ReCALL **22**(1): 3–19. 2010 © European Association for Computer Assisted Language Learning
doi:10.1017/S0958344009990164

Differences that make the difference: a study of functionalities in synchronous CMC

MARIE-MADELEINE KENNING
School of Language, Linguistics and Translation Studies, University of East Anglia, Norwich NR4 7TJ
(email: m.kenning@uea.ac.uk)

Abstract

This paper has a dual aim: to situate functionalities among the complex of factors that help shape online interactions and to explore the heterogeneity of audio conferencing and its implications. Following a critical discursive synthesis of the treatment of variables in the literature, attention focuses on the distinctive and diverse characteristics of synchronous audio environments. The existence of a synchronous-asynchronous dichotomy is questioned and the effects of a number of differences are highlighted. As a further illustration, the paper next undertakes a comparison of *Lyceum* and *Wimba Voice Direct*. The main features of the two environments are introduced and a detailed analysis of their text and voice chat facilities is carried out. In the last section the benefits of text-based CMC are revisited through the lens of voice chat technology. It is concluded that the benefits of CMC appear to have been overstated and that no general claims can be made about voice CMC, but that collectively, well documented studies could provide a useful source of information on the level of contingency of particular findings.

Keywords: Audio conferencing, CMC, functionality, language learning environments, variables, voice chat

1 Introduction

Overviews and meta-studies of CMC and CALL research have repeatedly expressed concern about a number of faultlines in the research literature. Researchers have been criticized for being too preoccupied with writing about the benefits and potentials of computer technology, instead of exploring the kinds of tasks that should be used and in what settings (Liu, Moore, Graham & Lee, 2003: 264); for producing studies lacking adequate theoretical underpinning (Huh & Hu, 2005), or with a flawed design (Hassan, Hauger, Nye & Smith, 2005; Lamy & Hampel, 2007); for using misleading titles (Felix, 2008); and for leaving out important information. According to Egbert (2005: 4–5), ''CALL research does not address […] differences in context well'', while, after reviewing 78 research studies on CALL, Hubbard (2005: 351) reports that ''researchers commonly leave out or fail to gather relevant data on [subject] characteristics''. In addition, essential information may be insufficiently foregrounded, leaving readers in doubt as to the type of CMC or interaction

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

(learner-learner or learner-teacher) involved. Whether the omission of some key information reflects a lack of appreciation of the complexity of the issues, or whether this is due to faulty reporting procedures, the outcome is something that falls short of a firm basis on which to build.

The premise on which this paper is based is that failure to appreciate the intricate network of interacting variables at work in online language learning constitutes a major obstacle to progress. Bearing in mind Kress' s (2003: 32) observation that "the sculptor who does not understand the potentials of the material with which he or she works is at an entire disadvantage in their work", the author takes the view that it is not possible to compare and contrast findings, to undertake replication studies, or to assess whether and how a particular application might be adapted for use in a different environment without adequate knowledge of the multifaceted nature of online language learning and its openness to the influence of a wide range of factors.

The treatment of variables in the CALL literature suffers from several shortcomings. The first is a tendency to neglect factors outside the technology itself. Bax (2003: 26) calls this the 'Sole Agency' fallacy. There is, at the same time, insufficient awareness of the divergences between different technological environments. This is reflected in, and compounded by, inconsistencies in the use of terminology. Terms like *channel*, *mode* and *modality* are not used in the same manner across the field, with many authors failing to make a distinction between *multimodality* and *multimediality*. Yet, as recognized by the e-learning and design literature, and as argued by Kress and Van Leeuwen (2001: 57): "Multimodality and multimediality are not quite the same thing. Radio [...] is multimodal in its affordances, because it involves speech, music and other sounds; but it is monomedial, since it can only be heard, and not seen, smelled, touched or tasted". There are also challenges in the use of *visual* to refer specifically to images, graphics or icons, as if written language was not perceived through sight. Following a similar evolution, *graphic* has come to imply the presence of symbol manipulation tools such as whiteboards and concept maps (e.g., Hassan *et al*., 2005: 2), as illustrated by Ciekanski and Chanier's (2008: 163) classification of multimodal synchronous online environments into "videoconferencing", i.e., "audio, video and text chat"; "audio-synchronous", i.e., "integrating audio and text chat"; and "audio-graphic conferencing", which integrates "audio, graphics and text chat". All this contributes to creating a confused picture in which one is often uncertain, if only momentarily, of what is being described and evaluated.

