phone number.

Lastly, for the sake of medium and long-term reusability, data collected will be converted into formats independent of their original platform, when the original formats were not open. Several international research associations, including CINES (2014) and Jisc (formerly the Joint Information Systems Committee), involved in the curation and archiving of research data provide clear information about such formats.

Expectations are even greater in regards to participants' interactions that are in text mode, either originally because they have been typed by participants or as the result of transcriptions of speech, for example. Their format will be machine-readable, even structured in order to detail information about an utterance

or a message and relate it to the properties of the environment that integrates this modality. For example, when an LMS includes a discussion forum, every message carries information, such as the author's ID, date of posting, title, message contents, thread, forum name, etc. Rationales for these expectations are related, firstly, to research analysis.

Ethics and rights for OpenData

Releasing a corpus as OpenData means allowing other people the possibility of free use, reuse and distribution. In other words, the user may extract part of the researcher's data, mix this part with data from other sources, add her/his own work to build upon the whole data set and distribute the entire result. Therefore, OpenData relies on two sorts of rights – those related to the data collection and those related to the data release. In other words, data collected need to be free of rights, and secondly the corpus creator should give the right to use the corpus to the end-user, through a licence that imposes minimal constraints. Indeed, internationally it is even recommended to avoid putting a licence that forbids commercial use and to waive intellectual property rights (IPR) (Open Knowledge, 2013). Waiving IPRs does not imply that the creators will not be cited or acknowledged. The full bibliographic reference of their work will become prominent in the corpora repository and will guarantee, in the academic world, that end-user researchers can clearly refer to the original creators when submitting their new analysis to a peer-review process.

Collecting data that are free of rights implies that the compiler him/herself has the right to use the resources included in the corpus and that participants waive their rights on what they have produced. Their agreements are obtained once they have individually signed a consent form, distributed after an "enlightenment" procedure (see Mackey & Gass, 2005). During this procedure, researchers have an open discussion with participants, where they explain drawbacks and benefits that may be expected from the course and the research process. For example, for research purposes on gestures, participants can give permission to be directly video-recorded without any post-process blurring. They will also be aware that if they change their minds, they can at any time ask for data that concerns them to be removed from the corpus (see Chapter 9, this volume).

The LETEC component that concerns Ethics and Rights contains two distinct parts. The private subfolder regroups all of the informed consent forms signed by the course participants, with contact information. This set of data is not included in the final version of the corpus but rather, due to its confidential nature, is conserved by the corpus compiler. In the second part, the corpus compiler includes

a blank example of the informed consent form signed by course participants, besides the corpus licence that details the conditions under which the corpus may be distributed (such as Creative Commons [2015] licences).

Organization of the global corpus

Once the four corpus folders (instantiation, research protocol, learning design, ethics and rights, see LETEC components Figure 10.1) have been organized, with preliminary treatment phases accomplished on the data where necessary, a general document is created. It contains descriptions of each corpus part and crosslinks pieces of information between the different parts (e.g., between the interaction data, research protocol and learning design). It also provides a full index of the resources collected. Each resource is listed, using the previously introduced resource IDS, and a summary of the resource's contents is given. This will help corpus end users determine what data to use, with relation to their specific research question(s).

Lastly, out of the global description, a short corpus description will be extracted so as to provide metadata in formats that website harvesters can recognize and save. The Mulce repository (2013) chose the format created by OLAC (Open Language Archives Community). It is compatible with European CLARIN standards for metadata. This means that metadata concerning all LETEC corpora, including bibliographic citations, appear in these international linguistic resource banks and can be searched for by Internet users.

Post research data and component

Post research often concerns transcriptions of multimodal interactions, in ways that will be presented below. These transcriptions produce a new set of data which will be assembled into a new LETEC, of a distinct type called a distinguished corpus (Reffay et al., 2012). Its size is much reduced, and corresponds to data assembled and produced by a researcher when s/he focuses on a specific research question and aims to publish an article on the specific topic.

A distinguished corpus includes a particular transformation of a selected part of the global corpus – for example, the transformation of a video file into an XML/text file of the transcribed interaction data and its associated metadata. Following transcription, data analyses can be performed. Data from the global corpus are not copied, but instead referred to, and the newly distinguished corpus only adds the transformed data for the specific analysis.

Distinguished corpora help sustain CALL research by valuing the analyses performed by the researcher. The data used for analysis can be presented in parallel with the analysis' results, and distinguished corpora can be cited and referenced in conference papers or published articles. Readers of a researcher's analysis can examine, or reuse the data analysis performed, whilst reading the report of the results.

Corpus publication

Once the content packaging of the corpus is finished, the compiler deposits the corpus in a repository that adheres to the requirements discussed earlier. This server will provide to the user open access to the information about each corpus stored in the repository with search facilities. It will be connected to harvesters so that its bank of metadata is searchable through each different harvester. It may also offer services such as permalinks to each corpus and data subset, which will identify them in a unique and permanent way.

The Mulce repository (2013) gives access to fifty LETEC corpora coming from more than ten different online learning situations that took place between 2001 and 2013. In May 2012, its size was the following: more than one million tokens, coming from 12000 audio turns, 17000 text chat turns, 3000 blogs, 2000 emails, 2700 discussion forum messages, plus more than 9000 non-verbal acts. The Mulce repository also gives access to more than 200 videos of online interaction sessions. These interactions were produced in a variety of environments (such as LMS, audio graphic systems, or 3D environments), by groups of learners from different countries, following a range of different pedagogical scenarios. A step-by-step tour of the repository is provided in the article entitled "Discovering letec corpora" on the Mulce documentation (2015) website. Needless to say, Mulce encourages other CALL researchers to deposit their corpora in the repository, provided they meet the general criteria outlined here, even if they do not exactly follow certain technical details to which the authors have alluded. Help and discussion will be offered to the depositor.

LETEC contributions to CALL research

The purpose of this section is to present how research on LCI may benefit from the existence of open access corpora. Research is a circular process. For example, LETEC corpora in the Mulce repository have been built out of online learning situations, starting more than thirteen years ago. Data have been reused several times and will be mixed into projects independent of Mulce as discussed previously.

Let us start with one of the very first steps in examining online multimodal interactions (see also Chapter 9, this volume), i.e., transcriptions of components of the instantiation part of a LETEC.

