Computer Assisted Language Instruction Consortium

CALICO JOURNAL (ONLINE) ISSN 2056–9017

Article

The role of qualitative approaches to research in CALL contexts: Closing in on the learner's experience

Mike Levy

Abstract

The article considers the role of qualitative research methods in CALL through describing a series of examples. These examples are used to highlight the importance and value of qualitative data in relation to a specific research objective in CALL. The use of qualitative methods in conjunction with other approaches as in mixed method research designs are a particular focus. It is argued that qualitative methods are especially effective when used in this way.

The discussion also aims to elaborate upon the role of qualitative approaches within CALL specifically, as a domain for research study with particular attributes that require a qualitative orientation. Here the use of electronic dictionaries is considered. Dictionary use is suited to such a discussion because it occurs frequently in everyday life, suggesting a qualitative approach, as well as in research studies that are strictly controlled, as in a quantitative approach. The contrast is instructive and helps demonstrate the respective strengths and limitations of each method.

Also central to the discussion is the language learner. A number of the studies described emphasize the importance of listening to the students' voice in the qualitative component (e.g., Conole, 2008; Jones, 2003). It is in the unpacking of what students actually do moment-by-moment in CALL tasks and activities that best illustrate the strengths of qualitative methods in enhancing our understanding of mediated learning and thereby driving productive research agendas.

KEYWORDS: CALL; MIXED METHODS; QUALITATIVE RESEARCH; RESEARCH DESIGN

Affiliation

The University of Queensland, Brisbane, Australia.
email: m.levy@uq.edu.au

CALICO JOURNAL VOL 32.3 2015 554–568

doi : 10.1558/cj.v32i3.26620

Introduction

New technologies in language education are being employed in many different settings around the world, each with its own particular context. Currently, for example, I have three doctoral students who are each positioning their research in entirely different educational settings, in Vietnam, Saudi Arabia and Japan. There are numerous constraints as well as many opportunities for using technologies to support and enhance language learning in these very different contexts of use. In each case, the doctoral student is carefully examining how contextual parameters help shape the practice of CALL in that particular setting. In Vietnam, recently implemented technology standards for language teachers are exerting a profound impact, especially under the direction of the National Foreign Language 2020 Project. A new organization, *VietCALL,* was established in 2014 with the plan of playing a major role in implementing the standards. In Saudi Arabia, at a private university with a superb technological infrastructure, student motivation is a priority as learners work within a predetermined curriculum within which technology use has to fit. In Japan, the flipped classroom is being implemented in the primary school classroom and the viability of such an approach in such a setting constitutes a focus. Work within the school setting and within the home in the cultural setting of Japan is of special interest. Each setting has its own history that has led to current understandings and initiatives. This in turn leads to particular goals, priorities and agendas both in research and practice. Curricula goals, pedagogical methods, examination requirements, and student needs all exert an influence on how technology might be employed and used.

For the qualitative researcher, for any given setting, we wish to understand more fully the learner's experience. As Heigham and Croker (2009) observe, 'As qualitative researchers … [the] research focus is on the participants – how participants experience and interact with a phenomenon at a given point in time and in a particular context …' (p. 7). This focus on understanding the experience of the learner is essential for the researcher who endeavors to tap into the learner's perspective through a range of qualitative methods and data-collection techniques. Further, as Rossman and Rallis (2003: 9) affirm, 'Qualitative researchers go to the people; they do not extricate people from their everyday worlds.' This is a key point for the CALL domain where increasingly learners are using their own powerful, personal technologies to work independently, outside the classroom and without the teacher. In CALL, a focus on better understanding the learner experience in settings where new technologies are being employed is critical.

An article written by Conole in 2008, points towards some important features and qualities of contemporary language learning environments with respect to new technologies. While the processes and products of language

teaching and learning twist and evolve, the learner voice can easily be overlooked. Technology is meant to be emancipatory and increasingly, through their personal technologies, learners can create their own occasions for language learning. To better understand these occasions and events, qualitative research methods are needed.