This paper seeks to bring greater clarity by highlighting some of the ways in which the many factors liable to influence online language learning interact with one another, with particular reference to the place and role of functionalities and affordances. Functionalities is defined here as features and capabilities (e.g., colour, a particular search facility, a shared tool, a voice activated command), whereas affordances refers to the potential for action offered by the attributes of an artefact, what it is for: " 'Affordance' refers to the perceived and actual properties of a thing, primarily those functional properties that determine just how the thing could possibly be used" (Pea, 1993: 51). For example, stairs are perceived in terms of their climbability (Gaver, 1991), while the affordances of ICT include (information) accessibility, communication and collaboration, and immediacy (Conole & Dyke, 2004). The paper begins with a critical discursive synthesis of the treatment of

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

variables in books and articles on CMC in language learning published since the mid 1990s. Attention then focuses on functionality, starting with an examination of the construct of synchrony before turning to the distinctiveness of virtual audio environments. Next, the paper draws a comparison between the functionalities and affordances of two multimodal environments, *Lyceum* and *Wimba Voice Direct*, before revisiting the literature on the benefits of text-based CMC through the lens of voice chat technology.

2 Variables in the literature

According to Lamy (2004: 521), many studies "fail to consider the totality of the meaning-generating experience of learners operating in synchronous CMC". This was certainly the case in early studies of CMC, which, as Warschauer (1996: 9) pointed out, made claims regarding the equalizing effect of electronic communication that were based on comparisons with whole class discussion, "where it is especially likely that shy students would be afraid to participate". It was obvious to a number of researchers that potential confounding influences needed to be controlled: "Comparisons of the participation patterns and the language and discourse produced in traditional classroom discussions versus electronic synchronous discussions need to ensure that task features and context variables, particularly group format and size, are truly comparable" (Ortega, 1997: 92).

As in communication studies, where references to a single electronic genre such as Netspeak (Crystal, 2001: 18) have given way to the recognition of a range of genres, there is now greater awareness of the complexity of online language use and language learning, and of the need to take into account a wide number of interconnected variables. Whereas, as underlined by Ortega (1997), early research (e.g., Chun, 1994) often failed to recognize the contingent nature of the results, the authors of recent studies are more cautious. They acknowledge the limitations of their studies and include caveats in the presentation of their findings. For instance, Jepson (2005: 85) admits that "because the participants most likely logged-in from different computer stations, it was impossible to control for environmental factors such as computer operating system, peripherals, including microphone and earphone, and Internet connection bandwidth. Variations in these environmental factors may have affected each participant's quantity and quality of participation". There are warnings that the results may not extend beyond the specific environment and population under consideration. Pellettieri (2000: 83), for example, writes: "because of the descriptive nature of this study, generalizations of these results to other populations should be made cautiously". Likewise, Smith (2004) emphasizes that, although the program used in the study was chosen for its simplicity in order to maximize the relevance of the findings, it cannot be assumed that learners using a different program will interact in exactly the same way. However, many authors do no more than warn readers, without delineating in detail the limitations of their studies. The fact is that while few researchers, if any, would contest that "research [on networked language learning] has indicated that there is no single *effect* of using online communication, but rather that processes and results vary widely depending on a range of logistical, pedagogical, and social factors" (Kern, Ware & Warschauer, 2004: 243, italics in the

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

original), the nature, and the potential and actual effects of these factors have been the object of little systematic treatment.

There are several possible reasons for the present situation. One is that examining variables is a daunting task, especially in CMC, due to the "unmanageably high number of variables" (Lamy & Hampel, 2007: 17) that affect human-human interaction. A second reason is that in an age of constant upgrades and improvements, the ground is forever shifting, so that any inventory or analysis involving functionalities is threatened with almost instant obsolescence. Also relevant is the slippery, all-encompassing nature of the term variable, which is used in the literature to refer both to factors that may differ from study to study but remain stable over the study (e.g., the language taught), and to aspects of the learning process liable to change over time (e.g., anxiety). How wide should the net be cast?