Separating transcriptions and analysis steps

Multimodal transcription is a topic discussed across disciplines, for example in Flewitt et al. (2009). In this article, the authors cite Baldry and Thibault who suggest, "multimodal transcriptions are ultimately based on the assumption that a transcription will help us understand the relationship between a specific instance of a genre, for example a text, and the genre's typical features" (as cited in Flewitt et al., 2009, p. 45). A straightforward interpretation of this statement may induce the idea that all approaches to multimodality should produce their own specific methodological approach to transcriptions. Indeed, the article illustrates various transcription methodologies, from several researchers, that adhere to distinct models of multimodality; in an ad hoc fashion, parts of texts, images, photos and hand-made pictures are intertwined in formats such as spreadsheets and word-processing documents, forbidding any kind of comparison or mixing of data. This interpretation of transcription confuses two steps in the research process – the actual transcription and the analysis.

Researchers involved in national or international consortiums on speech and/or multimodal corpora have special interest groups around interoperability (e.g., Huma-Num, 2014). The idea is that if one understands research as a cumulative process, idiosyncratic models need to be compared in order to enhance understanding of human interactions. This implies separating the transcription from analysis processes and using a variety of analysis tools with compatible output formats.

Figure 10.2 illustrates this point. It displays a window from the transcription software ELAN (Sloetjes & Wittenburg, 2008) that integrates the video screen capture of a Copéas session (red box). In this extract, three learners are working in a sub-group to complete a quiz provided by the tutor at the beginning of the session. The tutor comes into the virtual room while one learner is writing an ESL definition using the word processor. Several modes or modalities are being used: audio, text chat (label [3] in the red box) and the word processor (1), plus the iconic system (2), which lists the participants, their status, indicates who is talking, and allows simple communication (agreement, disagreement, raise hand, applause, etc.). The transcription process appears in the green box. According to the transcription code used (see Wigham & Chanier, 2015), the researcher defined one layer per participant and per modality (5), i.e., all Learner 1's text chat

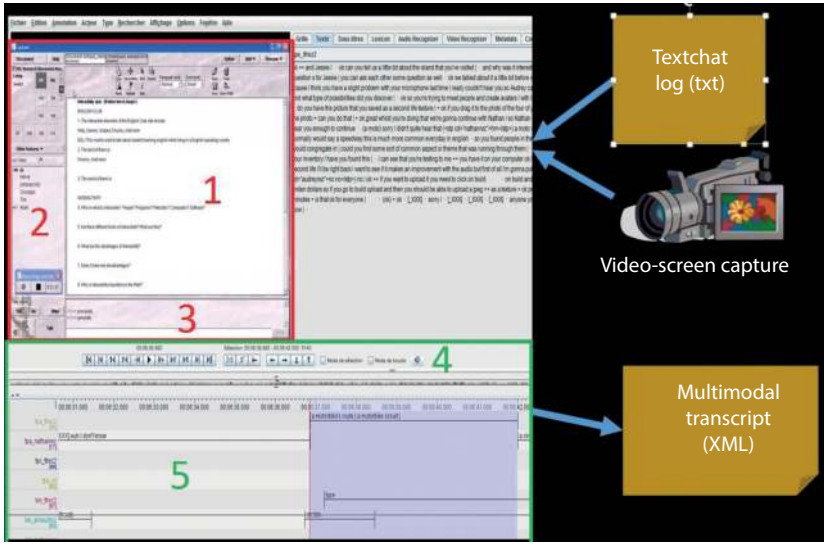


Figure 10.2 Transcript of a Copéas session through the software ELAN, with input and output files

turns are assembled on the same line, all Learner 1's audio turns on another line, and this is the same for the transcription of his/her actions in the word processor. Transcription is aligned with the video's time, and buttons in (4) provide different ways of selecting parts of the video and of moving between transcriptions layers. Once the transcription is completed, its contents are saved using a text-structured XML format that offers the possibility of later compiling it with transcriptions of other sessions from the same course and/or reusing the file with analysis software.

ELAN is a good example of open-access software. This asset, plus the interoperability one, allows any user, once the distinguished LETEC corpus has been downloaded, to rework on the transcription and add another layer, for example. It is largely used in the aforementioned community on multimodal corpora.

There is an even subtler methodological question where transcription is concerned: Are online interactions so complex that it is impossible to compare and make adjustments between transcription codes? Let us take an example and consider the code defined when transcribing online learning sessions in 3D environments where participants interact using avatars (Wigham & Chanier, 2015). Shih (2014) provided another approach to the same topic. Are these legitimate differences? Possibly, because it is a new area of research in CALL, where researchers have recourse to a variety of nonverbal communication frameworks. However, if CALL research aims to become more systematic in this area, then the situation may evolve in a manner similar to the area of speech corpora. Whereas textbooks

in second language acquisition or discourse analysis (e.g., Schiffrin, 1994) still give the impression that idiosyncratic codes for speech transcription are a normal methodological approach, a community of linguists specialized in speech corpora has developed a common way of transcribing speech (e.g., the CHAT format used in the open access CHILDES repository, MacWhinney, 2009) and has even included it in a more general framework designed for different text genres called TEI (Text Encoding Initiative [TEI, 2015]). With this extension of XML, a researcher who focuses on a new oral feature may code a new phenomenon whilst being compatible with the rest of the original coding scheme.