The Conole study involved a number of in-depth case studies and two of the case studies were described in detail. The case study is of course one of the primary qualitative methods. In the Conole study there were two main questions. How do learners engage with and experience e-learning (perceptions, use and strategies)? How does e-learning relate to and contribute to the whole learning experience? (Conole 2008: 125). The data collection comprised three sources, an online survey, audio logs and interviews, so the data was predominantly qualitative. The case studies used data from 85 audio log diaries and four semi-structured interviews. Two case studies were presented in the study report and the qualitative data collected provides for a detailed and differentiated description of technology use. The findings are complete with a rich discussion of the contexts of use with the tasks and usage matched to the technological tools employed. The net result is a detailed understanding of what tools are used and for what purpose.

The two Conole case studies provide a good example of how qualitative approaches can generate important detail on how learners choose to employ particular technologies in everyday, natural settings. Such work can also be helpful in addressing a potential problem in CALL research generally, as noted in the conclusion to a study by Steel and Levy (2013), '[T]here appears to be a gap or disconnect between what students are actually doing and where research directions in CALL are taking us' (p. 319). Careful interrogation of the student perspective can help ensure that research and practice remain aligned and connected.

The role and contribution of qualitative research

The design of three research studies is considered in detail in this section. The first study is large scale and aims to convey how qualitative data can extend and enrich results drawn from purely quantitative findings. The second study aims to evaluate a specific feature in a multimedia CALL design, in this case the effectiveness of visual and verbal annotations. The study findings show how the qualitative component illuminates the quantitative findings in important, revelatory ways that allow for much greater detail and elaboration around a specific focal topic. The third study is employed to facilitate a discussion concerning the important issue of control in a research design. As a field, CALL research includes studies that range in their design from those that are very carefully controlled, in artificial, experimental, laboratory settings, to those

that aim to gather data in natural contexts approximating everyday life. This issue is important in an area such as CALL where learners are increasingly utilizing their own powerful, personal technologies in their language learning independently outside of class without a teacher present.

Levy and Steel (2015) – Use of electronic dictionaries

Undergraduate foreign language students enrolled at the University of Queensland were surveyed about their technology use as language learners in mid-2011 (inside class, outside class, or both). All students were, at the time of the survey, studying a formal language course provided by the university in a blended mode. Students ranked the top three technologies they perceived as most beneficial to their language learning. Overall, 587 (from 2,114 – 28% response rate) language students completed the online survey.

The quantitative findings demonstrate that students are not confined to using the centrally provisioned technologies made available through their institutional studies. In fact in this study they played a relatively minor role. Instead they are drawing upon their own personal technologies that are fit for disciplinary purpose. Additionally, the results clearly highlight that the technologies and tools that language learners now potentially have at their fingertips are powerful, expansive and changing. These results indicate the importance of listening to the learners' voice, as indicated by Conole earlier. For this to occur, qualitative approaches to research are key.

In this study, and the one that preceded it (Steel & Levy, 2013), it is clear that the electronic dictionary tool is a high use item for language learners. The qualitative results show that many students are sophisticated dictionary users and quite capable of assessing the strengths and weaknesses of different learning tools. The quantitative component of the paper was strengthened through the addition of qualitative detail, and this data was generated via the open questions in the survey.

In a perceptive discussion of open-response items in questionnaires, the qualitative component, James Dean Brown (2009: 201) says:

> such items are best suited for exploratory research, where, at the beginning, the researcher may not know what the central issues are on a particular topic … or even what specific questions need to be asked. Open response questionnaires provide a way to find out, in an unstructured manner, what people are thinking about a particular topic or issue. As such open response questionnaires often serve as the basis for further, more structured research.

Brown (2009, p. 205) continues, 'Answers to open response items can also provide striking examples and illustrative quotes, so they offer a far greater richness, adding more depth and colour to the data than answers to closed

response items.' This was very true in the study we have just discussed. In CALL more broadly, it is a perspective that is increasingly important because learners have their own technologies that are capable of supporting their independence. Thus, Cohen and White (2008) position language learners as 'informed consumers' who need to make informed judgements about their needs, goals and purposes across their learning experiences. They 'have a view of language learners who actively construct and fashion a way of learning for themselves based on the alternatives available' (p. 200). Only through qualitative research methods can such personal choices become visible.

Jones (2003) – Seeking the optimal design

Jones (2003) describes an experimental study that employs a control group and three treatment groups, and inferential statistics to determine the kind of multimedia annotations that best support listening comprehension and vocabulary acquisition. Jones does not rely solely on quantitative, experimental techniques alone, however. Her research design also involves a qualitative dimension that uses participant interviews to provide the students' 'voice', and this is used in combination with the experimental data to provide a richer and more complete interpretation of the phenomena in focus.