Notwithstanding these difficulties, some scholars have attempted to map out the field, albeit with different outcomes, owing to the multitude of factors at play, the diverse standpoints from which classification can be undertaken, and whether one is dealing with CALL or CMC. Thus Egbert (2005: 4–5) defines CALL as a process involving six basic categories of variables: learner, language, context, one or more tools, tasks/activities and peers and teachers, while Felix's (2008) meta-study of the effectiveness of CALL identifies eight umbrella categories (number of participants, research design used, technology used, educational setting, language taught, subject/skill taught, variable under investigation, reported findings), some of which cover several variables. Focussing on CMC, Lamy (2004: 525) argues that the definition of oral competence in a multimodal learning environment requires not only "accounts of the discourse that is produced by the interactants" but also "accounts of

- the materiality of the tool involved (interface and hardware);
- the pedagogical philosophies that underpinned the design of the tools (if it was locally created) or that guided its selection (if the institution leased or bought it);
- the social relationships between the interactants (e.g., of power, exerted within the institution where the learning is taking place);
- the materiality of the learning situation (e.g., in a language lab, or in the participant's office, or home)".

Lastly, according to Wang (2004: 375), CMC "encompasses the following characteristics:

- It can be text-based, oral, and/or visual [i.e., videoconferencing links][1]
- It can be either synchronous or asynchronous
- It can be one to one, one to many, or many to many[2]

[1] As well as graphical (i.e., pictures, as in *The Palace* and *Second Life*).

[2] Many to many is commonly used in the literature to refer to an environment in which any participant can communicate with any, several, or all of the other participants. In practice, however, many to many consists of multiple one to one and/or one to many communications, many of which involve only a small number of active participants interacting in front of an audience.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

- It can be learner to learner, learner to instructor, or learner to native speaker
- It can be time and place dependent or independent''.

Despite their differences which, at times, can be difficult to reconcile, broad categorizations provide a valuable overview of the landscape that highlights the wide range of influences that affect communication. As such, they offer a backcloth against which studies can be evaluated or applications devised.

A prime source of information on what scholars regard as significant is what they have chosen to investigate (e.g., in-class versus out-of-class in Sanders, 2006). Pointers can also be found in the comments made in the analysis of the data, e.g., discussions of age differences (Lee, 2004), of mismatches in foreign language proficiency (Belz, 2002; Lee, 2004) and, more widely, in the narratives of qualitative research. Generally speaking, studies investigating a particular variable, especially from an interactionist perspective, have concentrated either on variables that have been the subject of research outside CALL, for example task type (Blake, 2000; Pellettieri, 2000; Smith, 2003), or on salient technological differences, usually conceptualized, at least initially, as dichotomies rather than polarities: asynchrony versus synchrony, voice versus text-based communication.

3 Functionality

It is in relation to functionalities and affordances that one finds the clearest signs of increasing sophistication. This is not an uncommon phenomenon. Ong (1982: 3) points out that one of the effects of technological progress is to bring to the fore facts that had remained unnoticed hitherto: ''Our understanding of the differences between orality and literacy developed only in the electronic age, not earlier. Contrasts between electronic media and print have sensitized us to the earlier contrast between writing and orality''. In a similar way, the advent of voice and multimodal networked environments has contributed to the problematization of what used to be regarded as straightforward characteristics, and led to the making of finer distinctions.

Synchrony is a case in point. It has long been common practice to describe communication and environments as either synchronous or asynchronous, and it is indeed one of the few variables routinely specified in CMC studies. What has become clear, however, in recent years is the amount of temporal variation that each type covers.

Kenning (2007: 53) defines synchrony as ''the necessary availability of the participants for the whole of a communicative event''. In an electronic context, synchrony implies that the participants are logged on at the same time. However, contrary to what might be assumed, the fact that an environment is synchronous does not mean that production and reception necessarily take place (quasi) simultaneously, in real time, as happens in face-to-face. In the case of synchronous written networked environments, messages may appear on the recipient's screen as they are being typed, but they may also only appear when sent. This is pointed out by Pellettieri (2000: 61), who underlines that *Ytalk*, the software program in her study, differs from *Daedalus InterChange*, a commonly used program in language teaching, because recipients view messages ''letter by letter, as they are typed'', whereas ''in *InterChange* users