Analysis tools and conditions for scientific discussions

Resuming our Copéas example, let us now consider its analysis. Some of the questions the research team had in mind were: Do participants get lost among the multiple possibilities offered by this type of multimodal learning environment? Do they make consistent individual choices? Can they also make collective choices? In the particular sequence alluded to in Figure 10.2, the workload is distributed among the three learners: one learner types in the shared word processor in order to answer the quiz, and the two others help him orally. Whilst they hesitate on the spelling of a word, the tutor came into the room and typed his corrections into the text chat. This went unnoticed by the learners, and, in turn, the tutor leaves the room. Ciekanski and Chanier (2008) have explained the notion of context which is dynamically built by participants. Relying on this notion developed by Goodwin and Durranti in 1992, their analysis explained that the tutor had been out-of-context. Interestingly, Lamy (2012) imagined the same kind of situation, without referring to any precise data:

> Imagine that the tutor led his tutorial via postings in the text-chat while students talked about other topics in the audio channel. It is unlikely that the group would accept such a position for the tutor, and we draw from multimodal social semiotics to help explain why. (p. 12)

Discussing alternative explanations with different theoretical references is a very important issue in research, provided that data and analysis tools support it. Figure 10.3 illustrates our analysis with the open-access tool TATIANA (Dyke et al., 2011), for analysing online interaction from a Computer-Supported Collaborative Learning (CSCL) perspective. To the left of the red line, one can see the same video (top left) with the transcription (bottom left), simply converted from the ELAN-XML output to the TATIANA-XML version. On the right-hand side of Figure 10.3 appears another view of the desktop, with a view of the modalities used by each participant: one line per participant, one colour per modality (text

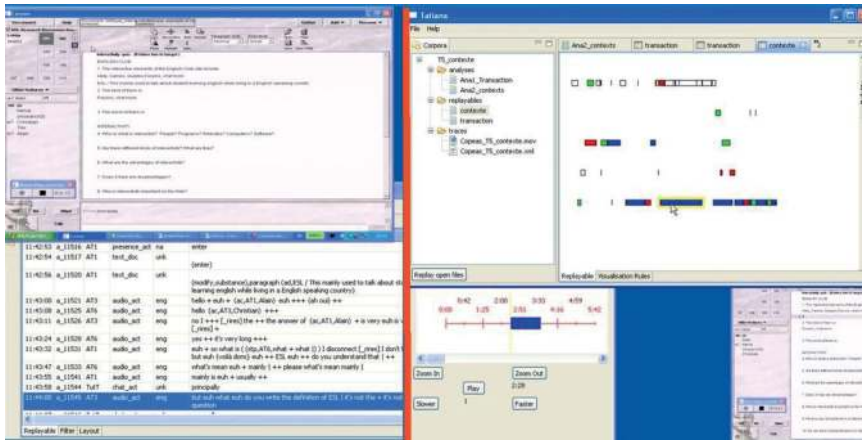


Figure 10.3 Being in and out of context in a multimodal environment. Follow up of example 2 analysed, thanks to the TATIANA software

chat, audio chat, word processor, etc.). This display helps visualize that participants may be out of context, that learners used the word processor in combination with other modalities, which highlights the strategic use of certain modes to facilitate the writing process. The learners also made consistent individual choices to participate in multimodal discourse and to make collective choices. Of course, this analysis has been achieved by examining the whole session, not only the aforementioned extract. The comparison with other sessions and several tools has been explained in Ciekanski and Chanier (2008). The analysis was possible because the output of a first transcription tool became input for a second analysis tool.

How opposite conclusions could be compared

Sindoni (2013) also studied participants' uses of modalities in online environments that integrate audio, video and text chat. She focused on what she termed "mode-switching" when a participant moves from speech to writing or the other way round. She collected dozens of hours of video screen online conversations that occurred in informal settings (hence not connected to a learning situation). When analysing transcriptions, she observed that participants could be classified according to their preferred interaction mode (oral or written). She also observed that "As anticipated, both speakers and writers, generally carry the interaction forward without mode-switching. This was observed in the whole video corpus" (Sindoni, 2013, Section 2.3.5). Hence, she concluded, "those who talked did not write, and those who wrote did not talk. Turn-taking adheres to each mode" (Sindoni, 2013, Section 2.3.5).

In analyses of the Copéas corpus, learners had a preferred mode of expression (oral or written), at least when they were of a beginner level. In contrast with Sindoni (2013), analyses of audio graphic and 3D environments show that learners were mode-switchers (even modality-switchers). Choices of mode/modality depended on the nature of the task that had to be achieved and on the tutor's behaviour (e.g.,Wigham & Chanier, 2015).

At this stage, one may expect that scientific discussions could take place between researchers studying online interactions, to debate contradictions, fine differentiations of situations, tasks, etc. In order to allow this, data from the different approaches need to be accessible in standard formats, with publications clearly relating to data and data analyses, and explicit information given about the format of the transcriptions, their codes and transcription alignments with video. However, Sindoni's (2013) data are not available. The inability to contrast data with other examples available in open-access formats is still holding back the scientific advancement of the CALL field.

Coming back to the topic of analysis tools, a researcher who has collected and structured her/his data now has at her/his disposal a wide, rapidly evolving range of tools for lexical processing, morpho-syntactic tagging, statistics, discourse analysis, etc. Should the researcher choose open-access tools with interoperable formats, s/he not only paves the way for circular, multi-analysis research processes but also contributes to the development of these tools; the teams of researchers who developed them are keen to improve them when confronted with requests based on actual data and analysis attempts. This interface between data-collection and analysis tools is at the heart of what Gray calls "e-science" (cited in Reffay et al., 2012, p. 12) and represents a priority in many different disciplines within the Humanities.

LETEC contributions beyond research in CALL: CMC training for language teachers and linguistics

The need for pedagogical corpora

Extracts of LETEC are currently being developed into resources to train language teachers in how to use CMC tools in their teaching practices. Training teachers out of authentic situations, built upon multimodal materials, is not simply a concern of the language learning field. Wigham and Chanier (2014) have detailed the extensive experience of the use of classroom video footage in teacher preparation and professional development in face-to-face contexts coming from the fields of physical education, educational sciences, and mathematics, and described the

production of several classroom footage video libraries. In the video libraries cited, the resources include two different types of data: (a) raw materials collected during the learning situation (curricular, student work, course planning, instruction and assessment resources), and (b) other *records of practice* (Hatch & Grossman, 2009). These resources include post-course interviews with teachers or observation notes made by researchers or trainee-teacher mentors during the class that was filmed. The aim is to give video viewers a sense of what the video footage may fail to capture or details that may have been obscured.

Whilst in other fields, importance is given in teacher training to combining raw materials from experienced teachers' classrooms with research materials, within CALL, CALL-based teacher education is most often delivered through confrontation with research findings and action research (Guichon & Hauck, 2011).

In the first approach, when trainers want students to gain skills in developing online learning situations based on interactive, multimodal environments, they have recourse to the reading of CALL literature disconnected from actual data. Pre-service teachers will not necessarily take the time to question the findings, taking research conclusions as a given. Indeed, the development of an analytic approach to the reading of research literature takes time, and during training courses, educators do not necessarily have enough time for this process to mature.

