

Evaluation of ICT Use in Language Education: Why Evaluate, Where to Look and with What Means?

Faiza Derbel

Abstract The aim of this chapter is to discuss how best to go about evaluating the use of Information and Communication Technologies (ICTs) in language education and to propose guidelines for future research in this area. It begins by deconstructing “evaluation” as a concept and reviewing the status of evaluation studies in English language teaching (ELT) drawing first on meta-analyses and reviews of research on Computer-Assisted Language Learning (CALL) and distance education, and then focusing on the shift in evaluation practice in light of emerging areas of ICT application facilitated by mobile devices and Web 2.0. Evaluation of ICT use will be examined in terms of its foci and emerging methodologies to highlight the need for taking into account the complexity of learning and teaching with ICTs to measure and confirm the “alleged” benefits of ICT use and also to develop knowledge and understanding of the use of ICTs in language education in various contexts. The paper ends by pointing to directions for evaluation studies of relevance to researchers in countries of the Middle East and North Africa (MENA) that this Handbook is targeting.

Keywords Evaluation · Frameworks · Research methodologies · Incorporation of ICT · Web 2.0 · MENA region · Language education

1 Introduction

This chapter delves into the issue of evaluation of the use of Information and Communication Technologies (ICTs) in education with specific reference to English language education. It discusses what methodology to adopt to carry out evaluation studies involving the use of ICTs. That is, how to frame the questions,

F. Derbel (✉)

Faulty of Letters, Arts and Humanities, University of Manouba, Manouba, Tunisia
e-mail: fderbel26@gmail.com

how to determine their purpose, and what research designs to adopt in light of current technological advances. For a start, what is meant by ICTs today (2015) and how is evaluation defined? According to the World Bank Group (2000), ICTs include any form of “hardware, software, storage, processing, transmission, and presentation of information (voice, data, text, images) (n.p.).” Recognized ICTs include, and the list is not exhaustive, multimedia authoring tools, various forms of distance learning platforms, video conferencing systems, mobile technology, social network platforms (Facebook, My Space, Skype, etc.), and various interactive and engaging devices. Educators are understandably attracted to engaging in what is now known as “ICT-enhanced,” “ICT-supported” or “ICT-based” instruction. Judging from the meta-analyses and research reviews of technology use available (Bax, 2003; Burston, 2015; Chappelle, 1998, 2003; Cox & Marshall, 2007; Debski, 2003; Liu, Moore, Graham, & Lee, 2003), teachers have been experimenting for more than half a century with hardware, software, learning systems, networks, mobile devices, downloadable applications and the abundant resources available on the web and in the cloud (Dudeney & Hockly, 2012). Educators are pressed to evaluate the new tools and learning spaces their students are using but are not certain whether the current methodologies are adequate for the task (Goodwin-Jones, 2005a, b, 2011; Greenhow, Robelia, & Hughes, 2009).

Technology-using teachers are generally attracted to using ICTs with their students and to accessing the wealth of digital resources and authoring tools available to them to introduce new input or to design innovative learning activities (Greenhow et al., 2009). Indeed, some teachers found that individualized Skype sessions with team members in the process of completing project work can be integrated within their courses to facilitate communication with their students (Goodwin-Jones, 2005b). Tasks such as reserving an air ticket, checking a bank account, reading news online, and applying for a job can become more authentic thanks to ICT use (Smyth & Mainka, 2006).

However, while teachers may embrace technology, they also assume the responsibility for any decisions they make in their own classrooms in terms of what can be revealed as the impact/outcome effect of ICT use on their learners’ school results. What if the taken-for-granted “beneficial” aspects of ICTs do not reflect on their performance? What if the virtues of the application cannot be appreciated by the learners? Is there a way to measure the educational potential of ICT and gauge the degree of its success/failure? One way of resolving the uncertainties is to evaluate the use of ICT in context. That is, investigate whether the ICT-supported instructional mode is leading to improvement based on “proof” from the learners themselves and/or following an appraisal of the situation by applying a given set of measurable criteria. Ideally, evaluators (whoever they may be) will try to establish a relationship between ICT use and learning effects (conceptualized as gains in achievement, improved motivation or changed attitude). Evaluation can be tied up to assessing compatibility with “desired” learning objectives, curriculum specs, and performance targets. This chapter delves into the issue of evaluating ICT use in language education which is, on hindsight, a straight forward task but has become complex in light of the advances in technology and the widening of the scope of

teaching/learning with ICTs. It will be argued in this paper that data-driven context-specific evaluation studies carried out by teachers-as-researchers is the most promising approach.