This study examines a hypothesis and a research question: the hypothesis says that students will recall more propositions from the text of a listening comprehension text activity when they have a choice of visual and verbal (multimodal) annotations compared with students who complete the same listening tasks with single type (unimodal) annotations (visual or verbal), or no annotations at all; the research question examines how students best acquire vocabulary under the different conditions and details of the variation from treatment to treatment. Accordingly, 171 English-speaking college students in second semester beginner French participated in this study.

The quantitative results of the study showed that students remembered word translations and recalled the passage best when they had access to visual and verbal annotations, reasonably well when they had access to one or the other, and poorest when no annotations were available. However, the qualitative results revealed much more. Beneath this general finding there was much valuable detail documenting precisely the conditions under which the different kinds of annotations worked effectively, and it was qualitative research methods that led to these insights and qualifications. For example, even though the combination of visual and verbal annotations produced the best results, students also expressed, through the interviews, a desire for choice, and options for viewing material in both a visual mode and a verbal mode separately, according to the particular situation. Another interesting result was that many students believed that the verbal annotations (i.e., translations) were easier

to process than images, requiring less mental effort to clarify meaning. However, those who accessed visual annotations outperformed those who did not, showing the value on this occasion of learners choosing the more difficult path for increased learning (p. 58).

The qualitative dimension is extremely important in the Jones study. The results indicate well just how such additional data can assist the researcher in contextualising quantitative findings that may at first appear straight-forward and conclusive. In this instance, the student's comments arising out of the interviews shed further light on the use of annotations in CALL design. For example, as well as the students obtaining the lowest scores on all the dependent measures in the control group, the students also had 'the lowest opinion' (p. 59) of the listening activity without annotations: in fact, Jones reports that one student referred to this activity as 'cruel and unusual' (p. 48). This is a significant finding, though we are not speaking of 'significance' in quantitative terms. One cannot ignore a student's attitude towards the tasks they are asked to complete because their attitude may affect their performance. If possible, knowing a student's views towards the task should be considered part of the research. In the Jones study, the transcripts from the interviews showed that a number of students in the control group were very frustrated by the task.

Student comments from the interviews highlighted some contradictions with the experimental findings. For example, the students' performance was low when only verbal annotations were made available; certainly some students in the verbal group believed that the translations of keywords helped their understanding of the aural passage. So there was evidence of differences within the group that would have been overlooked if only the quantitative statistical data had been collected. In this study, as in others, qualitative data from student interviews provided insights into individual differences. Whatever the broad conclusions of the research study on statistical grounds, adding a qualitative dimension enables the researcher or the designer to be made aware of the variation that often arises as a result of individual characteristics and behaviours.

Tono (2011) – Dictionary look-up behaviour

The Tono (2011) study on EFL learners' dictionary look-up behaviour is deserving of attention because of key features in the study design and their implications for research in CALL. Tono constructs a small-scale, controlled experiment with eight participants, and utilizes an eye-tracker in a carefully controlled design. Tracking software, video capture and eye-tracking techniques are being used increasingly in CALL research studies. The research design involves three cameras, two associated with the eye-tracker, and each subject was required to look at a PC monitor with their head as still as possible

on a chin-rest while searching for the information on the display. The task the students had to complete was also carefully planned using a specially devised dictionary entry with keyword signifiers in caps alongside each separate definition and the word 'make' that the students had to look up in the dictionary was supplied beforehand.

It is clear from this set-up that the experiment is carefully controlled. Clearly it has to be if the researcher wishes to use a precise measurement instrument like the eye-tracker and obtain satisfactory results. The technique can provide richer information on the look-up process in a dictionary once a user has located the relevant headword, as in the sequence in which segments of the text are read, and the process of elimination that is required to identify the correct sense of a word that has many meanings is completed. In this case, the students took very different pathways through the text, one successfully arriving at the correct answer, the other not. Overall, the study provides many insights into how individual readers process the text. But there are serious difficulties with this kind of study also.