The second approach focuses on action research with pre-service teachers participating in experiments and adopting either the role of learners or tutors. Here there is either the assumption that trainees will naturally understand what they need to do or, if greater guidance is given, reflective feedback sessions are often conducted with the trainees. In the latter case, attempts to use the same methodology for both data collection and training purposes are often difficult to manage; trainers face the issue that student materials are often heterogeneous and quickly extracted from the on-going experiment, and pre-service teachers may only consider his/her individual practice.

In the CALL field, training pre-service teachers in CMC out of online learning situations, built upon multimodal materials (carefully analysed with respect to theoretical viewpoints), alongside other records of practice/research data and findings, would be very helpful. Therefore, from the notion of LETEC, which are purely used for research investigations, arose the notion of pedagogical corpora.

An example of pedagogical corpus

Each pedagogical corpus includes a selection of materials from a LETEC corpus and a series of structured teacher-training tasks that have been developed from these materials, based on leads that had been identified in research papers for

which the analyses utilized the same data. To illustrate this concept, let us look at a pedagogical corpus, entitled *reflective teaching journals* that was developed from the research Copéas corpus (Wigham & Chanier, 2013).

From the course data and research articles about the project, the need of encouraging trainee-teachers to foster reflective practice through the writing of teaching journals was identified. Journal writing is a prerequisite for developing reflective practice, but it is not a sufficient condition. It only offers a one-sided view of the course situation. A more objective standpoint may come from confronting the journal with other perspectives. In order to make pre-service teachers aware of this situation, the pedagogical corpus focuses on tutors' and students' differing views of successful or unsuccessful collaboration and different perceptions of their online course. The objectives of the corpus are for trainee-teachers to do the following:

- identify language tutors' and students' differing views of successful online collaboration;
- summarize the characteristics of successful collaboration and produce a list of implications for practice;
- appraise the advantages of keeping teaching journals; and
- compare and contrast reflections from a teaching journal with naturally occurring data (interaction tracks) and researcher-provoked data (student feedback) to analyse whether teachers should base reflections about teaching practice solely on journal entries and personal reactions.

In the pedagogical corpus, the corpus users are guided through a series of reflective activities based on personal experience, extracts from the LETEC: interaction data (audio and video-based), learner questionnaires and both learner and tutor post-course interviews. The online corpus gives the instructions for all tasks, the timing guidelines and suggested student groupings. All tasks can be completed either online or in a teacher training classroom. Figure 10.4 shows a sample task in which users identify characteristics of successful collaboration through the tutor's discourse, using extracts of the reflective journal the tutor kept throughout the Copéas course and an extract of the audio post-course tutor interview.

Such pedagogical corpora offer a kind of expert viewpoint (but an expert viewpoint based on research analysis, i.e., coming from a scientific research cycle). Practice in teacher training, coming from the aforementioned fields, shows that it is not enough. Students need to bring their own data (extracts of live sessions and reflective writing) in order to confront these with expert views and other views from classmates as well, the whole process being integrated into a discussion framework, whether online (Barab, Klig, & Gray, 2004) or face-to-face. Furthermore, it cannot be a one-shot process but a progressive one. Becoming a

Activity 3.1
First of all, consult the following resources (rtjounrals-int-TutT-ext1-mp4, rtjounrals-int-TutT-ext2-mp4) that present the tutor's impressions of whether the activities he proposed were collaborative or not. In your notebook, take notes about the characteristics of successful collaboration the tutor gives. Remember that any points he gives about unsuccessful collaboration can be turned on their head to provide pointers for successful collaboration. What reasons does the tutor give for them? Note any examples he gives to illustrate the characteristics you have identified. Do any of the characteristics match those you listed in activity 2?
Resources:
– rtjournals-diary-TutT-pdf This is the tutor's journal that he kept throughout the Copéas course and in which he reflects about tutoring the course online. The journal is in English.
– rtjournals-int-TutT-ext1-mp4 This is a mp4 video of an extract of the audio post-course tutor interview with slides to guide the viewer. A researcher in French conducted the audio interview. The slides are in English. The video lasts 10 minutes 30 seconds.

Figure 10.4 Sample task from a pedagogical corpus (Wigham & Chanier, 2013)

teacher implies moving from a peripheral participation to a more centred one, and this process must be recognized as legitimate by the community (see Lave & Wenger, 1991). Of course, the teacher training period will not suffice, but the idea is to involve students in a rich process during which they confront expert and novice viewpoints.

Currently, two pedagogical corpora have been developed from two different global LETEC corpora. They can be downloaded from the Mulce repository. They have not yet been used to train teachers. For another approach to using corpora in teacher training, see Chapter 9, this volume.

From learner- to general user-computer interactions

In this chapter, several references have been made to works and methodologies adopted in linguistics, or corpus linguistics, which influenced CALL research on data. Is this a one-way flow? Does CALL have something to say that could benefit the linguistics field in general? A first refinement of the question could be: Do the language, discourse and texts produced by participants (learners, teachers, etc.) bear similar features (apart from the obvious differences due to the development of the learners' interlanguage, their errors) to those studied in general by linguists interested in computer-mediated discourse?

In order to answer the question, let us consider one type of environment, for example text chat. In the field of linguistics, descriptions of texts and language

exist in prototypical works, such as Crystal (2004) in the chapter "The Language of Chatgroups" and its section on synchronous groups. This study aims to give a very general overview of what is actually "the Language of the Internet" as reflected by the book's title. However, when considering text chat coming from CALL, the contents of the turns are strikingly different on both lexical and syntax levels (lexical diversity, use of emoticons or other interaction terms, structures of clauses, of utterances, turn lengths, etc.). The discourse organization is also very different. Whereas nicknames play an important role in informal text chats where users constantly change their nicknames in accordance with their current activities, moods, etc., this phenomenon rarely occurs in learning situations. Turns and their combinations (exchanges, transactions, etc.) are managed and structured in a very different manner. In order to support language production in an L2, turn-taking conventions are often adopted.[2]

Considering another mode would bring us to the same conclusions. For example, when skimming through corpora where speech is used, either in bimodal environments (text and audio chats) or in richer environments (audio graphic conferencing systems, 3D environments), discrepancies with informal L1 online conversations can be noted concerning a variety of features. To take but one example, speech overlaps in turn taking are not frequent in learning situations. Rationales explaining these differences in the different modes are quite obvious; language teachers organize scenarios beforehand, and tutors interact in ways that support language learners' productions, helping them take risks in a new language while simultaneously alleviating other tasks. CALL research has also begun to show that the orchestration and use of modes and modalities are different to non-educational situations, as previously exemplified in the discussion of Sindoni's work. To some extent, it could be said that multimodality can be "decomposed" to allow some specific modes and modalities to be used in order to focus on specific tasks (for an example, see the focus on writing in Ciekanski & Chanier, 2008). To sum up, the CALL experience of online interactions, supported by its specific corpora, can be of general interest to the whole linguistic community.