2 Defining Evaluation

It has become difficult to define the term evaluation or delimit its scope. As Harland (1996) rightly points out, “the term itself has been stretched and stretched to encompass an ever-widening range of activities, undertaken for an ever-increasing range of purposes (p. 91).” Simonson, Smaldino, Albright, and Zvacek (2009) suggest “evaluation (...) is the systematic investigation of the worth of an on-going or continuing distance education activity (p. 349)” which highlights the merit of an online program in terms of strengths, weaknesses, benefits and drawbacks. Therefore, setting criteria and benchmarks are important for the practice of evaluation as well as deciding who is well-placed to carry out the evaluation. As the title of this paper suggests, there is first a need to consider the reason for the evaluation, define the nature of the task and decide what (methodology) can be used as proof of “good use” or value-added to instruction and learning. Evaluators need to look for evidence (sources of data) and select the strategies to collect the needed evidence and reach results by using the necessary (and viable) analytical tools. By way of illustration, suppose that the teaching of a particular course is to be done through the use of a software “extolled” by its developers and marketers for its ability to make learning and teaching effortless, enjoyable and manageable within record time. In this case, a clear method of evaluation and fixed criteria for acceptance (features and properties) can help the decision maker (teacher or administrator) decide whether such software is relevant to the curricular goals in her specific context (Reiser & Kegelmann, 1994). Evaluation in that sense consists of determining the quality of the software beforehand but leaves open the question of who will do the evaluating, what information is needed and what procedures can be used to collect it.

There is certainly an expert side to evaluating instructional software as the process will entail knowledge of the technical and teaching skills required to exploit the useful aspects of this software. One way of carrying out the evaluation is to experiment with the software with a small number of students and determine its worth. In a scenario described by Reiser and Kegelmann (1994), the learners can be invited to use the software, the teachers observe them as they use it, and then ask them to express their opinions about the software. Indeed, many teachers I know engage in this off-the-cuff evaluation working with a trial version. As reflective practitioners (Nunan & Lamb, 1996), teachers are constantly confronted with situations that require evaluation of new tools (e.g., interactive whiteboard) or new content (e.g., a reading from an online magazine or a conversation thread on a blog) in order to reach decisions about the applicability of the ideas. Therefore, evaluation of ICTs is part of the day-to-day activities of technology-using teachers while planning, delivering and assessing the degree of success of their attempts.

Besides engaging in informal evaluation as potential users of ICTs in their work, teachers can find themselves subjected to evaluation as part of an institutional plan to introduce ICTs. In this case, the evaluation consists in appraising the progress of an intervention by an outsider looking for data evidence (e.g., coming from observation of a teacher in action) by an “evaluation expert” to determine whether and to what extent the teaching strategies the teachers are using match the “desirable” practice promoted by the intervention (Hedberg, 2011, pp. 9–12). Evaluation of interest to policy makers and funding bodies generally revolves around macro-level concerns related to cost effectiveness, programme fidelity, availability of infrastructures, specific applications, access, enrolment and drop out levels, etc. The evaluation methods in these circles are somewhat canonized considering the long history of what is now established practice within funding organizations (e.g., The World Bank, UNESCO, and OECD). Funding bodies tend to generalize their own methods of information gathering, discursive practices, measurement techniques and interpretative frames of reference (see Wagner et al. 2005). This type of evaluation work serves the purpose of *monitoring and evaluation* whereby the constructs are broken down into *intended outcomes* (teachers’ technical and pedagogical skills and learners’ desired outcomes, information skill, attitudes, and so on) and *Moderating factors* (level of community support, access to ICTs, and availability in the home or community) (Wagner et al., 2005, p. 8). Survey data can be collected and submitted to statistical analysis to obtain results about student outcomes by way of indications of increased knowledge of school subjects and/or improved attitudes about learning. *Monitoring and evaluation* is a frequent occurrence in the evaluation discourse of international funders because the function of the evaluation, for ill or for good, is to warrant the good management of the projects they are funding. Evaluation in this case serves a control function (Harland, 1996) and is not necessarily conceived to explain practice or entangle intervening factors.