How far this set-up might approximate to real world dictionary use or look-up behaviour remains an open question. Tono acknowledges, 'the use of the eye-tracker makes look-up performances different from those in normal settings' (p. 126). The lab set-up is very precise, and essentially quantitative in terms of the data it generates. Its sequential nature and the differing size of the circles that are used to indicate concentration, moments of rest or inaction in gaze data give away the fundamentally quantitative nature of eye-tracking as a research approach. But in 'real life', language learners do not use specially-written dictionary entries, or have their chins supported on head-rests to keep them still while they read them. While the Tono study does generate valuable results, it is going to be the qualitative approaches to research that are most likely to lead to insights into the particulars of spontaneous dictionary use by language learners in everyday life. Thus, a mix of methods may be in order, or perhaps a sequence of research studies (see Stickler & Shi, 2015). Another dictionary study by Hamel (2012) helps us explore this issue further.

Hamel (2012) conducted a study to test aspects of the usability of an online learner dictionary prototype. Like Tono her study was small scale, with six participants, and conducted under lab setting conditions using computer tracking (not eye-tracking). However, what marks it as different to the Tono study in terms of design is that in the Hamel study there was no time constraint for completing the dictionary tasks: Hamel reports, on average, participants took less than two hours to complete the experiment (p. 346). Participants had the choice to use the dictionary whenever they wanted. Thus, while this study was still conducted under controlled conditions, it differed in important ways from Tono (2011). In Hamel (2012) subjects simply turned to the dictionary when

they needed it, more like the circumstances of real life dictionary use. In the future, Hamel (2012: 359) is planning a 'classroom intervention' rather than a lab experiment. Such a path forward indicates an awareness of the limits of a controlled experiment and a recognition of the contingencies of everyday life where dictionary use occurs spontaneously as a by-product of a primary task.

A controlled experiment in a lab setting inevitably requires modification of the natural world of things. While carefully controlled studies as Tono on dictionary look-up behaviour using eye-tracking do get us closer – in one sense – to the learner experience, we do not engage with such questions as why and under what conditions learners look up a word in the dictionary in the first place (in-class/out-of-class, under direction or without it), what they might learn through the process and ultimately whether they are successful in finding what they want. Interestingly, in his discussion Tono (p. 149) found that one third of the dictionary searches resulted in failure. Such studies do not answer a whole suite of questions relating to learner motivation and intention. It is here that qualitative research methods must come into play for a deeper and broader understanding.

Closing in on the learner experience

In the studies exemplified by Conole (2008) and Jones (2003), we are using qualitative methods to dig deeper into the learners' experience by giving 'voice' to it. However, there is another sense in which the learners' experience needs to be accurately captured and represented when technology is involved, as when it is used to mediate interactions between human participants, as in synchronous computer-mediated communication (SCMC) for example. What is essential here is to capture the salient features of the particular, mediated learning environment and, by implication, to avoid falling into the trap of assuming that research findings deriving from co-located, face-to-face interactions can be transferred straightforwardly and without complication to mediated learning contexts such as SCMC. As noted by Levy (2000: 190): 'For the CALL researcher, technology always makes a difference; the technology is never transparent or inconsequential' (see also Levy, 2006).

A good example of this principle in action is described by Smith (2008) in a study that examined computer-mediated communication (CMC) between pairs of students learning German. Smith's study is relevant here because of its methods of data collection and analysis and the consequences for the findings of the research. Instead of only using a chat log file as a data source – formerly the typical approach – this study captured a video file of the whole interaction on screen. This is critical for two very important reasons. First using a video recording of the screen allowed the *history* of the construction to be captured, including any changes that have been made before the learner's text

becomes public. Chat logs omit this critical information. Second, recording such information illustrates a critical difference between online text chat and face-to-face (FtF) communication. In online chat participants construct their utterances in two stages: they can change their mind and make alterations privately before the text appears to their partners. In face-to-face conversations, this is not possible: all re-starts or slips of the tongue are heard by both participants, and hesitations and pauses matter (see Levy & Gardner, 2012). For the qualitative researcher such alertness to the context of the interaction and the effort to capture more of it is compatible with the goal of always wishing to get closer to the reality of the learner experience.