A common model of CMC interactions

Common interests between CALL and corpus linguistics also concern more abstract levels, such as models of online interaction. Following lessons learnt from

2. The reader interested in comparing such differences could access, for example, an informal text chat corpus from Germany (Dortmund Chat Corpus, 2003–2009) or a CALL text chat corpus (Yun & Chanier, 2014).

the Mulce project (Reffay et al., 2012), researchers are now collaborating with corpus linguists. At a national level, the CoMeRe project (Chanier et al., 2014; CoMeRe, 2015) has brought together corpus linguists and CALL researchers. The acronym (in French) stands for network-mediated communication, an extension of CMC, in order to include communication through phones, networks and devices. The CoMeRe project has built a kernel corpus in French that represents a variety of network interactions. Several LETEC corpora have been included and structured in the same model alongside corpora of SMS, tweets, Wikipedia discussions, blogs and text chat interactions. The whole set of corpora are released in an open-access format.

The CoMeRe team is also working with European researchers specialized in CMC to develop the Interaction Space model (TEI-CMC, 2015) through which to structure these interactions. Briefly, an Interaction Space is an abstract concept, located in time (with a beginning and ending date with absolute time, hence a time frame), where interactions between a set of participants occur within an online location. The online location is defined by the properties of the set of environments used by the set of participants (e.g., Chanier et al., 2014). Thanks to this model, corpora from learning and non-learning contexts can, on the one hand, use the same set features to describe the structure and properties of the environment where interactions occurred, the participants (individual, groups), the method for collecting data, for measuring time and durations, etc. On the other hand, in the body of the corpus, the interactions are listed in formats corresponding to their modes (written, oral, or non-verbal). The model is designed by a European group which aims to extend the text model of the Text Encoding Initiative (TEI, 2015) (currently very rich as it encompasses types such as manuscripts, theatre, literature, poems, speech, film and video scripts, etc.) in order to integrate CMC.

Conclusion

When studying LCI in ecological contexts, there are a number of variables that cannot be controlled. These variables make the comparison of scientific results difficult and the replication of a given learning and teaching experience near impossible. This chapter proposed one possible staged methodology to structure raw data from LCI situations into corpora so as to render them comparable, re-analysable and available to the whole research community. The case-study approach adopted allowed us to present the constitution and diffusion of LEarning and TEaching (LETEC) Corpora, using the example of the online Copéas course. In this presentation, we examined the ethical implications of producing corpora

as OpenData and suggested ways in which the transcription of LCI and their analysis can become more systematic and comparable.

The LETEC methodology is one methodological proposition to help the CALL field better meet the principles of scientific validity and reliability that are fundamental cornerstones of the scientific method, yet difficult to achieve in ecological learning situations. More systematic organization of data and its processing is often perceived as time-consuming. However, it requires a mind-set shift whereby individual researchers do not think of producing one-off analyses on individual learning situations but instead look towards long-term team research projects in which corpora, rather than data, are re-used for new analyses, produced from different perspectives, and are reconsidered and cross-referenced from one LCI experiment to another. This would encourage, firstly, a more circular and multi-analysis research approach within the field and, secondly, scientific debate within CALL and more largely within corpus linguistics, which is based on the possibility to reanalyse, verify and extend original findings and to contrast data with other examples from other research teams and different online environments.

References

Baldry, A., & Thibault, P.-J. (2006). *Multimodal transcription and text analysis.* London: Equinox.

Barab, S. A., Kling, R., & Gray, J. H. (Eds.). (2004). *Designing for virtual communities in the service of learning.* Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511805080

Bellik, Y. D., & Teil, D. (1992). Définitions terminologiques pour la communication multimodale. In *Les actes des 4èmes Journées sur l'ingénierie des interfaces Homme-Machine*, IHM'92, Telecom Paris Publ., Paris, 30 Nov. – 2 Dec., 229–232. Retrieved from <https://perso.limsi.fr/bellik/publications/1992_IHM_2.pdf>

Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Gozdz-Roszkowski (Ed.), *Explorations across languages and corpora* (pp. 563–580). Frankfurt, Germany: Peter Lang.

Chanier, T., Poudat, C., Sagot, B., Antoniadis, B., Wigham, C. R., Hriba L., … Seddah, D. (2014). The CoMeRe corpus for French: Structuring and annotating heterogeneous CMC genres. *Journal of Language Technology and Computational Linguistics*, 29(2), 1–31.

Chanier, T. (2013). *EUROCALL 2013, Survey on CALL in the digital humanities: Considering CALL journals, research data*. Paper presented at *EUROCALL 2013*, Évora, Portugal.

Chanier, T., Reffay, C., Betbeder, M.-L., Ciekanski, M., & Lamy, M.-N. (2009). LETEC (Learning and Teaching Corpus) Copéas [corpus]. Mulce.org: Clermont Université.

Ciekanski, M., & Chanier, T. (2008). Developing online multimodal verbal communication to enhance the writing process in an audio-graphic conferencing environment. *ReCALL*, 20(2), 162–182. doi:10.1017/S0958344008000426

CINES (2014). Description of resource formats eligible for archiving. National Computing Center for Higher Education. <https://www.cines.fr/en/long-term-preservation/expertises/formats-expertise/facile/>

CoMeRe (2015). Repository of computer-mediated communication (CMC) corpora. [webservice] Ortolang: Nancy. <http://hdl.handle.net/11403/comere>

Creative Commons (2015). Licences for publishing knowledge and data in open access formats [website]. <http://creativecommons.org/>

Crystal, D. (2004). *The language of the Internet*. Cambridge, United Kingdom: Cambridge University Press.