This chapter is written with the teacher-as-researcher and research student in mind and not the “expert” evaluator working for an international body. As academic and researcher functioning in a technology-challenged educational context, am often challenged by teachers in the audience questioning whether technology makes a difference and cynically ask: What kind of language are the learners likely to “pick up” from peers during online collaborative work? A more encouraging question came from a learning system provider who asked: “What can be the best way to evaluate what we are doing in Tunisia?” Of interest to me and my students is the evaluation of ICT use to investigate “the potential of technology for language learning” (Chapelle, 2003, p. 36). Evaluation can focus on learners working with ICT to pinpoint, for instance, what strategies they use to complete a task to provide evidence for the quality of the occurring exchanges which can, in turn, serve as evidence for the learning taking place as a process or as product. Evaluation of teaching with ICTs means, in this paper, any activity involving the exploration of specific situations to see whether the use of ICTs is anchored in criteria of “good practice,” “good results” and assumptions about quality teaching/learning (Simonson, Smaldino, Albright, & Zvacek, 2009). The sense making to result

from the exploration and through the examination of a software, courseware or teacher-developed materials and the learning arising during implementation are taken to be context-specific and personalized by the teachers in-action (Schön, 1983). In this type of evaluation, there is as much to gain for the teachers who will see for themselves how they fared as users of ICT. In the title of this paper I ask: *Why evaluate, where to look and with what means?* That is, the evaluator is required to determine the purpose of the evaluation as practice, fix targets for the evaluation (where to look) and determine and deploy the research instruments and analytical tools needed. By doing so, it is hoped that this chapter gives the inspiration, conceptual understanding, and practical guidance for prospective evaluation researchers to proceed on scientific grounds and explore with confidence situations of interest in their own context.

3 Conceptualizing Evaluation

Most teaching resource books used in teacher education and training include a section at least on evaluation. Authors (e.g., Reece & Walker, 1992; Simonson et al., 2009; Smyth & Mainka, 2006) encourage teachers to engage in evaluation of their learners and their own teaching by providing them with easy-to-use checklists, ready-made questionnaires, and sample interview protocols for them to apply step by step. These teacher educators/trainers are keen on evaluation because they see in the practice an opportunity for teachers to see for themselves how they are faring as teachers and how their learners are progressing. These resource books are recommendable as starting points for novice researchers interested in evaluation studies in school contexts. More specialized evaluation books (e.g., Hopkins, 1989) are also helpful in explaining the distinctions between different approaches to evaluation.

As far as evaluating ICT use, researchers would need to distinguish between formative and summative evaluation. Summative evaluation means that the software is evaluated while used by teachers and learners in order to determine what the outcomes are as revealed in test scores while formative evaluation means trying out software before designing and delivering instruction so that a decision is reached about its potential for learning (Lan, Sung, & Chang, 2013; Levy, 1999; Shaughnessy, 2003). Alternatively, Chapelle (2003, p. 81) argues that the evaluation can be equally focused on the *product* by measuring *what* has been learned as specified in the objectives of the instruction or *process* of learning by collecting information about *how* the software was used by learners and to do what learning activities. She adds that these distinct evaluation types can, when combined, contribute to a more exhaustive description of the situation of ICT use and help sharpen views of the issues surrounding the practice of teaching with technology. Furthermore, Ruhe and Zumbo (2009) propose with reference to distance education, focusing on the course's *underlying values* and *unintended consequences* as additional features of the implementation to uncover the “hidden” aspects of a distance course to complete the picture. The underlying values of a course, they

suggest, can be inferred from the language used to describe it (found in course outlines, descriptions of the goals, standards, and assessment principles). Labels like “innovative” and “cutting-edge” can be used to infer the intent and contrast it with the outcomes to see if there is a match. Unintended instructional consequences can be technical flaws, high dropout rates, or witnessed discrepancy among the course components. As for the unintended social consequences, these include, for instance, isolation of the distance learner and the need for the tutor to play the role of coach and facilitator (Derbel, 2013).

Judging from developments in the field of CALL (Chapelle, 2001, 2003; Levy, 1997), evaluation practice focusing on the measurement of *effects, impact, outcomes, efficiency*, was considered the legitimate way of reaching answers about ICT use which was at that time needed to make the case for technology or to convince administrators, parents and teachers to generalize the experiences of high profile projects (Chapelle, 2003, pp. 70–73). Indeed, in the early days of CALL (1960s) high profile projects in the US and UK were evaluated with comparisons with traditional teaching situations. The main purpose of the evaluation was to establish whether the PLATO learning system, for instance, had a positive or negative effect on learners’ achievements and attrition rates compared with traditional classes. The field moved on to develop more contextualized and sophisticated methodologies. The approach changed in the 1980s and 1990s as teachers began to play a bigger role in programming and developing their own materials, taking advantage of the availability of user-friendly authoring tools (Levy, 1997, p. 43). The objective of evaluation is no longer the justification of spending but rather exploring and understanding the learning processes as they unfolded during implementation (Garcia-Villada, 2009). As a result, evaluation criteria and methodological solutions were needed to help teachers make sense of the situations they were documenting and analysing from learning/teaching frames of reference. Chapelle (1994), for instance, illustrated how CALL activities can be researched to determine the quality of the communication resulting from their use, drawing in this case on an interaction analysis framework to produce a description of the learner-computer interactions, the language produced and the learner’s level of engagement with the materials.

