Diversity in modalities

Marie-Noëlle Lamy

Introduction

Why include a chapter on modality in this book? Computer-aided language learning (CALL) practitioners do not generally use the word “modality” when instructing their students (for instance they don’t say “make sure you carry out this interview using the audio modality,” or “share your thoughts on your partner’s ideas using the commenting modality”). On the other hand, researchers do use this word, all too often failing to define what they mean by it, or by the word “mode.” The idea of “diversity” in such a context comes close to taking on overtones of “mix” or even “mess.” It is my intention, in subjecting the notion of “modality” to a critical examination, to show how lack of clarity about this term is damaging published research in CALL, and to suggest how we can think about “modality” so as to facilitate our understanding of the educational benefits to be expected from CALL, whatever it is that “technology” throws at us next, in its increasing diversity.

In the literature of technology-mediated language learning, the term “modality” is rarely defined (although for a brief overview see Chanier and Vetter, 2006) and it is used inconsistently. For example the following list contains items that have variously been offered as illustrations of different modalities: “audio,” “text,” “icons,” “webcam,” “image,” “written,” “oral,” “video,” “voice-only,” “voice-image,” and “textchat.” However, one might equally say that a “video” is a material object or an electronic object; an “icon” is an artifact as well as a sign; a “webcam” is a tool; and “written” is one of the two forms that human language takes. The polysemy and semantic heterogeneity of this list make the word modality unusable as an operational concept.

In the field of multimodality research, a neighboring discipline to CALL, a consistent definition is available. Multimodality research seeks to understand how we make meaning, through the diversity of communicative forms – language, image, music, sound, gesture, touch, and smell – that surround us. With the exception of smell, these can all be found within the experience of learning and teaching online, which is why I propose to draw some definitions from the work of multimodality research’s main theorists, Kress and van Leeuwen, who
developed a theory for understanding the meanings communicated to us by objects such as adverts and posters, which use language as one – but not as the main – resource for meaning-making.

Let us start with the notion of “mode.” “Words processors,” these authors tell us, “must systematize such things as the thickness and positioning of the lines that separate sections of text, and develop a metalanguage, whether visual or verbal, for making these choices explicit” (2001, p. 79). Anyone who has had the opportunity to compare a page from a magazine in the English-speaking press with one from the French press knows that choice of typeface and positioning of headers, annotations, captions and so forth are systematized differently in these cultures. The culture-dependent systematization of graphic resources amounts to a specific “grammar” of visual communication, which Kress and van Leeuwen call a “mode.” The physical tool (a computer) changes the physical media (paper and ink) into a mode (a culturally intelligible page layout). Other, much more immediately obvious, culturally intelligible systems include language (written, spoken), the visual (figurative and non-figurative or coded, such as icons), sound (figurative and non-figurative such as music, or coded such as signals), and body-language. Modes are basic to meaning-making (hence the adjective “semiotic” is frequently used in the rest of the chapter) and they work together in complex ways to bring about our understanding of how objects communicate to us certain ideas or emotions such as “adventure,” “authority,” “humor,” “hope,” and hundreds of others. For example, a piece of architecture, an art installation, or an immersive game screen are made up of physical matter fashioned into cultural styles which we recognize as conveyors of the different ideas or emotions in question. “Modality,” therefore, is the relationship between modes and the culturally intelligible object that they underpin, and “multimodality” is:

the use of several semiotic modes in the design of a semiotic product or event, together with the particular ways in which these modes are combined – they may for instance reinforce each other (“say the same thing in different ways”), fulfill complementary roles … or be hierarchically ordered, as in action films, where action is dominant, with music adding a touch of emotive “color” and synch sound adding a touch of realistic “presence.” (Kress and van Leeuwen, 2001, p. 20)

For the purpose of our discussion of CALL, we will in a moment add the notion of “tool” to this definition, adding to the number of elements that can be combined into a particular “modality.” For these reasons it has seemed difficult to construct this chapter around a “list” of modalities available in CALL. Rather, the approach will be a discovery-based one, using a corpus of published CALL studies through which I hope to outline the challenges associated with the concept of modality in our field, the risks associated with misunderstanding modality, and an alternative way of addressing these problems.
Overview and general issues