DARIAH (2013). Digital research infrastructure for arts and humanities [website]. <http://www.dariah.eu/>

Dortmunder Chat-Korpus (2003–2009). German corpus of informal text chat [website]. <http://www.chatkorpus.tu-dortmund.de/>

DWDS (2013). Das Digitale Wörterbuch der deutschen Sprache [website]. <http://www.dwds.de>

Dyke, G., Lund, K., Jeong, H., Medina, R., Suthers, D. D., van Aalst, J., … Looi, C.-K. (2011). Technological affordances for productive multivocality in analysis. In H. Spada, G. Stahl, N. Miyake, N. Law, & K. M. Cheng (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning (CSCL 2011)* (Vol. I, pp. 454–461). Hong Kong, China: International Society of the Learning Sciences.

Flewitt, R., Hampel, R., Hauck, M., & Lancaster, L. (2009). What are multimodal data and transcription? In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 40–53). London, United Kingdom: Routledge.

Goodwin, C., & Duranti, A. (1992). Rethinking Context: An introduction. In C. Goodwin & A. Duranti (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 191–227). Cambridge: Cambridge University Press.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 123–145). Amsterdam, Netherlands: Rodopi.

Guichon, N., & Hauck, M. (2011). Teacher education research in CALL and CMC: More in demand than ever. *ReCALL*, 23(3), 187–199. doi:10.1017/S0958344011000139

Halliday, M. A. K. (1989). "Part A" in M. A. K. Halliday & R. Hasan (Eds.), *Language, context and cext: Aspects of language in a social-semiotic perspective* (pp. 1–49). Oxford: Oxford University Press.

Hatch, T., & Grossman, P. (2009). Learning to look beyond the boundaries of representation. *Journal of Teacher Education*, 60(1), 70–85. doi:10.1177/0022487108328533

Huma-Num (2014). French national corpora consortiums in various areas of arts and humanities. <http://www.huma-num.fr/service/consortium>

IMS Learning (2004). Description of learning scenarios (IMS-LD), and data packaging (IMS-CP). IMS Global Learning Consortium. Retrieved from <http://www.imsglobal.org/specifications.html>

Francis, W. N., & Kučera, H. (1964). *A standard corpus of present-day edited American English, for use with digital computers*. Providence, RI: Brown University.

Laks, B. (2010). La linguistique des usages: de l'exemplum au datum. In P. Cappeau, H. Choquet, & F. Valetoupolos (Eds.), *L'exemple et le corpus quel statut? Travaux linguistique du Cerlico, 23* (pp. 13–28). Rennes, France: Presses Universitaire de Rennes.

Lamy, M.-N. (2012). Click if you want to speak: Reframing CA for research into multimodal conversations in online learning. *International Journal of Virtual and Personal Learning Environments*, 3(1), 1–18. doi:10.4018/jvple.2012010101

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511815355

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. New York, NY: Routledge.

MacWhinney, B. (2009). Manual of the CHAT transcription format used in the CHILDES project. Retrieved from <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1181&context=psychology>

Mulce documentation (2015). Documentation on Mulce repository and Mulce methodology [website]. <http://mulce.org>

Mulce repository (2013). Repository of learning and teaching (LETEC) corpora [webservice]. Clermont Université: MULCE.org. <http://repository.mulce.org>

Open Knowledge (2013). Definition of "open" with respect to knowledge and data. <http://opendefinition.org/okd/>

O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511497650

Reffay, C., Betbeder, M.-L., & Chanier, T. (2012). Multimodal learning and teaching corpora exchange: Lessons learned in 5 years by the Mulce project. *International Journal of Technology Enhanced Learning*, 4(1), 1–20. doi:10.1504/IJTEL.2012.048310

Schiffrin, D. (1994). *Approaches to discourse*. Malden, MA: Blackwell.

Shih, Y.-C. (2014). Communication strategies in a multimodal virtual communication context. *System*, 42, 34–47. doi:10.1016/j.system.2013.10.016

Sindoni, M. G. (2013). *Spoken and written discourse in online interactions: A multimodal approach*. New York, NY: Routledge.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation*, 816–820.

TEI (2015). Text Encoding Initiative. [website]. <http://www.tei-c.org/>

TEI-CMC (2015). Computer-Mediated Communication working group of the TEI consortium [website]. <http://wiki.tei-c.org/index.php/SIG:Computer-MediatedCommunication>

The Open University. (2001). *Research methods in education*. Milton Keynes, United Kingdom: The Open University.

Wigham, C. R., & Chanier, T. (2013). Pedagogical corpus: Reflective teaching journals. [corpus] Mulce.org: Clermont Université. <http://repository.mulce.org>

Wigham, C. R. & Chanier, T. (2014). Pedagogical corpora as a means to reuse research data and analyses in teacher-training. In J. Colpaert, A. Aerts, & M. Oberhofer (Eds), *Proceedings of the sixteenth international CALL conference, 7–9 July 2014* (pp. 360–365). Antwerp, Belgium: University of Antwerp.

Wigham, C. R., & Chanier, T. (2015). Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world Second Life. *Computer Assisted Language Learning*, 28(3), 260–283. doi:10.1080/09588221.2013.851702

Yun, H., & Chanier, T. (2014). Corpus d'apprentissage FAVI (Français académique virtuel international). Banque de corpus CoMeRe. [corpus] Ortolang.fr: Nancy.

Zourou, K. (2013). Research challenges in informal social networked language learning communities. *eLearning Papers*, 34, 1–11.