In her 2003 book, *English Language Learning and Technology*, Chapelle succinctly summarizes evaluation research in CALL and demonstrates how researchers have successfully used frameworks from allied disciplines (Second Language Acquisition, Discourse Analysis, Testing and so on) and analytical tools fit to the focus and purpose of the evaluation (learner, teacher, task, the interactions, or output). By way of illustration, Chapelle (2003, pp. 82–108) explains that researchers can investigate how learners use the options in the software, observe and document whether learners use the help functions (e.g., subtitling or annotations) and the feedback built into the software. Results of this evaluation can be used by software developers to improve the design and increase its learning potential. Similarly, she points out that content analysis has been “successfully” used to analyse the language of learners as they were performing computer-mediated communication (CMC) tasks and describe the characteristics of

the register used, the written interactive discourse and other moves. Inferences can then be made about the value of the task and the language abilities of the learners.

The evaluation work helped with its cumulative effect to establish the field of CALL and the value of teaching/learning with technology. The studies overviewed in Chapelle's (2003) book can inspire novice evaluators when looking for an area of focus (learner, teacher, communication tasks, CALL materials, and multimedia activities), types of evaluation (formative/summative or product/process), research designs (experimental, case study or action research), framework for analysis and interpretation (content analysis, interaction analysis, discourse analysis) and positioning of the evaluator(s) (insider or outsider).

Most importantly, the shift starts by asking the question differently (Cox & Marshall, 2007). Instead of asking whether teaching with technology is "better than" teaching face-to-face or "more effective" in improving learners' test scores, it is more informative to ask questions about how the experience of learning L2 with CALL went and find evidence for its contribution to learning language. The evaluator can examine the situation from multiple perspectives and collect quantitative as well as qualitative data about the on-going implementation of technology-supported lessons. Jamieson et al.'s (2005) study illustrates a possible design they used to evaluate the *Longman English Online (LEO 3)* during implementation. Chapelle's (2001) six criteria of CALL evaluation (language learning potential, learning focus, learner fit, authenticity, positive impact, and practicality) were used to examine the materials from the perspectives of the developer, teacher and the learners. Their scheme of data elicitation included questionnaires, reflections, and interviews based on specific evaluation questions for each criterion. The "numeric summaries (p. 32)" obtained from test scores and questionnaire analysis were supplemented with interview and guided reflection data. Conclusions were then reached about how *LEO 3* matched the six criteria and how it was appreciated by this group of learners and the teacher. To evaluate teacher developed reading materials while used by learners, Derbel (2001) resorted to a screen capturing software (*Camtasia*) to observe the learners' "navigation paths" and "look-up behaviour" which were later corroborated with the self-reporting data obtained from learners by means of a stimulated recall protocol. Strengths and weaknesses of the materials and design options were revealed and juxtaposed with the teacher's pedagogical intent.

Another evaluation study reflecting the shift is Lan et al.'s (2013) action research study focusing on a proposed mobile-supported cooperative reading system. Their design reflected an attempt to bridge evaluation and pedagogical concerns by incorporating formative and summative evaluation within a two-loop action research cycle. They collected data before and while teaching, followed by repair cycles of the system's functions and interface. The teachers taught two units and administered pre- and post-tests to measure their students' attitudes examined along five dimensions of the interface and functions of the system.

Evaluation work of teacher practices is also a possible focus. To trace and construct the learning processes of teachers making the transition from traditional teaching to teaching with interactive whiteboards, learning objects and network links, Hedberg (2011), for example, developed the "concerns-based adoption

model” (p. 4) by drawing on the literature on teacher acceptance of innovation and was able to trace the teachers’ learning curves and changes in their pedagogy with the whiteboard from the narratives he collected over the semester. He was also able to determine the factors shaping the quality and pace of their learning to teach with the interactive whiteboard and learning objects. This study design reflects a shift in focus from the technology itself to the teacher implementing it. Similarly, Starkey (2011) developed “the digital age learning matrix” (p. 24) as a research tool to evaluate how digitally-able beginning teachers taught lessons using digital technologies. He used the matrix as a reference to evaluate whether the teachers’ practice is aligned with the principles of constructivism (taken to be the learning theory compatible with the requirements of 21st century digital learning).