In order to learn a language online and to communicate (including online) in that language, you need to orchestrate a variety of resources. These include natural language in its written and spoken forms and they always also include visual resources. Understanding even a simple written text requires sensitivity to two modes: a linguistic mode (written language) and the visual mode (the choice of fonts and the organization of spaces on the screen). In CALL, many resources are used, such as icons, images, colors, and shapes, in combination with each other. For example a tool such as a floating tag – the small window that opens as the cursor is floated over specific parts of the screen – can support the written language. But it can also support pictures (as with some online bookshops, where the floating tag reveals an enlarged picture of the book cover) or even a search form (which is a further tool supporting written language). The floating tag tool in itself has no “meaning” of interest to a language learner or teacher. On the other hand the written text, picture, or search form may convey many meanings, some of which are relevant to the language learning situation. For example the text facility accessible via a floating tag may be used to support a vocabulary gloss, an image may be used to illustrate a cultural reference, and a search box facility may be exploited as part of a grammatical, lexical, or other type of exercise. On “multimodal” platforms the physical material includes keys to be pressed for text creation, and pads or mouse attachments to be tapped, stroked, or clicked for opening up the audio channel in order to hear or speak. In other words, the learning objective will be jeopardized if the learning design has not fully taken account of the materiality of the tool and of the specific mode or modes involved.

To sum up, then, in the context of CALL semiotic resources are made up of (a) material tools; (b) modes, such as written language, spoken language, or visual language; and (c) language learning objectives materialized through educational designs. The conjunction of these three elements makes up modality, as exemplified in Table 7.1.

Not all technologies that have proved of interest to language learners and teachers are listed in the table. For example neither blogs, multi-user domain object oriented (MOOs), virtual worlds, nor social networks are featured (see Chapter 5 for a discussion of these technologies). This is because in the terms of this chapter, they are not tools in themselves but rather they are made up of combinations of tools and modes, much like the piece of architecture or art installation mentioned in the Introduction.

The central issue when discussing modes and modality

If there was a one-to-one correspondence between physical tools and the other resources that make up modality, e.g. if synchronous messaging tools were
Diversity in modalities

Table 7.1 *Examples of modality relationships in CALL*

<table>
<thead>
<tr>
<th>Tools</th>
<th>Modes</th>
<th>Language learning objectives facilitated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browsers</td>
<td>Multiple modes</td>
<td>Access to authentic material in L2</td>
</tr>
<tr>
<td>Word processors</td>
<td>Written text mode</td>
<td>Productive text-based activities</td>
</tr>
<tr>
<td>Photo editing software</td>
<td>Visual modes</td>
<td>Productive image-based activities</td>
</tr>
<tr>
<td>3D artifact manipulation</td>
<td>Visual modes</td>
<td>Cultural or geographical learning outcomes</td>
</tr>
<tr>
<td>GPS tools (mobile devices)</td>
<td>Visual and iconic modes</td>
<td>Geographical or other location-dependent learning outcomes</td>
</tr>
<tr>
<td>Automatic recording and screen capture tools</td>
<td>Multiple modes</td>
<td>Reflective activities (revisiting tutorial sessions)</td>
</tr>
<tr>
<td>Hot buttons</td>
<td>Iconic modes</td>
<td>Communication and interaction (e.g. vote button); create tele-presence; help with turn-taking (e.g. raised-hand button)</td>
</tr>
<tr>
<td>Asynchronous sending/receiving (fora)</td>
<td>Mainly written text mode, may have some element of visual modes</td>
<td>Written work. Peer collaboration; feedback; meta-commenting (extended, reflective commentary)</td>
</tr>
<tr>
<td>Synchronous messaging tools (chat)</td>
<td>Mainly written text mode, some element of visual modes</td>
<td>L2 messaging styles. Peer collaboration; feedback. Also group bonding.</td>
</tr>
<tr>
<td>Voice-over Internet (telephony, audio)</td>
<td>Mainly spoken language mode, with textual, visual and iconic modes</td>
<td>Oral communication and interaction; collaborative work; feedback</td>
</tr>
<tr>
<td>Webcam-based video-conferencing</td>
<td>Spoken and visual modes, with iconic mode.</td>
<td>As above</td>
</tr>
</tbody>
</table>

only ever used with the written language mode and the only learning objective was to practice informal writing, this chapter would not be necessary; we could then happily talk about “the chat modality.” However, it became clear above that a single tool may be associated with more than one mode, and more than one educational meaning. Even a synchronous messaging facility can support several different modes (for instance meaning can be made within a chat window using smileys, which belong to the iconic mode, or with thumbnail photographs, which partake of the visual modes).

Take the example of how dispersed distance learners used a chat tool in a multimodal platform project (reported in Lamy, 2006) where the task instructions required students to:
give priority to the audio channel and communicate orally;
• use the chat facility in case of sound breakdown, writing in their mother
tongue in order to avoid compounding difficulty.

