Lang. Teach. (2010), **43:1**, 66–74 © Cambridge University Press 2009
doi:10.1017/S0261444809005850 First published online 18 May 2009

Plenary Speeches

The spread of computer-assisted language learning

Carol A. Chapelle Iowa State University, USA
carolc@iastate.edu

This paper argues that the vertical spread of computer-assisted language learning (CALL), i.e., a spread throughout language materials and curricula, makes it difficult to draw a clear distinction between CALL and other language materials. In view of the emphasis that teachers, researchers, and administrators have placed on evaluating CALL, I argue that some valuable lessons about materials evaluation can be drawn from reflection on issues in CALL evaluation. In particular, I discuss the opportunities for professionals to reconsider assumptions held about comparative research, draw upon research perspectives and methods from applied linguistics in materials evaluation, and include critical perspectives which examine the opportunities that materials offer language learners to engage in language and culture learning.

1. Introduction

The expression 'computer-assisted language learning' (CALL) refers to a variety of technology uses for language learning including CD-ROMs containing interactive multimedia and other language exercises, electronic reference materials such as online dictionaries and grammar checkers, and electronic communication in the target language through email, blogs, and wikis. These varied technologies used by language learners have spread over the past several years across many language classrooms and beyond. This spread can be thought of in two ways. One is the GEOGRAPHICAL (or horizontal) spread that takes technology to so many language learners around the world. The other is the VERTICAL spread throughout the language curriculum. The horizontal spread is an important area of inquiry as the social, political, and economic realities intersect with access to technology for learners in different parts of the world. However, my focus in this paper is the vertical spread of technology throughout the language class and curriculum.

In the past, the language curriculum and classroom were examined in terms of the textbooks used as well as the language and behavior of the teachers and students. Today, however, the textbook often comes with a CD-ROM and has a companion website. Some textbooks are intended to be used with online materials in a learning management system that the publisher maintains and the instructor may have a WebCT or Moodle course set up

A revised version of a plenary paper presented on 6 June 2008 at the Canadian Association of Applied Linguistics Conference, Vancouver, British Columbia.

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

as well. The course activities on the Web would not necessarily be limited to these, as learners might use email, instant messenger and Skype to communicate with other learners or other speakers of the target language at a remote location. In fact, the students and teacher may not ever meet in a classroom together, but instead may meet virtually on a course website in a distance-learning program.

In view of today's realities in the United States, Canada and many other countries, the historically-constructed line between applied linguists who work in CALL and those who produce other forms of language learning materials is difficult to maintain and not very useful. In a sense, today almost anyone who is working on materials for classroom language learning is working in CALL. This vertical spread of CALL and the considerable activity surrounding its development and evaluation prompts reconsideration of all language learning materials. This paper discusses three lessons learned from work on CALL that may be relevant for all language learning materials.

2. Evaluating language learning materials

The amount of published work on materials evaluation is surprisingly small in view of the impact that materials have in the instructional process. Tomlinson's edited collections (1998, 2003b, 2008) help to define the area of materials development, but each of the volumes contains some discussion of evaluation (e.g., Ellis 1998; Littlejohn 1998; Tomlinson 2003a; Masuhara & Tomlinson 2008) beyond the evaluative perspectives implicit in the advice about materials development. Tomlinson's (2003a) own paper on evaluation lists nineteen principles extrapolated from research on L2 acquisition. Three examples of these are the following:

- Materials should achieve impact (through novelty, variety, attractive presentation and appealing content).
- Materials should expose the learners to language in authentic use (ideally to rich and varied input which includes unplanned, semi-planned and planned discourse which stimulates mental response).
- The learners' attention should be drawn to linguistic features in the input (so that they are alerted to subsequent instances of the same feature in future input). (p. 21)

On the surface, these suggestions appear similar to the types of criteria that have been suggested for development and evaluation of CALL (e.g., Chapelle 1998, 2001; Doughty & Long 2003). However, in contrast to CALL, work on language learning materials includes little exploration of research methodologies for providing evidence about the quality of materials. Beyond Tomlinson's work, the scholarship in this area consists of just a few more books (e.g., McGrath 2002, McDonough & Shaw 2003) and articles describing evaluation (e.g., Sheldon 1988; Barnard & Randall 1995; Chambers 1997; Tomlinson et al. 2001).