The studies reviewed so far indicate that there are generally two perspectives: The psychometric type and the more interpretative type. A note of caution is due, however. It has been pointed out earlier that experimental designs fell out of favour in the 1980s and 1990s among teachers interested in process research but meta-analyses and reviews of research focusing on impact/effect of ICT on learning indicate that this research tradition has not been abandoned (Cox & Marshall, 2007; Liu et al., 2003; Ting, 2005). Despite the detected methodological “flaws” and shortcomings in a big number of studies which did not qualify for inclusion, these authors merely call for the sharpening of the research practices by collecting more in-depth data and increasing the length of time spent in the field (i.e., carry longitudinal studies). Felix (2008) praises Allum (2004) and Nutta et al. (2002) for incorporating delayed post-tests and self-reporting data to measure learners’ attitudes towards computers and school subjects. Liu et al. (2003) mention Brett (1997) who compared multimedia technology and simple audio and video equivalents to prove the superiority of multimedia CALL. There must be new concerns for researchers who find themselves working with the Web 2.0 generation and facing new challenges. I outline the following observed trends.

Greenhow et al. (2009) note that the most prominent change consists in the broadening of the conceptualization of “classroom” and that “[t]oday learners have more choices about how and where to spend their learning time (e.g., in online settings or in private, public, or home school options) than they did 10 years ago” (p. 247). Therefore, to recall part of the question I am asking in this paper (“*where to look?*”), it seems that any attempt to evaluate the use of Web 2.0 tools (e.g., social networks, media sharing, blogs, etc.) will imply data collection from different electronic and physical spaces. More innovative data collection methodologies (also outlined by Goodwin-Jones, 2011) are emerging to evaluate learning in Web 2.0. like web surveys, digital photography and movies, voice recognition tools, and more powerful data aggregation tools.

Evaluating learning in Web 2.0 can be guided by questions about the learning opportunities (e.g., creation of content), the role(s) teachers play in the process considering the open-endedness of virtual spaces, the education value of learners’ participation in Web 2.0 and how they bridge the learning opportunities in different spaces (in and out-of-school and on social networks). Currently, researchers started to dispute the impact of mobile technologies and constant internet connectivity on

academic performance. Researchers are asking whether new working styles are developing and whether these are conflicting with the learning objectives proposed by teachers. For example, Rosen, Carrier, and Cheever (2013) attempted to assess the level of distraction during task performance and Lepp, Barkley, and Karpinski (2015) whether learners are able to self-regulate to resist the distraction. Thus, they are seeking through their work to reach answers and clarifications about the opportunities and the risks for academic performance. Other researchers are more positive in their outlook to the use of mobile phones in L2 learning. Burston (2015), for instance, singles out Chen, Chang, Lin, and Yu (2009), Liu (2009) and Tai (2012) as exemplifying the use of mobile phones in communicatively-oriented activities. Therefore, it appears that more up-to-date methodologies are not necessarily being used with recent technological innovations.

Evaluation work can help make the case for technology use in language education; i.e., establish benefits, gains in achievement, learning potential associated with a particular situation of ICT-supported teaching. Evaluation can lead to better understanding of the processes facilitated by the introduction of an ICT tool and how it contributes to language learning. More informative evaluations are the ones which yield thorough descriptions of the tools used, level of learners involved, the relevance for the curriculum or lesson, the support mechanisms, and the learning theory embedded in the materials and implemented by the teacher. Based on the discussion in the previous sections, what is known today has been made possible by revising the methodologies used and especially by theorizing from the ground naturally occurring teaching (Guba, 1981) with the involvement of the main actors in the research process. The ultimate goal of evaluation is to provide answers and give clarifications about the use of ICTs where “mixed feelings” about their worth are shared among funders, administrators and/or teachers and parents. Evaluation studies can also reach unintended, unexpected or conflicting results (Felix, 2008).

4 Recommendations for Future Research

In the beginning of this paper, I had hoped that the analysis would lead to providing guidelines for novice teachers they can use as points of departure for the design of evaluation studies that tap into the implementation of ICT-supported teaching/learning in the MENA region. I will start from my understanding of my own context in Tunisia and hope that readers in the region can establish parallels relevant to their own contexts. What has been done on the issue of ICT use in education in Tunisia is somewhat patchy. Much of it has not been documented even though a good number of teachers I know have been trying ideas for a number of years and some are quite advanced technology users. Moreover, little information is available on the official policy related to ICT use in education. A report by Hamdy (2007) included a listing of the state-led projects and policy decisions from 2001 upwards focusing on the introduction of ICT as a compulsory subject in schools. Another book by Chabchoub and Bouraoui (2004) included elements of the story of

state policy such as the Family Computer project in 2000 or background to the creation of the Virtual University of Tunis (UVT) in 2002 and subsequent plan(s) to “digitize” 20 % of the courses at university by 2006–2007. UVT has now made strides into course development according to the information on their websites (number of courses online) and has put in place a teacher training program for tutors. Moreover, Master’s level research studies are relatively few (Ben Gayed, 2007; Ben Hammed, 2012; Charfi, 2010; Klibi, 2014; Lachheb, 2013) and generally focused on “effects,” “motivation” and “hurdles and obstacles”. Therefore, there is a lot left to explore about actual implementation.