AFTERWORD

Engineering conditions of possibility in technology-enhanced language learning

Steven L. Thorne
Portland State University, USA & University of Groningen, The Netherlands

This is a tremendously exciting time to be a language educator, applied linguist, or language technology specialist in part because technology has come to mediate all manner of professional, recreational, interpersonal, and educational activity. As Bryan Smith and I noted in a recent publication (Thorne & Smith, 2011), second and foreign language researchers and educators have long recognized the potential of digital technologies to provide access to input, practice, and rehearsal (audio recordings, video, tutorials, drills, mini games), to amplify possibilities for meaningful and creative expression (text and media processing), to extend existing and create new opportunities for interpersonal communication (synchronous and asynchronous messaging, online intercultural exchange), to collaborate in (often) linguistically rich multiparty interaction in the 'wild' (i.e., naturally occurring and non-institutionally located online environments and communities), and to construct relevant presentations of self in digital media environments. Indeed, independent of the issue of successful integration of technologies into formal educational spaces, late modernity is increasingly defined by the seeming ubiquity of mediated engagement as a routine and unmarked dimension of life activity.

The interest in computer-assisted language learning (CALL) has steadily grown for more than three decades (Hubbard, 2009). Throughout this period, authored and edited books and academic journals such as Language Learning & Technology, ReCALL, and the CALICO Journal (among others) have provided robust scientific studies and pedagogical and curricular innovations that indicate the effectiveness and efficacy of online interaction and digital environments for language learning. This said, and acknowledging the continuing relevance of the digital divide in many parts of the world, it is a debatable issue as to whether new technologies have transformed the reality of language learning in most instructional contexts. Many institutional settings continue to employ pedagogical orientations and activity types that educators from 50 or 100 years ago would find largely familiar. In some cases, when new technologies are included, they serve as

DOI 10.1075/lsse.2.11tho

a digital simulacrum of earlier analog practices, and while some uses present opportunities for new forms of engagement and communicative interaction, others were developed in good part because of their recognizability or due to an unreflexive faith in the efficacy of, to take one example, patterned repetition of the sort that has informed language learning drills and worksheets for decades.

To be fair, what might be termed traditional methods and curricula within language education have achieved significant results for diligent and committed students. This acknowledged, what I found to be a compelling feature of this volume was its consistent encouragement for empirical investigation and iterative CALL design that has the potential to ameliorate outcomes for a larger number of learners. In reading through this manuscript, I particularly appreciated that the introductory chapter in the volume (Caws & Hamel) emphasized human-computer interaction processes, or if you will, technology-mediated relational dynamics in motion, which Caws and Hamel further specify as 'learner-computer interaction' (or LCI) in order to more precisely focus on the human developmental processes that inform CALL interventions. They leverage the notion of engineering, both in practical and adaptive application as well as metaphorically, as a framework that has the potential to make more rigorous the process of designing and implementing effective and robust conditions for language learning.

The first section of this book focused on 'frameworks guiding research', which in this case implicitly referenced a praxis approach that emphasized the dialectical union of research with the design of technology-mediated learning environments. Design-based research, which unites empirical analysis with learning theory-driven design (e.g., Caws & Hamel, this volume; Levy & Caws, this volume; Rodriguez & Pardo-Ballester, 2013), usefully informs an expansive view of language learning that helps to contextualize discrete system components and learner actions within a more holistic developmental framework. In their introduction, Caws and Hamel succinctly stated a fundamental question that informs all of the chapters in this volume: "design [is] critical for the success (or failure) of any intervention. And if good design can lead to better learning, we ought to ask ourselves this simple question: how can we design good, sustainable learning ecosystems that are mediated by technology?" Their response was to urge CALL practitioners to explicitly take on the role of an engineer and in so doing, to scientifically explore developmentally fecund opportunities presented by the profound human capacity to adapt and modify their cognitive, communicative and material environments through the creation of new, and use or adaptation of existing, mediating artefacts (in this case, digital technologies in the service of language use and learning).

Catalyzed by the engineering metaphor, the first section of the volume focused on theoretical frameworks that interface CALL design and pedagogical

interventions with contemporary approaches to ontology, epistemology, and methodology.

Educational ergonomics, defined as analysis of the interaction between learning and educational interventional design, is a perspective that is unfortunately rarely visible in the field of CALL. Caws and Hamel (Chapter 2) effectively outlined the merits of this approach as it helps to reveal the dynamics of learners engaging CALL tools in relation to efficiency, effectiveness, and the iterative redesign of technology-enhanced learning environments. The recuperation of educational ergonomics within CALL underscores the dialectical relationship of humans and technology and emphasizes the multiple culturally informed identities of technologies as a function of their situatedness in often heterogeneous webs of social practice (Thorne, 2003, 2009, 2016). Aligning with following chapters, the discussion of educational ergonomics was usefully integrated with the notion of affordances, mediated goal-directed action, cultural historical theories of development, and the data driven observational evaluation of the efficiency and effectiveness of CALL interventions.

Blin (Chapter 3) described the theory of affordances, generally associated with the ecological psychology movement (e.g., Gibson, 1979; Norman, 2002; see also van Lier, 2004), which posits that diverse environments differentially enable, or create agentive opportunities, for human action. Blin worked to clarify the messy history of the concept of affordances and aligned this post-cognitivist and ecological view of human action with complexity, activity theoretical, and distributed views of language learning that could, and in my view should, more substantially inform CALL research and development.

Schulze and Scholz (Chapter 4) outlined a compelling research paradigm that argued for understanding CALL as a complex adaptive system, one that is nonlinear, interconnected, characterized by variability, and which integrates learner-computer interactions with language learning processes and theories. I was particularly pleased to see explicit attention to innovative theories of language structure and development, such as construction grammar, usage-based linguistics, and emergentism, which they contend are commensurable with complexity theory as an ontological base, and with sociocultural theory as a framework for understanding and analysing processes of human development.

Levy and Caws (Chapter 5) ambitiously addressed the challenge of integrating macro contextual factors (curricula, preexisting levels of technical competence and varying levels of technical support, the availability of technology, systems thinking, school policy, and the like) with learners' discrete educational experiences in technology-mediated interaction. In particular, the authors advocated for a movement toward normalization (e.g., Bax, 2011) that would situate CALL as a routine, supported, expected, and tightly integrated aspect of the broader

system of constraints and affordances that comprise instructed and institutionally located language learning. This preferred future view of CALL as seamlessly interwoven with non-CALL activities is a obvious telos for instructed language learning environments, and while not explicitly mentioned in the chapter, achieving normalization in this sense has the potential to tremendously increase the ecological validity of language educational practice (Thorne, 2013) in view of the ubiquity of mediated and non-mediated cognitive and communicative activity that increasingly comprise everyday contexts.