This work spans the theoretical, empirical and practical. However, in view of the amount of time learners spend engaged with their language textbooks, materials evaluation is a topic worthy of more attention from applied linguists who are concerned with instructed L2 acquisition. In many language courses, textbooks guide and shape the choices teachers

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

make about the syllabus, exercises, content, and explanations. What are the perspectives and practices that shape the content of the textbooks? A few anecdotes suggest that this question is worthy of investigation.

One anecdote comes from authors of a beginner-level foreign language textbook who developed exercises modeled on those shown in research to facilitate students' learning of the grammatical points that they were teaching. When the draft materials were sent out for review, reviewers indicated that they did not like the exercises because they were unfamiliar, so the publisher asked the authors to omit them. A second anecdote comes from the topical content of ESL (English as a Second Language) books. Editors screen the topics carefully to be sure that they are not too controversial, discriminatory, boring, or distant. However, anyone from the Midwest United States would have to wonder about the content of one exercise that portrays a farmer in the Midwest as stupid and uncooperative by his role as an obtuse interlocutor in a dialogue in which the student is to form one question after another seeking information from him. A third anecdote comes from the authors of a CD-ROM for French learning. They submitted their materials to the review section of a professional journal, and were told that the journal was not interested in reviewing it because of the French Canadian language varieties it contained!

These anecdotes suggest that evaluation processes ARE used in the development and selection of materials. In the first two cases, evaluation was orchestrated by publishers who call upon the judgments of authors, editors, and reviewers to develop materials that teachers and students like so that they will sell. In the first example, the reviewers' preferences outweighed any research findings. In the second example, the content sensitivity review was undoubtedly done by editors without any sensitivity to cultural stereotypes pertaining to the Midwest United States. In the third example, the personal backgrounds and opinions of individuals came into play as well, as the editor decided that Canadian French was automatically disqualified from the reviews section. The question in these and many other cases is how professional knowledge and practices in applied linguistics can be brought to bear on the issues of materials evaluation in order to diminish the idiosyncratic personal opinions that dominate. Considerably more evaluation work has been conducted on CALL. Without suggesting that these issues have been completely solved in theory or practice, I can find some useful lessons from work on CALL for all materials evaluation.

3. Evaluating computer-assisted language learning

CALL is defined as 'the search for and study of applications of the computer in language teaching and learning' (Levy 1997: 1). The precise teaching and learning practices encompassed by this definition have evolved with computer technology to include use of interactive CALL programs, linguistic and informational resources, and communications programs, but the most telling aspect of the definition remains: the definition portrays work in CALL as inquiry which includes the activities of development, discovery, selection, use, and evaluation of language learning activities that draw upon technology. The definition aptly captures the sense of charting new territory – a factor that is relevant for the distinction made between the needs of CALL evaluation relative to those of evaluation of materials.

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

Hubbard (2006: 313) pointed out an important practical factor that may be at the basis of the distinction: 'many current language teachers have limited experience with CALL software from the learners' perspective and may be novices as well using technology for teaching'.

The perception of those in CALL is that the development and use of CALL is different from that of textbooks, workbooks, videos, and other materials for language teaching. This perception is supported by the fact that the options for language learning with technology are considerably greater than those that can be put between the covers of a textbook or on an audio or videotape. Moreover, when CALL began, it was something novel and expensive. Today, cost is an issue for all professionally-produced materials which contain a combination of digital and paper media. As a result, the professional literature in CALL contains considerable research and reflection on evaluation. This work appears in numerous journals focusing on technology and language learning including *CALICO Journal*, *ReCALL Journal*, *CALL Journal*, *Technology & Language Learning*, and *ALSIC* (*Apprentissage des langues et systèmes d'information et de communication*). It is from this body of work that I have culled three perspectives on evaluation that seem equally relevant to language learning materials, whether they are paper or digital.