As this chapter is meant to guide researchers interested in evaluating ICT use in educational contexts, it is important to remind them that research methodology is a set of options and solutions to assist them in finding answers to their questions. Therefore, if research students (or practitioners for that matter) are interested in exploring the “worth” or “outcome” of applying a software, using an authoring tool, or adopting a social network platform in their own school, they can find guidance in these reviews. Other possible studies of the evaluation type can be, for instance, conceived to answer a question like: “*What happens when learners perform a writing task on a class blog?*” The researcher can resort to digital data recording (using an outside camera focusing on learners and a “tracking” software focusing on what happens inside the computer) collected during the performance of the writing task. The operational data can be transcribed and coded according to a system which is likely to serve the researcher’s purpose (see Chapelle, 2003; Derbel, 2001; Jamieson & Chapelle, 1987). Such rich data can help uncover on-task/off-task behaviour, strategies for task completion, the teacher’s role/presence (if the teacher is involved), and online/offline dialogue with peers and teacher (whether instances of negotiation of meaning and request for help do occur). Taken together, these details can help researchers in the process of interpretation (relying on an appropriate theoretical framework) in order to make inferences and reach conclusions about the quality of the learning experiences (the aim of the evaluation). In the end, the results should contribute to clarifying how blogging can be considered a “good” or “beneficial” way of teaching writing supported by evidence from the data. The following sections are suggestions for possible studies to evaluate ICT use in language education.

I begin by suggesting that evaluation studies focus on ICT-supported teaching of the traditional language skills (listening, speaking, reading and writing). The research design should be powerful enough to capture the intricacies of the situation. A mixture of process and product data and examine them drawing on theoretical models, (e.g., Goodman’s (1968) interactive reading model), to design and determine the quality of the instruction and learning experiences. Focus can be on task variation, diversification of the reading resources, individual learning strategies, or group dynamics. Second, the object of the evaluation can be a specific ICT tool and its application to teach learners of different ages or of different specializations (e.g., medical doctors). A third area can be based on Greenhow et al.’s (2009, p. 250) suggestion to study modes of self-expression (e.g., when learners produce digital content by means of a specific authoring tool) and using it to

communicate in various learning spaces in both formal and informal settings. I have learnt when invited to a Tech Age Teachers' meeting in 2014 (see <https://www.irex.org/projects/tech-age-teachers-tunisia>) that teachers across sectors and levels are being trained to integrate technology in their teaching and are out there creating teaching/testing materials, class websites, blogs and wikis. A multiple case study of these teachers will be of interest and a good starting point for the appraisal of the philosophy of empowering teachers with technological know-how and their experience with incorporating technology in their teaching in natural settings (Guba, 1981). Finally, but not exclusively, researchers/teachers can evaluate attempts to blend face-to-face teaching with online collaboration projects on social networks with a special emphasis on the process and product. Chapelle's (2003, p. 110) advice about generating CALL text from CMC exchanges can be extended to the new electronic tools and spaces teachers are using today.

5 Conclusion

I have, in this chapter, discussed evaluation as a concept and illustrated how it was used to appraise attempts to implement ICT-supported teaching. I have argued that evaluation can be conceptualized differently and that methodologies have to be compatible with the questions asked and explained that to this day researchers continue to look for answers about the value of technology use for learning but methodologies evolved over the decades as researchers grappled with the questions (Bax, 2003; Chapelle, 2003; Dudeney & Hockly, 2012; Levy, 1997). Theories, conceptual frameworks and analytical tools from already established fields have proven useful (Chapelle, 2003; Liu et al., 2003). However, completely innovative practices may be needed to match the nature of evolving uses and practices and concerns (Greenhow et al., 2009; Goodwin-Jones 2011; Starkey, 2011). I will end with Levy's remark about CALL research in 1990s: "(...), evaluation is crucial if CALL is not to be entirely technology-led, and if we are to identify and build upon prior successes" (p. 41). In Tunisia, for instance, teacher educators and practitioners are cast in the role of recipients of training in technological skills with no research activity to follow up implementation and teacher mediation. The research I am suggesting can help bring into focus the role of teachers in ICT use in language education.