An analysis of usage showed that students did prioritize the audio channel but
that they ignored the instructions constructively, appropriating the chat tool in
a range of ways, forming “parallel talk” and “cross-talk” conversations in the
target language instead. Conversation patterns included the following:

• parallel talk (i.e. interaction between the oral and the written language) went
on in audio and in chat on topics that were:
  • unrelated (e.g. students communicated orally about the task while text-
chatting about the possible whereabouts of a latecomer);
  • identical (e.g. students communicated orally about the weather in one
of their locations while text-chatting about the weather in another
location);
• crosstalk (i.e. interaction between the written and spoken modes) used topics
that were:
  • divergent (e.g. the oral talk was about the task but the text-chat box dis-
played a call for help from a participant with a sound problem, immedi-
ately triggering an oral debate about how best to help him out);
  • convergent (e.g. the spoken conversation was about the Alliance Française
while text-chatters were writing about the British Council, until the group
started comparing both organizations orally in the audio channel.

In these parallel and cross talks, the meaning to which the chat function had
been assigned by the task designers rarely came to the fore. Rather, socio-
affective discourse and other types of discourse arose, as summarized in
Table 7.2.

The issue for our understanding of modality is as follows: we cannot say
that we are dealing with one modality (“the chat modality”) as this would be
a reduction of the diversity that is going on in this chat-based class. The chat
software tool alone supports three different modes and a variety of learning
interactions, both in terms of content (discourse type) and group dynam-
ics (parallel talk, crosstalk). Increasingly learners and teachers are working
with more than one tool at a time, as do the above with a chat window and
an audio tool, which makes it even more difficult to accurately describe the
process.

Example

Comparative methods have proved useful in semiotic analysis (Neergaard and
Ulhøi, 2007) for identifying the features of an object, by contrasting them with
Table 7.2 *One tool, three modes and many meanings in text-chat*

<table>
<thead>
<tr>
<th>Discourse type (most frequent first)</th>
<th>Mode</th>
<th>Content of parallel talk and crosstalk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socio-affective</td>
<td>Linguistic (free, written), iconic (coded: smileys)</td>
<td>Greetings, jokes, compliments, apologies</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Linguistic (free, written), linguistic (coded: URLs)</td>
<td>Suggesting ideas, websites, documentation</td>
</tr>
<tr>
<td>Methodological</td>
<td>Linguistic (free, written)</td>
<td>Requesting clarification of instructions, or assignments</td>
</tr>
<tr>
<td>Linguistic</td>
<td>Linguistic (free, written)</td>
<td>Checking or suggesting vocabulary, requesting corrective feedback</td>
</tr>
<tr>
<td>Technological</td>
<td>Linguistic (free, written)</td>
<td>Requesting or offering help</td>
</tr>
</tbody>
</table>

one that is broadly but not entirely similar. Thus Lamy and Hampel (2007) created four meta-studies, each based on comparing two projects that had a tool in common, e.g. pairs of studies that used forums, chats, MOOs, and multimodal platforms. The authors hypothesized that a comparative approach would make it easier to derive a rich understanding of the modality relationships at work through this literature than by examining cases singly. In the following section, I sum up their main findings.

**Study pair 1**

The first pair of studies in the survey (Savignon and Roithmeier, 2004; Weasenforth *et al.*, 2002) used forums. The starting point for Savignon and Roithmeier’s research was learner uptake of form, positioning the study on the cognitive end of online language learning research (the opposite end of the spectrum being more socio-constructivist). Savignon and Roithmeier were looking for evidence that “the collected bulletin board postings on a single subject qualify as a cohesive, coherent text” (2004, p. 269). Their second research question focused on discourse as well: “What discourse features can be identified that reflect participant engagement in terms of sustaining a collaborative dialogue?” (2004, p. 269). The setting for the project was a class of German students of English at a German secondary school, eight years into their study of their second language (L2). Working with them was a class of US students of German enrolled in their third year at a high school. They discussed societal topics (the American Dream, the death penalty, drinking and driving, the Kosovo conflict) over a three-week period. All three research questions received a positive answer. Uptake and peer-sharing of lexical items
were shown to have occurred “implicitly or explicitly” (2004, p. 272). The issue of the sustaining of a collaborative dialogue received an equally positive answer, based on observations about both the use of strategies to mitigate potential conflict and the co-constructing of a coherent and cohesive text: “[t]hrough the incorporation of previously used lexical items, ideas, and even entire postings, participants show they were following the discussion” (2004, p. 284).

The second forum study, by Weasenforth et al. (2002), aims to determine under what conditions teachers may be “[r]ealizing constructivist objectives through collaborative technologies.” The authors use constructivist principles as a framework, which positions them on the other end of the continuum of pedagogical philosophies compared to Savignon and Roithmeier. They evaluated a three-semester process involving fifty-two advanced-level university reading/writing students of English as a second language (ESL), all international graduates. Course requirements were explicitly oriented to participation patterns (e.g. “[D]uring semester 1, students were required to introduce a new thread each week and to participate in a total of twelve discussions about course content” (2002, p. 61). The asynchronous nature of the threaded discussions, according to Weasenforth et al., “made [the course assignment] particularly useful for the promotion of coherent discussion. The additional time available for reading and composing postings encourages reviewing and responding to classmates’ arguments” (2002, p. 74). The asynchronous nature of the tool was also credited with enhancing the reflective learning style of quiet students, and with introducing flexibility through the possibility of extending assignment preparation time whenever necessary.