3.1. Comparison research

Early research on CALL was designed to show comparisons of learning outcomes from a control (classroom) group with those obtained from an experimental (CALL) group. Today the idea that research on CALL should deliver results allowing for comparisons continues to be presented. Moreover, journals regularly publish studies comparing the outcomes from a CALL teaching unit (Allum 2002) or curriculum (Chenoweth & Murday 2003; Scida & Saury 2006) with a comparable one that does not incorporate computer technology. Researchers interpret results in terms of the particular programs investigated, but such research prompts researchers to attempt to extend beyond decisions about whether or not to use CALL to conclusions about the potential of technology for replacing some aspects of classroom instruction.

The quantitative methodology for systematically combining the results obtained across studies is a meta-analysis. Grgurovic, Chapelle & Shelley (in preparation) conducted such a meta-analysis of research on CALL from studies published from the early 1970s through 2006. Forty-two studies were identified as having a methodology for comparing CALL with a non-technology condition that allowed for their inclusion on a meta-analysis. (The original database of articles included is in Grgurovic 2007). Of those 42 studies, 14 had found that the two groups were equivalent at the time of the pretest. These 14 studies produced a total of 32 effect sizes because of multiple tests given in some of the studies (e.g., assessment of both vocabulary and listening improvement). These 32 effect sizes produce a statistically significant difference in favor of the students using technology over those who had the classroom instruction alone.

This type of meta-analysis that attempts to summarize knowledge for the field about the effects of technology for language learning is of interest to many people. In particular, this is an interesting result for program administrators and teachers charged with making decisions about purchasing computers and designing the curriculum. Applied linguists are less enthusiastic about such a summary because such results fail to provide any information

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

about how or why learning took place through technology. After years of attempting to design such research and interpret the results of comparison studies, CALL professionals recognized serious limitations in this approach to learning about language learning through technology. Pederson's (1987: 106) conclusion from her research expressed the frustration that can result from such work: 'Comparative research on computer-assisted versus non-computer-assisted language instruction is incapable of providing generalizable results'. She discussed three reasons: the impossibility of replication, the uncertainty about the causes of outcomes, and the lack of connection of such research with language learning theory. Today, over twenty years later, one might add to these reasons the fact that technology is often used to change and expand the intended learning outcomes rather than to increase the level of performance in exactly the same areas as those targeted by classroom instruction. For example, learning outcomes from CALL can include students' increased competence in using the technology as a source of help, which can afford them greater autonomy, as well as students' increased intercultural competence through communication with members of the target culture.

Despite the serious limitations in the insights provided by technology vs. classroom studies, applied linguists need to be able to provide such answers that are of interest to some people without failing to recognize the limitations. It is therefore useful to identify the range of audiences for research on CALL (Chapelle 2007). The audiences, which might be placed on a continuum from insiders to outsiders, have their own interests in questions about CALL and their own perspectives from which to interpret results. In sum, applied linguists need to be able to provide results of comparison research, but also move beyond such research to offer evaluation methods that take into account professional knowledge about research on language learning.

3.2. Professional knowledge

Dissatisfaction with experimental methods used to evaluate CALL has prompted researchers to explore alternatives by drawing upon professional knowledge about language and language learning. For example, evidence is gathered to indicate whether or not CALL activities engage learners in tasks that demonstrate authentic language or simulate authentic tasks. These and other positive qualities (e.g., Chapelle 2001) that are investigated in CALL research are very similar to those that have been suggested as guidelines for all language materials. The difference with research on CALL is that researchers have attempted to draw up empirical research methods from applied linguistics to investigate the extent to which evidence exists for such qualities.