The second section of the volume moved toward methodologies, approaches, and case studies that all built upon the intensive capture and analysis of learner behaviours and/or communicative activity. Specific approaches included learner modelling, screen capture, eye-tracking, video-based analysis of gesture in video conferencing settings, and corpus analytics. Heift (Chapter 6) described learners' varying rates of the use of instructional scaffolding (here, help features, answer look-up behaviour, and preemptive feedback in an intelligent CALL environment) that reveal insights resulting in better individualizing instruction to accommodate diverse learner types (or learner personas) while also facilitating meeting the needs of individual learners. The three subsequent chapters focused on process-oriented learner data. Hamel and Séror (Chapter 7) discussed video screen capture (VCS) as a way to document learner behaviours that assisted with the continuing development of an online dictionary prototype. They emphasized insights that emerged from the keystroke-by-keystroke process of learners engaged in composition and VCS as a way to objectify and reflect upon the L2 writing process. While the argument is cogently made that VCS is useful as a form of usability testing (and hence helpful for researchers, teachers, designers, and technologists), I was particularly struck by its potential pedagogical value to learners themselves as a way to objectify and more powerfully self-regulate their composition process, potentially taking the form of students producing a think-aloud account as they watch their own composition process unfold.

As has been discussed elsewhere (e.g., O'Rourke, 2008; Smith, 2010), the interaction record of online communication is important (a log file of text chat, for example), but also somewhat thin in that the temporality of allocation of attentional resources is largely unknown. Eye-tracking techniques (Stickler, Smith, & Shi, Chapter 8) contribute to better understanding gaze within the visual field as it relates to reading, text production, and by proxy, aspects of language processing, all of which open up new possibilities for investigating real-time L2 use and learning. In a related vein, Cohen and Guichon (Chapter 9) embraced the issue of developing more holistic units of analysis for investigating video-based online intercultural exchange. They addressed methodological approaches, specifically multimodal analysis, which incorporate spoken interaction with gesture, gaze,

and bodily orientation in order to more fully situate text-based corpus data for purposes of SLA research. Chanier and Wigham (Chapter 10) continued the focus on uses of corpora in CALL and proposed a staged methodology for structuring and sharing (in an open data repository) LCI for purposes of teacher professional development, SLA research, and more broadly, corpus linguistic investigations of computer-mediated communication.

As will be apparent to readers, this is an ambitious volume that presents fresh and innovative perspectives. It repositions CALL as a design-based process involving the engineering of technology enhanced learning environments and their subsequent iterative improvement via the empirical investigation of learner-computer interaction data. Numerous methodologies and digital tools support LCI research in the service of ameliorating the efficacy of technology-enhanced learning. As described in this volume, these include eye-tracking, intelligent CALL environments, video screen capture, and procedures for creating multimodal annotations of corpus data, all of which help to emplace conventional CALL data sources, such as the textual interaction record, in more fine-grained context. Theoretically, this volume is tightly aligned with contemporary approaches to language structure and human development, with significant and well-integrated treatments of dynamic systems theory, usage-based linguistics, ecological psychology, cultural historical theories of artefact mediation, educational ergonomics, and the interplay of micro and macro dimensions of learner-computer interaction.

For practitioners and researchers working in the areas of applied linguistics, CALL, and L2 education, this volume has provided numerous sign posts guiding us forward on the path of creating more developmentally effective technology-mediated learning environments. The hard work, of course, begins now.

References

Bax, S. (2011). Normalisation revisited: The effective use of technology in language education. *International Journal of Computer-Assisted Language Learning and Teaching*, 1(2), 1–15. doi:10.4018/ijcallt.2011040101

Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.

Hubbard, P. (2009). General Introduction. In P. Hubbard (Ed.), *Computer Assisted Language Learning, Volume 1: Foundations of CALL*. Critical Concepts in Linguistics (pp. 1–20). New York: Routledge.

Norman, D. A. (2002). *The design of everyday things*. New York, NY: Basic Books.

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer-mediated communication and language learning. *Computer Assisted Language Learning*, 21(3), 227–251. doi:10.1080/09588220802090253

Rodríguez, J. C., & Pardo-Ballester, C. (Eds.). (2013). *Design-based research in CALL*. CALICO Monograph Series, Volume 8. San Marcos, TX: CALICO.

Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (Vol. 11, pp. 79–97). Dordrecht, Netherlands: Springer. doi:10.1007/978-90-481-9136-9_6

Thorne, S. L. (2003). Artifacts and cultures-of-use in intercultural communication. *Language Learning & Technology*, 7(2), 38–67.

Thorne, S. L. (2009). 'Community', semiotic flows, and mediated contribution to activity. *Language Teaching*, 42(1), 81–94.

Thorne, S. L. (2013). Language learning, ecological validity, and innovation under conditions of superdiversity. *Bellaterra Journal of Teaching & Learning Language & Literature*, 6(2), 1–27.

Thorne, S. L. (in press, 2016). Cultures-of-use and morphologies of communicative action. *Language Learning & Technology*, 20(2).

Thorne, S. L., & Smith, B. (2011). Second language development theories and technology-mediated language learning. *CALICO Journal,* 28(2), 268–277.

van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective*. Boston: Kluwer.

Author index

Subject index

This book focuses on learner-computer interactions (LCI) in second language learning environments drawing largely on sociocultural theories of language development. It brings together a rich and varied range of theoretical discussions and applications in order to illustrate the way in which LCI can enrich our comprehension of technology-mediated communication, hence enhancing learners' digital literacy skills. The book is based on the premise that, in order to fully understand the nature of language and literacy development in digital spaces, researchers and practitioners in linguistics, sciences and engineering need to borrow from each others' theoretical and practical toolkits. In light of this premise, themes include such aspects as educational ergonomics, affordances, complex systems learning, learner personas and corpora, while also describing such data collecting tools as video screen capture devices, eye-tracking or intelligent learning tutoring systems. The book should be of interest to applied linguists working in CALL, language educators and professionals working in education, as well as computer scientists and engineers wanting to expand their work into the analysis of human/learner interactions with technology communication devices with a view to improving or (re)developing learning and communication instruments.

ISBN 978 90 272 5751 2

John Benjamins Publishing Company