Lamy and Hampel’s comparison of these two studies shows that the combination of forum as a tool and the written language as a mode afford flexibility of the type that language educators need, facilitating the success of such divergent learning designs. The question then is whether all possible learning designs could be facilitated by forum-based work. If the answer was yes, we could say that we had identified a stable modality.

Study pair 2

Lamy and Hampel also looked at synchronous tools as used in a study by Blake (2000) and compared it with Thorne (2003). Blake, who has written extensively about task types offline, set up networked learner–learner discussions via a synchronous chat program associated with other tools. Fifty intermediate learners of Spanish took part in the study, carrying out a series of online task types: jigsaw, information-gap, and decision-making. Blake’s main finding was that well-designed networked tasks “promote learners to notice the gaps in
their lexical interlanguage in a manner similar to what has been reported in the literature for oral learner/learner discussion” (2000, p. 132). The only benefit of the online situation compared to the face-to-face one, for this author, is the remote linking of learners, i.e. a “communications” advantage rather than a learning one.

The other paper in this pair, Thorne (2003), is based on telecollaborative exchanges between US and French students via chat facilities. Thorne presents a case study of students of French at Penn State University and engineering students at the Ecole Nationale Supérieure de Télécommunication de Bretagne. The study focused on US students’ perception of different tools, and in particular their preference for synchronous communication via instant messenger (IM) rather than email. It showed that although the exchange was supposed to facilitate intercultural communication between these students, the way it was set up did not take account of the “cultures-of-use” of the communicative media. That is, it created a mismatch between how students use these tools in everyday life and how they were supposed to use them in the educational context. In order to talk to their peers outside class, the students tended to use IM rather than email. For them, email was a communication tool to be used with parents and teachers, that is, between power levels and generations, and therefore inappropriate for an exchange with peers in class and for building a relationship with them.

The turning point of Thorne’s article, initially an exploration of sociocultural conditions, is the active resistance to email in favor of IM by the US (and some of the French) participants. By re-orienting his research towards a wider socio-historical focus (“cultures-of-use” and their influence on learning), Thorne is able to return his reader to a level of reflection that goes beyond a limited interrogation of the possibilities of the tool. Modality, as a relationship, is seen here to include not only the immediate educational context but the socio-historical context within which those in the educational institutions operate. The contrast between the two approaches to modality in study pair 2 is stark: while Thorne contributes a refinement to our understanding of how modality works, on the other hand Blake’s results appear to be offered regardless of whether there is an electronic tool involved or none, telling us a lot about task design but nothing about modality in CALL.

Study pair 3

MOOs have characteristics in common with synchronous text-based systems such as chat, which they combine with asynchronous written systems such as fora, but also have distinctive features. Peterson (2004) explains that “these environments are designed around a hierarchy of user privileges that enables the creators of a MOO to structure the environment to meet the
needs of a particular learner group … Typically MOOs contain numerous virtual rooms, linked together by entrances and exits … Some MOO environments contain numerous learning objects including virtual projectors, lecture spaces, notes, web pages and recording devices” (Peterson, 2004, pp. 40–44).

Kötter (2003) reports a 1998 project with fourteen German-speaking learners of advanced English at the University of Münster, and fifteen US students of intermediate German at Vassar College, who met for twice-weekly interaction, for seventy-five minutes at a time, in a MOO where they collaborated in a total of eight sub-groups of three or four students, to complete projects of their choice. For example, they created self-profiles and discussed the cultural implications of this experience with their partners. All were asked to present their work to the other groups during the final sessions of the exchange. Kötter researched four questions, one of which directly addresses the specificity of the MOO. (“How do students who meet in a MOO rather than in person deal with the apparent “virtuality” of their encounters, that is, which “MOO-specific” tools and strategies do they employ to express themselves and exchange information?”) Kötter observed that many participants:

had begun to make themselves at home in the MOO by creating their own rooms even before they met with their tandem partners for the first time. Some learners had fitted these rooms with objects ranging from a sofa or a carpet to a piano or a refrigerator, and several students had additionally composed elaborate descriptions of these purely text-based locales … Equally important, many of the remarks that the students made to each other documented that they conceptualized the MOO as something with a spatial dimension. One learner commented upon arrival in her partner’s room that it “looks okay here,” while another stated that she preferred her peer’s room to the MOO’s entrance … Moreover, many learners exploited the notion of space in the MOO by engaging with things they found in these rooms (e.g. Jack settles down in a comfortable chair) to create a pleasant atmosphere for their encounters even if they had no previous experience with MOOs. (Kötter, 2003, p. 152)

However these socio-affective exchanges remained confined to a pre-sessional “settling in” phase and did not sustain as part of the core learning opportunities available when the course itself unfolded. Kötter thus had no opportunity to develop his thinking on these specific MOO modality issues.