Based on hypotheses about the value of interaction for language learning, researchers gather data about the amount and quality of interactions learners engage in as they use technology for language learning. Interactions can refer to interaction between the learner and the computer (e.g., Plass et al. 1998) or between the learner and other interlocutors who are communicating through computer-mediated communication (Sanaoui & Lapkin 1992). Modeled on L2 classroom research, many studies have sought evidence for learners' noticing by gathering linguistics data showing negotiation of meaning (e.g., Blake 2000). Research seeks evidence for authenticity through the use of a variety of discourse analysis

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

and introspective methods (Chapelle & Lui 2007). Research examining impact is conducted through the use of questionnaires and discourse analysis (e.g., Belz 2003; Jamieson, Chapelle & Preiss 2005). Some research on impact has investigated how the cultural ideologies and behaviors intersect among participants and in some cases fail due to cultural aspects of the partnerships (Belz 2001). Research on the language and the evidence it provides for the development of intercultural competence offers an example of the way in which research on technology tasks draws upon research methods in applied linguistics to provide evidence of the quality of particular tasks.

These methods and combinations of them in mixed-methods studies reflect approaches for gathering empirical evidence in applied linguistics research (e.g., Dornyei 2007). Such methods have been used along with the judgmental approaches that characterized much practice in other materials evaluation. Empirical methods provide a basis for supporting or refuting the types of claims that are made about language learning activities, and therefore one can imagine a day when outrageous claims such as 'Guaranteed to be the best way to learn a language' will be met by critical questions such as 'What evidence do you have to suggest that?'.. While such empirical methods are useful for investigating the claims that one might make explicitly about learning materials, a more critically-oriented analysis is needed to investigate other aspects of language learning materials and activities.

3.3. Critical perspectives

A third area that research on CALL is helping to point out is the relevance of critical perspectives on language learning materials. Critical perspectives have been developed largely due to the recognition that 'pedagogical choices about curriculum development, content, materials, classroom processes, and language use, although appearing to be informed by apolitical considerations, are in fact, ideological in nature, with significant implications' (Auerbach 1995: 9). Implications can fall beyond what materials developers have in mind as they design pedagogical tasks and select topical content for language learning materials. The expanded set of options for materials developers afforded by technology reopens questions such as how best to select the topical content of materials and how to prepare learners for intercultural exchanges using the Internet.

In the United States, national guidelines for foreign language teaching indicate that topical content should be chosen in a way that allows students the opportunity to learn about the cultural contexts where the language is used. Internet-mediated intercultural foreign language education helps to meet this goal by helping learners practice their language with native speakers their own age, from whom they might ideally gain more than language practice. Belz and Thorne (2006: x) define 'the hallmark of Internet-mediated intercultural foreign language education' as 'the inclusion of living, breathing human representatives of the languages and cultures under study in classroom-based FL [foreign language] instruction . . . the Internet is the mediator of this process'.

A critical look at the cultural content of learning materials raises important issues, including how choices about the particular cultural contexts are made. Schulz (2006: 13) points out, 'In the United States, as yet there are no commonly agreed upon minimal cultural contents that

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

textbook authors are expected to include in materials . . . For instance, there is no agreement among teachers of German as to which representatives of German culture their students should know about [and] what events related to German history they should be familiar with'. In view of the importance of the materials in educating students about who speaks the language, the selection of contexts and people depicted in the language materials is important. Writing about the need for critical cultural awareness in the development of language materials, Risager (2007: 236–237) argues that selection of content in language materials 'demands that education authorities, language teachers, and authors of teaching materials seriously consider their image of the world and remain open to the diversity of contexts that might be relevant and interesting for learners'.

In French teaching materials in the United States, for example, one might expect authors to recognize the relevance of French-speaking Canada and provide ample materials from contexts in Quebec and French-speaking Canada. A study of nine French textbooks used in large public universities in the United States found some mention of Canada in each, but the amount of coverage of Canadian topics and context was not substantial (Chapelle 2009). Moreover, in view of the interesting lessons that US students might learn from Canada about language, the choices of topics missed some significant opportunities. Topics with the potential for expanding students' understanding of language include the value of bilingualism, the political nature of language, the linguistic borrowing and code switching that occurs in multilingual contexts, and the role of language in identity. These are important lessons for students in the United States, where the dominant ideology treats monolingualism as the norm and language as apolitical (Robinson, Rivers & Brecht 2006). US students might be better prepared to communicate with and learn from students outside the United States having had a glimpse at alternative language ideologies and practices.