This idea of getting closer to the learner experience is key. O'Rourke (2008, 2012) provides further insights when he discusses all that may easily be overlooked when researching the detail of synchronous CMC. Such features may include body posture, gaze data, audible self-speech and so on. Not only does the discussion point to the limitations of output logs as an information source concerning an online interaction, but it also indicates that the use of such techniques as eye-tracking only take us a little closer to the learner experience; they by no means reveal the whole of it. In fact, in research to date, O'Rourke (2008: 233) argues that many aspects of the learning environment relevant to learning in SCMC have been neglected, namely:

(a) users' paralinguistic and non-linguistic behaviours – gestures, spoken utterances, posture, etc.;
(b) 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);
(c) drafting processes – i.e., editing of input prior to sending; and
(d) attentional focus – i.e., what users are actually attending to at a given moment.

The papers written by O'Rourke on method are worthy of careful reflection. This work shows just how different synchronous CMC may be compared with FtF conversation conducted in the same physical space. He explains why output logs are 'impoverished' (p. 236), and excludes entirely the 'private space' in which students construct their utterances in text chat. He concludes, '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 behavior – we need to "zoom in" and examine the texture of interactions with SCMC systems as experienced by the individual' (p. 247). The same applies to video communication as in Skype, for example, where the relative position of the webcam and the actual position of the eyes of the participants differ.

Generally speaking, in endeavoring 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 (1986), where they dedicate a whole chapter to this topic. As they say, ' "Doing" is an interpretation within a background and a set of concerns' (p. 143). The qualitative approaches referred to in this section are capable of providing a much richer description of the student experience. Such data may not provide the basis for an 'explanation', and it may not be generalizable, but as Krathwohl (1993) explains 'creative description' is required to map the terrain first before it becomes possible to design meaningful experimental, quantitatively-oriented studies. The study by Smith (2008) and the discussion by O'Rourke (2008; 2012) show just how important this is.

Mixed methods

In the discussion so far, the reader will have undoubtedly noted that the majority of the examples described involve mixed methods. Qualitative research methods work very well as a complement in CALL, although obviously they have their place independently also. In this section, drawing where possible on the examples described earlier, some key ideas on mixed methods in CALL will be described. The reader is advised to seek such texts as Heigham and Croker (2009), chapter 7 (Ivankova & Creswell) for a discussion on mixed methods for qualitative research in Applied Linguistics; and to the chapter on mixed methods approaches by Ware and Rivas (2012) for a discussion of the topic in relation to researching classroom integration of online language learning projects.

Both Ivankova and Creswell (2009) and Ware and Rivas (2012) provide an overview of the key features of mixed methods research designs and they each provide useful, complementary perspectives. Both describe the principal ways in which study designs vary along two axes. The chronological axis refers to the order in which the quantitative and qualitative components are conducted (i.e., sequentially or simultaneously). The 'emphatic' axis refers to the weighting of the component, that is, which component is dominant. The standard nomenclature, that is, for example, 'QUAN/Qual', is used to denote the relevant weighting, in this case the quantitative component. Ware and Rivas also provide a description of five example studies of research of online language exchanges with the key elements summarized in a table together with the quan/qual sequencing and weighting in each case.

Ware and Rivas (2012) make some valuable observations of relevance to the topic of this paper and the role of qualitative research methods. They explain that thus far there are few mixed methods studies for online language learning, and that for those that have been conducted, qualitative research designs

are often preferred because the exchanges are heavily contextualized both culturally and socially. They continue, 'Additionally, because of the cultural and social layers of context, key features of quantitative research, such as the need for standardized measures, generalizability of findings, and large sample sizes, serve as logistical as well as ideological barriers' (p. 110). They also highlight the predominance of simultaneous designs for practical reasons such as the need to complete a study within the span of a single semester. Their conclusion on mixed methods in research comes with an astute, practical reminder:

> Each tradition of enquiry, whether qualitative or quantitative, is grounded at the ideological level in a set of values and worldviews, and at the methodological level in a set of tools, techniques, and terminology. Most researchers require years of training to follow only one of these methodological paths, and many graduate programs therefore must emphasize training in one methodology over the other to cultivate experts within a particular methodological tradition. (p. 127)

Their practical response to this dilemma is to advocate team-work and collaboration with colleagues, a strategy with which I wholeheartedly agree.

Strengths and limitations

Every research study, whatever its orientation or point of departure, circumscribes an area of concern and a point of focus. Necessarily, the study will foreground certain aspects and background others, and it will inevitably include and exclude different factors that potentially may impact upon the phenomenon under investigation. As a result, the study will always have limitations and (we hope) strengths. The three studies discussed in this paper are no exception.