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

view only the final version of their partners' composed utterances". It is an important difference since, as Pellettieri observes, participants in *Ytalk* can interrupt each other. A second important consequence of the immediate appearance of what is typed is that participants cannot edit their messages in private, in contrast to *InterChange*, where "as students feel that their messages are complete, they 'publish' them by sending them to the other members of their 'discussion group'" (Kern, 1995: 458). It can therefore be argued that *InterChange* represents a safer environment for learners than *Ytalk*, with all the implicational consequences this may have for their emotional states. On the other hand, the delayed appearance of recasts in systems that only display completed messages may contribute to their going unnoticed (Lai & Zhao, 2006: 113)[3].

There are similar variations in the case of oral networked interactions. To begin with, as with written CMC, transmissions may be subject to a slight, but noticeable, time lag which makes production and reception less co-terminous than in face-to-face. Blake (2005: 509), referring to bimodal CMC, reports that users find the persistent problems of time lag confusing, at least for the first few sessions. Like silences in telephone conversations, time lags prompt participants to undertake channel checks in order to ascertain whether their interlocutors are still there (e.g., can you hear me?), and interfere with the flow of the conversation. Second, interactions are affected by the nature of the link, which may allow signals to flow in both directions at once (full duplex), thereby opening up the possibility of backchannel, overlaps and interruptions, or only in one direction (half duplex), with turn-taking regulated through the use of icons (thereby excluding vocal backchannel, overlaps and interruptions). The difference has significant implications for the incidence of, for instance, negotiation for meaning and self correction, since the only point at which an interlocutor can intervene in a half duplex system without resorting to another tool (e.g., text chat), provided one is available, is when the speaker releases the floor. In effect, half duplex reduces an interlocutor's opportunities for initiating negotiation and gives more opportunity for self correction and self repair to the speaker. This should be taken into account in any study of negotiation in a half duplex online environment.

Although unavoidable, the use of generic labels is potentially misleading. While it is convenient to refer to environments as synchronous or asynchronous, one must not overlook the fact that the relationship between synchrony and asynchrony is best approached as a matter of degree, as done by Ciekanski and Chanier (2008: 173). Rephrasing and expanding their description of audio as more synchronous than text chat, which in turn is more synchronous than a shared word processor, we may say that face-to-face offers greater simultaneity than audio networks, audio than text chat, and text chat than a shared word processor.

The synchrony/asynchrony issue provides a glimpse into the way in which the functionality of a tool can facilitate and constrain certain types of interaction, and indicates the need for fine-grained studies in which this kind of information is

[3] The appearance of input letter by letter has ceased to be widely used; however, the example aptly demonstrates the potential implications of an apparently minor variation.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

provided. Only through ethnographic studies and "thick" descriptions (as described by Ponterotto, 2006) can one come to an appreciation of the differences and similarities between technological environments, identify the effects of other contextual factors, and make a distinction between findings that are generalizable and those that are specific to a particular environment.

One of the most beneficial effects of the spread of voice networks to language teaching has been to put the spotlight on oral communication and draw attention to the way in which oral interactions are affected by the context of use. This was not the case in studies of text-based CMC, which have often used face-to-face as a comparator or as a frame of reference, without always giving due consideration to the contextual dimension of interactions. So much so that despite the widespread use of the telephone in everyday life, if not in teaching, oral interactions, oral medium, oral mode and face-to-face have occasionally been used interchangeably (Fernández-García & Martínez Arbelaiz, 2003; Kitade, 2000). Scholars involved in research into the applications of voice-based and multimodal environments have been at pains to emphasize the distinctive nature of these environments. Developing Erben's (1999: 232) argument that "one cannot readily take for granted that face-to-face and electronically networked immersion classes are the same", they stress that speaking online, in front of a screen, typing, dragging or clicking on a mouse is significantly different from having a conversation face-to-face or on the telephone (Hampel, 2006; Lamy, 2004). They point out that the presence of technological mediation changes the nature of interactions and alters discursive practices in ways which vary with "the specificity and materiality of the tool itself" (Lamy, 2004: 520). This is a key point, for, in facilitating or constraining certain types of interaction, the functionality of an environment affects how people handle language, and promotes or impedes the display of different types of competence, which, arguably, in turn impacts on SLA.