4. Conclusion

The three lessons prompted by work in CALL suggest directions for expanding perspectives from applied linguistics to the practice of materials development and evaluation. These are challenges for us in applied linguistics. We need to take the lead in moving materials evaluation into a more research-oriented framework, where legitimate claims are made about materials on the basis of evidence from research rather than solely from the creativity of marketing departments. Until applied linguists are prepared to offer concrete suggestions about feasible research that can be used in materials evaluation, we may need to be circumspect in criticizing publishers. In applied linguistics, one can find ample mention of the villainous publishers whose profit motives preclude responsible development and evaluation of materials. My experience in working with publishers on materials development projects presents a different picture.

I meet extremely dedicated and creative former ESL/EFL teachers who attempt to develop materials that will be engaging and educational for the many students who will use them all over the world. Despite the good intentions and hard work of these materials developers, the fact is that they learned very little of the relevant professional knowledge in their MA programs in TESOL and applied linguistics, that is knowledge about language, L2

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

acquisition, assessment, CALL, and research methods. The picture I see presents a challenge to applied linguists to increase the sophistication of the knowledge that is transmitted in graduate programs in TESOL and applied linguistics, and in doing so to strengthen the connection between professional knowledge and materials development and evaluation. Any improvement in the practice will have as its basis a strong coherent basis in our profession. In my view, the spread of CALL contributes to such a basis by providing both the opportunity to learn about materials evaluation and the necessity to break down the distinction between CALL and other language learning materials.

References

Allum, P. (2002). CALL in the classroom: The case for comparative research. *ReCALL* 14.1, 146–166.

Auerbach, E. R. (1995). The politics of the ESL classroom. In J. W. Tollefson (ed.), *Power and inequality in language education*. Cambridge: Cambridge University Press, 9–33.

Barnard, R. & M. Randall (1995). Evaluating course materials: A contrastive study in text book trialling. *System* 23.3, 337–346.

Belz, J. A. (2001). Institutional and individual dimensions of transatlantic group work in network-based language teaching. *ReCALL* 13.2, 213–231.

Belz, J. A. (2003). Linguistic perspectives on the development of intercultural competence in telecollaboration. *Language Learning & Technology* 7.2, 68–117.

Belz, J. A. & S. L. Thorne (2006). Introduction: Internet-mediated intercultural foreign language education and the intercultural speaker. In J. A. Belz & S. Thorne (eds.), *Internet-mediated intercultural foreign language education*. Boston, MA: Thomson Heinle, iix–xxv.

Blake, R. (2000). Computer-mediated communication: A window on L2 Spanish interlanguage. *Language Learning & Technology* 4.1, 120–136.

Chambers, F. (1997). Seeking consensus in coursebook evaluation. *ELT Journal* 5.1, 29–35.

Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology* 2.1, 22–34.

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge: Cambridge University Press.

Chapelle, C. A. (2007). Challenges in evaluation of innovation: Observations from technology research. *Innovation in Language Learning and Teaching* 1.1, 30–45.

Chapelle, C. A. (2009). A hidden curriculum in language textbooks: Are beginning learners of French at U.S. universities taught about Canada? *The Modern Language Journal* 93.2, 139–152.

Chapelle, C. A. & H-M. Lui (2007). Theory and research: Investigation of 'authentic' CALL tasks. In J. Egbert & E. Hanson-Smith (eds.), *CALL environments* (2nd edn.). Alexandria, VA: TESOL Publications, 111–130.

Chenoweth, N. A. & K. Murday (2003). Measuring student learning in an online French course. *CALICO Journal* 20.2, 285–314.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Doughty, C. J. & M. H. Long (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology* 7.3, 50–80.