In the qualitative data drawn from a large-scale survey, and through the responses to the open questions, Levy and Steel (2015) report on what 587 students *say* they do when using electronic dictionaries. This reportage does not necessarily reflect what students *actually* do, nor how the associated processes may or may not contribute to language learning (in a measurable way). This is a limitation of this study. Smaller-scale studies are needed to complement and enrich the findings of the present study. In both cases, the data collected may be qualitative.

Hamel (2012: 342) provides a useful analysis of empirical studies on dictionary use by language learners and lists some of the data collection instruments and alternatives. These include: think aloud protocols, stimulated recall interviews, introspective journals, computer logs and screen capture. Many of these qualitative techniques and approaches also have the potential to enable a more direct insight on learner cognitive processes also. For instance, Cotos (2011) provides appropriate and valuable introspective evidence (in addition to quantitative) to support the argument about the significance of the

qualitative data in the study of the learner perspective in the CALL design and evaluation of a web-based automated writing evaluation program.

When we weigh the relative value of quantitative and qualitative methods, I consider it unlikely that we would likely reside wholly on one side or the other. Too much information is lost, especially in a complex, emerging field like CALL, if the research relies solely on one particular approach. The conditions dictated by controlled laboratory settings while valid for quantitative studies lie far from the reality of modern personal technology use by many language learners, and one is ill-advised to suggest that findings from research conducted in one setting necessarily apply to the other. Still, studies conducted in laboratory-like conditions (e.g., Tono, 2011; Hamel, 2012) do provide valuable information on patterns of use that may be followed up in qualitative research designs conducted in natural settings. The reverse also holds true.

There are also material limitations. Techniques such as eye-tracking, while feasible on larger-screen desktop or laptop computers become much more challenging when users employ tablet or phone-sized screen technologies. Still, other data collection methods and techniques such as video-screen capture with talk-aloud protocols, stimulated recall and so forth can be used in tandem with eye-tracking, for example, to strengthen the validity of the data. Each method, both separately and in combination has its strengths and limitations. Research in this area will inevitably require different, complementary techniques and approaches. Ultimately, large and small-scale studies are necessary to provide breadth and depth, sometimes through mixed methods approaches, to reach a deeper understanding of the processes involved.

Conclusion

I would like to close this discussion with some reflections based upon a quote from Martyn Hammersley, a recognized leader in qualitative research methods. In his book, 'What is qualitative research?', Hammersley (2013), concludes:

> Finally, given that, in practical terms, the difference between qualitative and quantitative methods is a matter of degree, I do not believe that 'qualitative research' is a genuine or useful category – any more than is 'quantitative research'. While, at present, we cannot avoid reliance upon this distinction, we need to move toward a more adequate typology, exploring the various options open to social researchers as regards how they formulate research questions, engage in research design, collect and produce data, analyse it, and report their findings. (p. 99)

I agree. In this short paper, we have seen studies that are essentially quantitative getting closer to the learner experience, as in Tono's use of eye-tracking.

This study was conducted with just eight subjects, a small sample which is not typical of many quantitative studies; Hamel's study using video capture was conducted with just six participants. Quantitative studies are not necessarily large scale, and qualitative studies not necessarily small (e.g., qualitative component of Levy & Steel, 2015). Hamel's study was also conducted in controlled laboratory setting conditions without time constraints on the dictionary tasks (p. 346). The degree of control adopted by the researcher and implemented through the research design is often more of an issue than whether a study is labelled quantitative or qualitative. Of course there are constraints on task-completion times in most normal settings, in a regular language class for example. Further, as Tono points out, the use of the eye-tracker as part of a research design exerts an effect and changes the way users look up words in the lab setting as opposed to the everyday one. Great caution needs to be observed therefore before assuming results generated in controlled conditions bear relevance to parallel activities in natural settings – and yet we all know the temptation is there to make the jump, to make those assumptions. Qualitative designs are intended for research studies focusing on natural settings, but of course conducting research in these settings poses a great challenge. This applies particularly to a field like CALL. When we consider the increasing use of mobile technologies such as smart phones and tablets inside and outside the classroom, with and without a teacher present, we begin to recognize some of the limitations of studies that restrict themselves to strictly controlled lab-style conditions and a PC monitor.