References

- Allum, P. (2004). Evaluation of CALL: Initial vocabulary learning. *ReCALL Journal*, 16(2), 488–501.
- Bax, S. (2003). CALL—Past, present and future. *System*, 31(1), 13–28.

- Ben Gayed, L. (2007). *Effects of CALL on motivation of Tunisian learners of English for specific purposes* (Unpublished Master's thesis). University of Sfax.
- Ben Hammed, R. (2012). *Normalisation: Obstacles and solutions* (Unpublished Master's thesis). University of Manouba.
- Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25(1), 39–53.
- Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, 27(1), 4–20.
- Chabchoub, A., & Bouraoui, K. (2004). *Introduction à la pédagogie numérique [An Introduction to digital pedagogy]*. Publication de l'ATURED.
- Chapelle, C. A. (1994). CALL activities: Are they all the same? *System*, 22(1), 33–45.
- Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology*, 2, 22–34. Retrieved 09 December 2015, from <http://www.llt.msu.edu/vol2num1/article1/index.html>
- Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge, UK: Cambridge University Press.
- Chapelle, C. A. (2003). *English language learning and technology*. Philadelphia, USA: John Benjamins Publishing Co.
- Charfi, H. (2010). *A comparison of the metacognitive reading strategies applied by self-regulated EAP students in print and online environments* (Unpublished Master's thesis). University of Sunderland, UK.
- Chen, T.-S., Chang, C.-S., Lin, J.-S., & Yu, H.-L. (2009). Context-aware writing in ubiquitous learning environments. *Research and Practice in Technology Enhanced Learning*, 4(1), 61–82.
- Cox, M. J., & Marshall, G. (2007). Effect of ICT: Do we know what we should know? *Education and Information Technology*, 12, 59–70. doi:10.1007/s10639-007-9032-x
- Debski, R. (2003). Analysis of CALL (1980–2000) with a reflection on CALL as an academic discipline. *ReCALL*, 15(2), 177–189.
- Derbel, F. (2013). Facilitation of learning in electronic environments: Reconfiguring the teacher's role. In M. Ciussi & M. Angier (Eds.), *Proceedings of the 12th European Conference on E-Learning, ECEL* (pp. 94–100). Academic Conferences & Publishing Int., Ltd.
- Derbel, F. (2001). *The integration of CALL in the ESL classroom: Reconciling agendas* (Unpublished Master's thesis). Iowa State University, Ames, Iowa, USA.
- Dudeney, G., & Hockly, N. (2012). ICT in ELT: How did we get here and where are we going? *English Language Teaching Journal*, 66(4), 533–542.
- Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learnt in two decades of research? *ReCALL*, 20(2), 141–161.
- Garcia-Villada, E. (2009). CALL evaluation for early foreign language learning: A review of the literature and a framework for evaluation. *CALICO Journal*, 26(2), 363–389.
- Goodman, K. S. (1968). The psycholinguistic nature of the reading process. In K. S. Goodman (Ed.), *The psycholinguistic nature of the reading process* (pp. 15–26). Detroit, MI: Wayne State University Press.
- Goodwin-Jones, R. (2005a). Emerging technologies, messaging, gaming, peer-to-peer sharing: Language learning strategies & tools for the Millennial Generation. *Language Learning & Technology*, 9(1), 17–22. Retrieved 01 February 2015, from <http://llt.msu.edu/vol9num1/pdf/emerging.pdf>
- Goodwin-Jones, R. (2005b). Emerging technologies, Skype and podcasting: Disruptive technologies for language learning. *Language Learning & Technology*, 9(1), 17–22. Retrieved 01 April 2015 from <http://llt.msu.edu/vol9num3/pdf/emerging.pdf>
- Goodwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, 15(2), 2–11. Retrieved 01 February 2015, from <http://llt.msu.edu/issues/june2011/emerging.pdf>
- Greenhow, C., Robelia, B., & Hughes, J. (2009). Learning, teaching, and scholarship in a digital age, Web 2.0 and classroom research: What path should we take now? *Educational Researcher*, 38(4), 246–259.

- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*, 29, 57–91.
- Hamdy, A. (2007). Survey of ICT and education in Africa: Tunisia country report. The World Factbook, 2007. Retrieved 05 July 2015, from <https://cia.gov/cia/publications/factbook/geos/ts.html>
- Harland, J. (1996). Evaluation as Realpolitik. In D. Scott & R. Usher (Eds.), *Understanding educational research* (pp. 91–105). London, UK: Routledge.
- Hedberg, J. G. (2011). Towards a disruptive pedagogy: Changing classroom practice with technologies and digital content. *Educational Media International*, 48(1), 1–16. Retrieved 01 February 2015, from <http://dx.doi.org/10.1080/09523987.2011.549673>
- Hopkins, D. (1989). *Evaluation for school development*. Buckingham, UK: Open University Press.
- Jamieson, J., & Chapelle, C. A. (1987). Working styles on computers as evidence of second language learning strategies. *Language Learning*, 37, 523–544.
- Jamieson, J., Chapelle, C. A., & Preiss, S. (2005). CALL evaluation by developers, a teacher, and students. *CALICAO Journal*, 23(1), 1–44. <http://www.tesl-ej.org/pdf/ej64/a2.pdf>
- Klibi, A. (2014). *Implementation of computer assisted language learning in Tunisia: EFL teachers' perceptions and perspectives* (Unpublished Master's thesis). University of Manouba.
- Lachheb, A. (2013). *Information technology effects on Tunisian college students, Tunisian English majors as a case study* (Unpublished Master's thesis). Grand Valley State University, USA.
- Lan, Y.-J., Sung, Y.-T., & Chang, K.-E. (2013). From particular to popular: Facilitating EFL mobile-supported cooperative reading. *Language Learning & Technology*, 17(3), 23–38. Retrieved 01 April 2015, from <http://llt.msu.edu/issues/october2013/action.pdf>
- Lepp, A., Barkley, J. E., & Karpinski, A. C. (2015). *The relationship between cell phone and academic performance in a sample of US college students* (pp. 1–9). Jan-March: Sage Open.
- Levy, M. (1999). Design processes in CALL: Integrating theory, research and evaluation. In K. Cameron (Ed.), *CALL: Media, design, and application* (pp. 459–469). Exton, PA: Swets & Zeitlinger.
- Levy, M. (1997). *Computer-assisted language learning: Context and contextualization*. Oxford, UK: Oxford University Press.
- Liu, T.-Y. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning*, 25(6), 515–527.
- Liu, L., Moore, Z., Graham, L., & Lee, S. (2003). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990–2000. *Journal of Research on Technology in Education*, 34(3), 250–273.
- Nunan, D., & Lamb, C. (1996). *The self-directed teacher: Managing the learning process*. Cambridge, UK: Cambridge University Press.
- Nutta, J. W., Feyten, C. M., Norwood, A. L., Meros, J. N., Yoshii, M., & Ducher, J. (2002). Exploring new frontiers: What do computers contribute to teaching foreign languages in elementary school? *Foreign Language Annals*, 35(3), 293–306.
- Reiser, R., & Kegelmann, H. (1994). Evaluating instructional software: A review and critique of current methods. *Educational Technology Research and Development*, 42(3), 63–69.
- Reece, I., & Walker, C. (1992). *A practical guide to teaching, training and learning: A practical guide*. Tyne and Wear, UK: Business Education Publishers Ltd.
- Rosen, L. D., Carrier, M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task switching while studying. *Computers in Human Behaviour*, 29, 948–958.
- Ruhe, V., & Zumbo, B. D. (2009). *Evaluation in distance education and e-learning*. New York: The Guilford Press.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York, USA: Basic Books.
- Shaughnessy, M. (2003). CALL, commercialism and culture: Inherent software design conflicts and their results. *ReCALL*, 15(2), 251–268.
- Simonson, M., Smaldino, S., Albright, M., & Zvacek, S. (2009). *Teaching and learning at a distance: Foundations of distance education* (4th ed.). Pearson Education, Inc.

- Smyth, K. & Mainka, K. (2006). *Pedagogy and learning technology: A practical guide*. Edinburgh Napier University. Published online by Creative Commons, 2010. Retrieved 01 February 2015, from <http://www.creativecommons.org/licensing/by-nc-sa/2.5>
- Starkey, L. (2011). Evaluating learning in the 21st Century: A digital age learning matrix. *Technology, Pedagogy and Education*, 20(1), 19–39.
- Tai, Y. (2012). Contextualizing a MALL: Practice design and evaluation. *Educational Technology & Society*, 15(2), 220–230.
- Ting, S. E. (2005). The impact of ICT on learning: A review of research. *International Educational Journal*, 6, 635–650.
- Wagner, D. A., Day, B., James, T., Kozma, R., Miller, J., & Unwin, T. (2005). *Monitoring and evaluation of ICT in education projects: A handbook for developing countries*. Washington, DC: InfoDev/World Bank. Retrieved 01 December 2014, from <http://www.infodev.org/en/Publication.9.html>
- World Bank. (2000). *Monitoring and evaluation: Some tools, methods and approaches*. Washington, D.C.: World Bank Group. Retrieved 01 December 2014, from <http://www.worldbank.org/oed/ecd/>