Here it should be stressed that what is being referred to is more than broad features. The range of artefacts at the user's disposal (mouse, headset), the design of the interface (e.g., location and prominence of icons), voice quality, the possibility of anonymity and/or including photos, etc. are all factors that are liable to exert an influence on interactions. Observing that telecollaboration often takes place via tools embedded in learning management systems such as *WebCT*, *Moodle* or *Blackboard*, O'Dowd and Ritter (2006: 633) report that in their experience "even supposedly little differences between such systems, or the way they are set up, can have an effect on the exchange". Therefore, to simply refer to, say, *Blackboard*, in a study, is not enough.

Crucially, differences in usability will have differential implications in different contexts. They may have little discernible effect, or they may make a significant difference. In a similar vein, there is considerable evidence that a particular feature is not always viewed in the same way. Thus the loss of embodiment that occurs in an online environment without webcams can be perceived as advantageous, detrimental, or neutral: "half of the respondents found 'not seeing or not being seen by the other learners to be neither a positive nor a negative experience and some (5) even found it to be an outright positive event" (Hauck, 2007: 211). It is tempting to speculate, therefore, that the turn delays experienced in some forms of voice networks will be more frustrating, if only because they will be more noticeable, for fluent speakers used to a quick tempo than for beginners who need time to process

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

speech, just as a participant with a low level of digital literacy may cause frustration to those with better digital skills.

Among the more recent insights to be gained from the literature is the finding that the availability of alternative means of communication tends to alter the use to which a tool is put. Several studies (Blake, 2005; Hampel *et al*., 2005; Sauro, 2004) indicate that text chat is used differently when embedded in a multimodal environment. In such environments text chat tends to be employed to signal or circumvent sound problems, to clarify without interrupting, to check things that seem easier to check in writing, or for off-topic conversations. Significantly, a similar kind of strategic discursive specialization is mentioned by Kern (1995: 459), who reports that students sitting together in an *InterChang*e session talk to each other (face-to-face) but that "their talk tends to be metalinguistic ('how do you say…?') or metadiscursive ('that's the same argument you made before') in nature". In both cases it can be seen that learners' behaviour is aimed at making the most of the modes of communication at their disposal. The purpose may not be the same in all cases. For instance, in the case of the dyad studied by Sauro (2004), text chat was used tactically by the less dominant speaker to maintain a foothold in the conversation and reposition herself. Generally, where there is a range of options, one would expect learners to be influenced by a whole array of factors, notably their distinctive attributes, e.g., ease of access (whether you have to wait your turn or whether everyone can input at the same time), the nature of the task, and individual preferences.

One factor which has not received the attention it deserves is the impact of the evolution of technology on expectations. Learning cannot but be coloured by what happens outside education, so that the way in which learners approach CMC applications will be influenced by their experience of CMC in everyday life. Over the past few years virtual environments such as *Second Life* and social networking websites like *Facebook* have become a common context for the social use of language and the number and pace of interactions have seen significant increases. These changes in the dynamics of communication are bound to affect learners' behaviour and attitudes and, like the refinements and new functionalities and affordances brought by technological progress, can have a potentially adverse effect on the leverage of previous work.

4 Comparison of Lyceum and Wimba Voice Direct

As an environment that figures prominently in the existing body of research on CMC, *Lyceum* is a prime candidate for inclusion in a comparison of audio conferencing online environments combining text and voice chat. Not least because its functionalities and affordances have been described and discussed in detail, and there is plentiful information on how *Lyceum* is used and on user feedback (Hampel, 2003; Hampel *et al*., 2005, Hampel & Hauck, 2006; Lamy, 2004; Rosell-Aguilar, 2007; Vetter, 2004, etc.). Unlike *Lyceum*, *Wimba Voice Direct* does not appear to have been much used in the published research. However, in contrast to *Lyceum*, an institution-specific groupware system providing voice conferencing together with synchronous, visual workspace tools (Buckingham Shum *et al*., 2001), it is commercially available and can be integrated into popular learning management systems

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

such as *Blackboard* and *WebCT*. Given this very different positioning, it is impossible to ignore the issue of the transferability of *Lyceum* applications to a *Wimba* environment. Put plainly, to what extent is it possible for someone working with *Wimba Voice Direct* to adopt or customize applications, ideas and strategies outlined in the research involving *Lyceum*? The comparison below, which is based on the author's exploitation of *Wimba* tools at the University of East Anglia, is offered in part to try and get a better idea of what is a question of considerable practical and pedagogical interest.