The second MOO-based study is by Schneider and von der Emde (2006). Eleven German-speaking learners of English from the University of Münster and fourteen US students of German from Vassar College met twice a week for sixty minutes at a time, for intercultural work in small groups. The cultural content was derived from documents and synchronous discussions about two comparably emotional events: the shootings at a school in Columbine, Colorado in 1999 and at a school in Erfurt, Germany in 2002. Carried out by means of student-created “project rooms”, the “project work” phase can claim
to specifically exploit the functionalities of the MOO tool. In this phase, students built open-ended and interactive rooms in which to raise questions rather than give answers. These rooms “depict all the conflicting perspectives on the various topics in order to encourage the visitors to the rooms to draw their own conclusions” (2006, p. 191). For instance, one group “explicitly embedded the concept of conflict into the structure of their project by calling one of their rooms Missverständnisse ‘Misunderstandings’,” (2006, p. 192) and posting the following explanation: “This room has information about our discussions but it does not have any answers. Why? Because there aren’t any concrete answers” (2006, p. 192). Schneider and von der Emde commented on this use of the environment, but their focus on cultural learning (rather than on educational technology) prevented them asking questions that might bring a better understanding of the modality of MOO-based language work, such as “how did students use the meaning-making resources specific to MOOs; for example did the creation of the room called Missverständnisse support the learning project that the students had in mind when they created it (e.g. facilitating reflection and demonstrating a critical engagement with the cultural agenda of the course)? Schneider and von der Emde’s study touches on these questions, yet it falls short of problematizing the relationship between discourse objectives (critical engagement) and the MOO tool’s user-extendable spaces and graphics.

For Lamy and Hampel, what emerged from these two studies of MOOs was Kötter’s sole concentration on chat within MOO, and Schneider and von der Emde’s strong focus on pedagogy, rather than on the mediation of that pedagogy via the MOO tool. They concluded that insufficient notice has thus far been taken of the work of those who, as early as 1999 (Shield et al.) and 2001 (Peterson) were trying to identify the specific modalities that could be associated with the MOO tool. If over the years these have not been perceived as anything other than identical with those of chat tools, it is perhaps no surprise that MOOs have not been widely used in language learning and teaching online.

**Study pair 4**

Erben (1999) was a pioneer project using a synchronous multimodal platform. Lamy and Hampel contrasted this work with Svensson’s multimodal project in 2007. Erben’s work is a case study in an immersion context, involving a B.Ed. program at Central Queensland University, in which up to 80 percent of the curriculum is delivered through the medium of Japanese. Erben produces an understanding of participants’ behavior, which he expresses as “reduction” and “amplification” of the range of symbolic cues available to them. Amplification and reduction refer to “those classroom discursive practices which, because of the nature of the mediated interaction at a distance, participants need to modify … in order to achieve the same effect as if the equivalent cue, sign or behavior
was produced in a face-to-face classroom” (1999, p. 237). An example of amplification might be an instance of teachers having “to increase question wait time due to the fact that delayed transmissions from site to site may occur” (1999, p. 238), while reduction might be “the loss of learning opportunities through such technical hiccups, where the connection between sites may freeze” (1999, p. 239). Initially, Erben found that in the online setting teacher-led activities were amplified; after a while, however, teacher control was reduced. Private classroom communication was generally amplified.

Modifications of discursive practices occurred, as existing ones did not suffice to guarantee clarity of meaning in the new environment. Erben called this phenomenon “reconstruction” (1999, p. 240). An instance of a reconstructed sociocultural practice is that of “bowing in the classroom,” a necessity in Japanese culture, leading “to the use of different verbal cues or picture icons” to replace the bodily bowing in the online setting. The author notes that “as classroom participants adapted to the use of the platform, instructional processes came to be increasingly reconstructed in ways which represented a substantive shift away from how these processes occurred in face-to-face immersion classrooms” (1999, p. 240).

The final study in Lamy and Hampel’s survey is Svensson’s (2003) project, located at Sweden’s Umeå University, at Humlab, a then “state-of-the-art” technology laboratory for humanities students and researchers. The project brought together different language subdisciplines. It involved twenty-two advanced postgraduate students of English, the majority of whom did not have advanced computer experience, taking part over three years. The work involved producing a piece for assessment, equivalent in workload to the twenty-five-page essay required of offline students. However, going far beyond the communication modes used in the essay genre, this assignment involved linguistic modes (“free” as well as “coded” – such as displayed text) as one part of the whole rather than as the principal mean of communication, combining language with images, movement, and sound. The assignment was therefore truly multimodal.