Ellis, R. (1998). The evaluation of communicative tasks. In Tomlinson (ed.), 217–328.

Grgurovic, M. (2007). Database of comparison studies: Computer-assisted vs. classroom second/foreign language instruction. http://tesl.engl.iastate.edu:591/comparison/main.htm.

Grgurovic, M., C. A. Chapelle & M. Shelley (in preparation). Effectiveness of CALL: A meta-analysis and research synthesis.

Hubbard, P. (2006). Evaluating CALL software. In L. Ducate & N. Arnold (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos, TX: CALICO, 313–334.

Jamieson, J., C. A. Chapelle & S. Preiss (2005). CALL evaluation by developers, a teacher, and students. *CALICO Journal* 23.1, 93–138.

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press

Levy, M. (1997). *Computer-assisted language learning: Context and conceptualization*. Oxford: Oxford University Press.
Littlejohn, A. (1998). The analysis of language teaching materials: Inside the Trojan Horse. In Tomlinson (ed.), 190–216.
Masuhara, H. & B. Tomlinson (2008). Materials for general English. In Tomlinson (ed.), 17–37.
McDonough, J. & C. Shaw (2003). *Materials and methods in ELT: A teacher's guide* (2nd edn.). Malden, MA: Blackwell.
McGrath, I. (2002). *Materials evaluation and design for language teaching*. Edinburgh: Edinburgh University Press.
Pederson, K. M. (1987). Research on CALL. In W. F. Smith (ed.), *Modern media in foreign language education: Theory and implementation*. Lincolnwood, IL: National Textbook Company, 99–132.
Plass, J. L., D. M. Chun, R. E. Mayer & D. Leutner (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology* 90.1, 25–36.
Risager, K. (2007). *Language and culture pedagogy*. Clevedon: Multilingual Matters.
Robinson, J. P., W. P. Rivers & R. D. Brecht (2006). Speaking foreign languages in the United States: Correlates, trends, and possible consequences. *The Modern Language Journal* 90.4, 457–472.
Sanaoui, R. & L. Lapkin (1992). A case study of an FEL senior secondary course integrating computer networking. *Canadian Modern Language Review* 48.3, 525–552.
Schulz, R. A. (2006). The challenge of assessing cultural understanding in the context of foreign language instruction. *Foreign Language Annals* 40.1, 9–26.
Scida, E. E. & R. E. Saury (2006). Hybrid courses and their impact on student and classroom performance: A case study at the University of Virginia. *CALICO Journal* 23.3, 517–532.
Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal* 42.4, 237–246.
Tomlinson, B. (ed.) (1998). *Materials development in language teaching*. Cambridge: Cambridge University Press.
Tomlinson, B. (2003a). Materials evaluation. In Tomlinson (ed.) (2003b), 16–36.
Tomlinson, B. (ed.) (2003b). *Developing materials for language teaching*. London: Continuum.
Tomlinson, B. (ed.) (2008). *English learning materials: A critical review*. London: Continuum.
Tomlinson, B., B. Dat, H. Masuhara & R. Rubdy (2001). ELT courses for adults. *ELT Journal* 55.1, 80–101.

CAROL A. CHAPELLE, Professor of TESL/applied linguistics, is co-editor of the Cambridge Applied Linguistics Series. Her research explores issues at the intersection of computer technology and applied linguistics. Her books include *Computer applications in second language acquisition: Foundations for teaching, testing, and research* (Cambridge University Press, 2001) and *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology* (John Benjamins, 2003). Co-edited books focus on language assessment and research methods: *Assessing language through technology* (Cambridge University Press, 2006), *Inference and generalizability in applied linguistics* (John Benjamins, 2006), *ESOL tests and testing: A resource for teachers and administrators* (TESOL Publications, 2005). Her most recent books are *Building a validity argument for the Test of English as a Foreign Language* (Routledge, 2007) and *Tips for teaching with CALL* (Pearson-Longman, 2008).

https://doi.org/10.1017/S0261444809005850 Published online by Cambridge University Press