Such concerns really just begin to indicate some of the complexities involved in closing in on the learners' experience. Such complexities are being encountered directly by the three PhD students I referred to in the introduction as they aim to understand how contextual factors, at multiple levels, impact upon any implementation of CALL in a particular context and setting. Hamel talks of planning a classroom intervention as a next step. One could then easily envisage further studies on dictionary use outside of class in a range of natural settings. In these examples, qualitative research designs and methods are required. Such a sequence of moves to pursue a research agenda also signals the importance, not only of the single study, but of a series of studies, some perhaps quantitative, some qualitative, some mixed that move collectively and in a coordinated way towards a research objective, for instance how our learners utilize their electronic dictionaries for language learning. We will need to employ all the technical means at our disposal in order to capture, analyse and understand the subtle and sophisticated patterns of interaction in each context. Such means may include video-capture, eye-tracking, multimodal tracking techniques, image interpretation, gesture monitoring and so forth. We can all look forward to an exciting time ahead.

About the author

Dr Mike Levy is an Honorary Professor at the University of Queensland, Brisbane, Australia. His research work 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 the design and use of online dictionaries and electronic translation tools. He is on the editorial boards of *CALL*, *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 (www.worldcall.org).

References

Brown, J. D. (2009). Open-response items in questionnaires. In J. Heigham & R. A. Croker (Eds), *Qualitative Research in Applied Linguistics,* 200–219. New York: Palgrave Macmillan.

Cohen, A. D. & White, C. (2008). Language learners as informed consumers of language instruction. In A. Stavans & I. Kupperberg (Eds), *Studies in Language and Language Education,* 185–205) Jerusalem: The Hebrew University Magnes Press.

Conole, G. (2008). Listening to the learner voice: The ever-changing landscape of technology use for language students. *ReCALL,* 20 (2), 124–140. http://dx.doi.org/10.1017/s0958344008000220

Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal,* 28 (2), 420–459. http://dx.doi.org/10.11139/cj.28.2.420-459

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. http://dx.doi.org/10.1080/09588221.2011.591805

Hammersley, M. (2013). *What is Qualitative Research?* London: Bloomsbury.

Heigham, J. & Croker, R. A. (Eds). (2009). *Qualitative Research in Applied Linguistics*. New York: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230239517

Ivankova, N. V. & Creswell, J. W. (2009). Mixed methods. In J. Heigham & R. A. Croker (Eds), *Qualitative Research in Applied Linguistics,* 135–164. New York: Palgrave Macmillan.

Jones, L. C. (2003). Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students' voice. *CALICO Journal,* 21 (1), 41–65.

Krathwohl, D. R. (1993). *Methods of Educational and Social Science Research: An Integrated Approach.* New York: Longman.

Levy, M. (2000). Scope, goals and methods in CALL research: Questions of coherence and autonomy. *ReCALL*, 12 (2), 170–195. http://dx.doi.org/10.1017/S0958344000000525

Levy, M. & Caws, C. (In press.). CALL design and research: Taking a micro and macro view.

In C. Caws & M-J. Hamel (Eds). *Learner Computer Interactions: New Insights on CALL Theories and Applications*. Philadelphia, PA: Benjamins.

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

Levy, M. & Steel, C. (2015). Language learner perspectives on the functionality and use of electronic language dictionaries. *ReCALL*, 27 (2): 1–20. Published online: 5 January 2015. http://dx.doi.org/10.1017/S095834401400038X

Levy, M. & Stockwell, G. (2006). *CALL Dimensions: Options and Issues in Computer Assisted Language Learning*. Hillsdale, NJ: Lawrence Erlbaum & Associates.

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. http://dx.doi.org/10.1080/09588220802090253

O'Rourke, B. (2012). Using eye-tracking to investigate gaze behavior 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,* 305–341. Bern: Peter Lang.

Rossman, G. B. & Rallis, S. F. (2003). *Learning in the Field: An Introduction to Qualitative Research* (2nd edn). Thousand Oaks, CA: Sage Publications.

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

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

Stickler, U. & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO Journal,* 32 (1), 52–81.

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

Ware, P. & Rivas, B. (2012). Researching classroom integration of online language learning projects: Mixed methods approaches. In M. Dooly & R. O'Dowd (Eds), *Researching Online Foreign Language Interaction and Exchange: Theories, Methods and Challenges,* 107–131. Bern: Peter Lang.

Winograd, T. & Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Reading, MA: Addison-Wesley.