The environment used (under the Active Worlds program) combined sound, text, and graphics. In order to participate, users choose to be represented by an avatar – a person, an animal, or an object. Students worked in sub-groups to build a graphical world around themes (e.g. “weddings,” “the city,” “monstrosity”), to link objects in their “world” to web pages, which they created, and to represent, often in non-linguistic form, the concepts relevant to their theme.

The project is deliberately process- rather than product-oriented, and Svensson shows how students collaborate and negotiate in building their “world.” He stresses the importance of space: “visual and auditory means of expressions in distributed spaces have a very strong motivating and creative effect” (2003, p. 139, emphasis added). According to Svensson, such project
work is particularly useful for language learning, which “is about language, immersion in other cultures, communication, media, intercultural meetings and role-play, and virtual areas supply us with a place where all these can come together” (2003, p. 140).

Lamy and Hampel’s conclusion is that Erben has helped define some of the new learning practices of the online world (reductions, amplifications, and reconstructions) by mapping them to the known reality of the physical classroom, thus creating a generalizable model, i.e. one that could be interpreted by later researchers as a position from which alternative models could be constructed. In contrast, Svensson has offered a rich description of practice. Each approach has its merits. With Svensson’s we gain a “feel” for what the students working in that modality may have experienced. With Erben, our understanding of modality is more theoretical.

What have we learnt about modality through the discovery process of contrasting studies? Table 7.3 is an attempt to redevelop and revisit the scheme set out in Table 7.1 to highlight reflections about modality.

Implications

We have seen that in technology-mediated language learning and teaching, modality is at best a misunderstood and at worst an ignored concept. The survey in the preceding section showed that research that ignores it is in danger of missing out on ways of explaining the nuances in the learning process (study pairs 1 and 2), or of failing to exploit learning possibilities to the full (study pair 3). Researchers who integrate modality into their research (even if implicitly, as is the case in the studies surveyed) have been able to improve our understanding of the ecology of learning (e.g. Thorne) and to provide the sort of generalizable framework that is important for ensuring the relevance of local findings to the CALL community at large (e.g. Erben). The comparative survey has also shown how modality is affected by the materiality of the online environments and by the role that “space” (the physical spatial experiences of the user and the represented spaces on the screen) plays in them. In this final section of the chapter, we suggest a framework that may help to make these aspects of modality more visible to research.

Materiality

The materiality of the environment impacts on the dynamics of conversations. In two parallel conversations in chat, for example, it is not accidental that a tutor-initiated discussion is carried out orally while the students’ exchange about sound levels is conducted in text-chat: if we mentally invert the modal choices and imagine that the tutor led his tutorial via postings in the text-chat
Table 7.3 Modality: issues emerging from our corpus

<table>
<thead>
<tr>
<th>Pairs of studies/tools</th>
<th>Modes</th>
<th>Language objectives</th>
<th>Reflections on modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 Fora</td>
<td>Linguistic -written</td>
<td>Acquisition, collaboration</td>
<td>The success of both pedagogical projects seems to have inhibited the researchers from further investigating the role of modality.</td>
</tr>
<tr>
<td>Pair 2 Chats</td>
<td>Linguistic – written</td>
<td>Better task design, Cultural</td>
<td>Blake: no differences between F2F and online modalities are discussed. Thorne: the “mixed” outcomes of the learning project seems to have encouraged the researcher to extend his reflection on modality.</td>
</tr>
<tr>
<td>Pair 3 MOOs</td>
<td>Linguistic -written, iconic</td>
<td>Cultural</td>
<td>The researchers seem to have experienced difficulty in reflecting on the specific modalities at work in the MOO project, remaining within a discussion of chat modalities instead.</td>
</tr>
<tr>
<td>Pair 4 Multimodal platforms</td>
<td>Linguistic- written and spoken, visual, iconic</td>
<td>Professional development, Cultural, self-expression</td>
<td>Erben: the researcher engaged with modality issues and set out a generalizable framework to deal with them. Svensson: the researcher provided a rich description of the modality at work in the project (rather than a framework).</td>
</tr>
</tbody>
</table>

window while students talked about other topics orally, we see that such an arrangement is unlikely. In the next section I explain why.

These authors’ accounts of meaning-making in multimodal environments went further than the opening up of non-linguistic modes to analysis that we mentioned earlier. For them discourse is only one of four dimensions of modality, which also includes the design, production, and dissemination of any given multimodal artifact.

The impact of design on meaning-making

The design of an artifact embodies the ergonomic and esthetic reflections of its creator. In our context, the notion of “design” subsumes the choices that have determined what the users see on their screen, such as the architecture of the environment, the choice of colors, and the ergonomic harmony (or disharmony)
Diversity in modalities

of the frames, buttons, and icons as they appear on screen. For example, in some audio-synchronous platforms, a loudhailer icon signals the opening of the audio channel. In others, an icon representing a pair of fleshy lips carries out that function. The impact on meaning is not the same in both cases.

A fuller understanding of these meaning-making resources and their impact may be obtained by contrasting the experiences available to users on two platforms that have been used for language learning: Elluminate and Traveler. The metaphor underpinning Traveler is that of a fantasy world bringing together natural, supernatural, and built objects, peopled with avatars, reminiscent of video games. The visual design of Traveler is in complete contrast with the textual-visual design of an educational product like Elluminate, with its square workspaces, its stationery box full of pencils and erasers, and its hierarchy of “rooms” reminiscent of the serious educational purposes of a face-to-face campus. The ergonomics of these two environments are different too: while Traveler users work in three dimensions, using the keyboard’s arrows to walk or fly their avatar around, those of Elluminate use text or click icons to create new rooms. A final material difference adds to the contrast: while Traveler users may remain anonymous, Elluminate participants are always identified (via a nametag which appears automatically on log-on). Overall, then, Traveler gives learners opportunities for meaning-making through a playful design that connotes the “cultures-of-use” of the young, and Elluminate does so using the connotations of the academic life. Do these different approaches have a different impact? We have as yet no empirical learner data with which to answer that question, but we can gain a fairly good idea about how educators see the issue by noting that when the Mathematics and Computing Department of the Open University adopted Second Life (which has a ludic design similar to Traveler’s) to help its dispersed post-graduate students to collaborate, the decision was made to build an “island” dedicated to this group within Second Life and the design chosen for the island was, significantly, not the fantasy “look” that we mainly associate with Second Life, but one representing an idealized version of a university campus (http://virtualmphil.open.ac.uk; see Chapter 5 for further discussion of virtual environments such as Second Life).

The impact of production on meaning-making

In their notion of modality, Kress and van Leeuwen also included the production aspect of cultural artifacts. A simple example of the differences in meaning resulting from different conditions of production can be taken from the world of art, with such questions as “what different meanings inhere in an original painting as opposed to within its fake?” or “what different meanings inhere in a paper reproduction as opposed to within an electronic image?”
etc. In our context, an example is the extent to which the system has been produced to constrain the user to certain actions or how free the system allows the user to be. At one extreme the user cannot change anything and is dependent on the meaning-making resources produced by the system (a pre-interactive-TV television viewer would be a good example). At the other end of the continuum, the user has a lot of freedom to produce his/her own resources (MOOs, avatar-based worlds, graphic tools). Even in a “free” situation, the system provides many fixed features, so a “semi-constrained” position is what users mainly find themselves encountering, giving rise to creative usage. Here is an example of user creativity: in push-to-talk systems, users must press a button in order to ensure that their voice is heard by others, and one person only will be heard at a time. However, vocalizations from listeners (such as a happy “oh!,” an ironic “ha!,” or a sigh) are meaningful resources in conversations. Deprived of these paralinguistic possibilities in one push-to-talk system studied by Lamy (2006), users responded by switching modes, i.e. by co-creating a “library” of custom-build smileys which they input into the text-chat window.

The impact of dissemination on meaning-making

Finally, another dimension of multimodal semiotics for Kress and van Leeuwen relates to differences in meaning resulting from the way in which a cultural artifact is disseminated. Here is an example from a CALL project. In distance or blended educational projects, students participate from their homes. According to the specification and age of the user’s monitor, an image may be disseminated to them with higher or lower resolution, and the colors may be more or less true to the originals files input by the designers. In the following example, a role-play involved dispersed English learners of intermediate French looking at photographic stimuli on their screens. The group was required to discuss these images of six sea resorts and to decide on the basis of that discussion which resort would be most suitable for a fictitious family with specific needs. Students were to arrive at an agreed choice through oral consensus. One participant, misled (at it was discovered later) by the poor quality of the image on her browser, created confusion by talking about “a multi-storey car park” – which her peers did not recognize (they could see a pretty seaside villa instead). As a result, the negotiation of meaning did take place, but instead of hinging on the carrying out of a pedagogical role-play focused on the vocabulary of houses and seaside resort facilities, the learning objective became the disambiguation of a real-life misunderstanding. The tool and the mode remained unchanged in the planned version of the task and in the actual task that was carried out, but as the learning objective changed, so in our definition the modality can be said to have changed too.
Diversity in modalities

The impact of screen and perceptual spaces on meaning-making

While Scollon and Scollon (2003) acknowledged the importance of Kress and van Leeuwen’s design dimension, they reinterpreted and reinforced it in order to account for the role played by spaces, in a way that we have found useful for understanding how our users perform on multimodal platforms. Scollon and Scollon call their framework “geosemiotics” and they define it thus:

Geosemiotics is the study of meaning systems by which language is located in the material world. This includes not just the location of words on the page you are reading now but also the location of the book in your hands and your location as you stand or sit reading this. (Scollon and Scollon, 2003, pp. x-xi)

The authors structure geosemiotics into three sub-sets: the interaction order, visual semiotics (on which we will not elaborate here, as this concept comes close to Kress and van Leeuwen’s notion of design mentioned earlier) and space semiotics.

The interaction order provides a way of understanding how individuals perceive the interactional value of the space they choose to use. In their description of the interaction order, the authors include perceptual spaces and interpersonal distances. According to them in face-to-face situations dominant perceptual spaces are visual and auditory (“less noticed” ones are olfactory, thermal, and tactile). The notion of interpersonal distances allows geosemiotics to ask questions about the relationship between space, sound, and socialization. For example, the auditory space that I perceive, and my perceived intimacy or distance with the individual vocalizing the sound that I am hearing, together form the semiotic resource by which I embody meanings. Applying this framework to multimodal platform users, the question becomes: how do they co-construct interpersonal values (intimate, personal, social, public) into conversations that proceed simultaneously through visual spaces of varying salience and through an auditory space defined by the spatially and tactiley intimate device of an earpiece or headset? Here lies the answer to our earlier question about why an online tutor would not use the chat windows to carry out their teaching, while the students chatted off-topic on the audio channel.

Space semiotics, in Scollon and Scollon’s words, is the most fundamental part of geosemiotics, because it asks “Where in the world is the sign or image located?” and because it aims to account for “any aspect of the meaning that is predicated on the placement of the sign in the material world” (2003). For example if students have a choice of inputting a comment into a narrow chat window at the bottom of the screen or into a whiteboard that occupies most of the screen, what will determine their choice: the visual salience of the space, the type of information to be imparted, and/or the nature of the particular group they are working with? A theory of space semiotics applied to CALL situation
needs to address these questions, taking into account wider “ecological” issues such as the fact that any user is free to shrink the screen or to open other screens, which makes the visual message received very different from the visual message sent. In terms of multimodal electronic environments, space semiotics provides the basis for asking questions such as “how do users decode and encode meanings in a material situation involving their computer and its various peripherals (keyboard, mouse or keypad, webcam) as well as other stimuli around them (possibly another computer, a video screen, a person physically present in the same room as them)?”

To put this in semiotic (rather than in multitasking) terms, consider the difference in meaning-making between two situations that I recently experienced. One was when I created a blog at home using a template from a website and sheets of paper to help sketch out the blog’s design. The other was when I created another blog using a WiFi connection while sitting in the audience at a talk given by a speaker on the subject of creating blogs, then publicly revealed what I had been doing, as part of my contribution to an end-of-lecture debate on the pros and cons of allowing ubiquitous WiFi technology on university campuses. In the first situation, I was understanding instructions and carrying them out to create an artifact. In the second situation, I was creating evidence to support my position in a professional discussion. Although the tools and modes were the same in both activities, the impact of place was different, the objective of the person carrying out the activities was different, and therefore overall the modalities can be said to have been different.

**Conclusion**

In offering to the CALL domain a reinterpretation of modality which is based on semiotics, I am aware of stepping into a territory that is just beginning to be charted, although pointers have been provided from neighboring educational research. For example based on his work with virtual games worlds, Lemke (2006) sketches out a research agenda which, I suggest, is applicable to the context of CALL:

[I]t is possible to create real-time, synchronized video and computer log records of monitor display, keystroke and mouse or joystick input, and user speech and action. It is possible in this way to follow user activity in entering the gameworld, acting and moving within the primary gameworld and among various subsidiary “screens” or auxiliary attentional spaces, communicating within and parallel to the gameworld action … and on leaving the gameworld. Ideally we would also like to observe how people integrate or cumulate in-game meaning-making activity and meanings made with out-of-game life activities and identities … We would like to understand class, gender/sexuality, cultural and subcultural differences in which games people play, how, and why; the kinds of meanings they make and feelings they experience; and what persistent learning
Diversity in modalities

effects result. But we need to take such an ambitious agenda one step at a time. (Lemke, 2006, p. 11)

Providing Lemke’s word of caution about a gradual approach is heeded, I see a useful research agenda emerging: to test out, with a large volume of learner interaction data collected from a variety of technological environments, the methodological claims made in this paper according to which what appears to be the bewildering diversity of modality in CALL can be understood in a more systematic way through the synergistic use of social semiotics and geosemiotics. In other words, in this chapter I have been arguing for the development of a cultural approach to understanding the “CA” part of the acronym CALL.