It might be objected that a single-function tool like Wimba cannot match the possibilities for complex interactions offered by a multifunction groupware system. It must be stressed, therefore, that the intention is not to run *Wimba Voice Direct* down, but rather to highlight some of the issues raised by technological diversity. It is argued that the sharing of best practice among CALL scholars needs to go beyond the publication of significant research results to the ability to implement successful applications and good teaching ideas across the field. One possible way of achieving this is through the creation of reusable resources (i.e., software artefacts) of various degrees of granularity (size). Reusability, however, is a complex issue presenting many challenges (Littlejohn, 2003), including the fact that it is unlikely to have a high priority among the goals of language teaching academics, who tend to be primarily concerned with their own students. The problem is that, without direct access to the technological environment involved – a not uncommon situation – it can be difficult to ascertain the practical relevance to one's own work of a study discussing effective learning experiences developed in another place, and this difficulty acts as a barrier to the spreading of good pedagogic practice. The dissemination of ideas and results through books and articles is undoubtedly essential for a community of practice; but for novices to become experts, for those in less well endowed institutions to progress beyond peripheral participation, there must be means to judge to what extent successful ideas can be implemented in another context and environment. This will allow a larger number of students to benefit from the work and insights of the original researchers and developers, will be economically advantageous in increasing cost-effectiveness, and can help bridge, albeit modestly, a growing divide between institutions with elaborate systems and a strong research output and less favourable environments. From this perspective *Lyceum*, despite its planned replacement, and *Wimba Voice Direct,* are an appropriate choice, reflecting the situation at the chalk face and illustrating the challenges that arise in this kind of comparison.

Lyceum and *Voice Direct* have three important defining characteristics in common: both offer audio and text chat; neither offers video; and in contrast to the environments used in studies such as Sauro (2004) and Jepson (2005), they are tuition environments, rather than general chatroom systems. However, *Voice Direct* is a basic voice plus text system, whereas *Lyceum* comprises five tools: two communication tools (audio and text chat) and three shared editing tools (whiteboard, concept map, document (word processor)) (Hampel *et al*., 2005: 5). This means that instead of being confined to the linearity and inalterability of output that is inherent to audio and text chat, users of *Lyceum* can incorporate images and diagrams, control where items are placed, and make alterations. Some of these functionalities can be introduced into a *Wimba* environment, but this requires using *Voice Direct* in

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

tandem with other stimuli and tools (e.g., wikis). The feasibility of such an approach and the likelihood of a successful outcome depend on many factors that cannot be examined here, but the lack of integration of the various components can create interface and manipulation problems that interfere with communication and may jeopardize the activity. The room metaphor of the *Lyceum* screen, which is "designed to enhance the intuitive nature of the different tools, and to reduce the potential anxiety and inhibitions of students unfamiliar with such online environments" (Hampel *et al.*, 2005: 2) is another feature alien to *Wimba*, which, at least as it was set up at the university in question, does not allow learners to move around freely or groups to be easily reconfigured.

Hampel (2006: 112) describes the functionalities and affordances of *Lyceum* as comprising, *inter alia*:

> "– joint production of texts and images;
> – online connection allows users to import text or images from the World Wide Web;
> – texts and images can be saved;
> – simultaneity of certain modes (audio and text chat);
> – no privileges inbuilt into software for tutors: all participants have equal rights, at least technically;
> – existence of sub-conferences (which can be used for group work);
> – lack of body language".

Significantly, of these seven characteristics, only two (the 4th and 7th) apply to *Voice Direct*. Some of the others may be applicable if other tools are brought in, but they are not part of the basic software.

As might be expected, the differences between the environments which text chat and voice chat belong to have implications for the purposes to which they are put and for their relationship. Clearly, unless some kind of substitute is provided for the absent tools, text and audio chat cannot have identical functions in *Voice Direct* and *Lyceum*. For instance, since *Voice Direct* does not offer a whiteboard, students will not use text chat "to express in writing what they see on the whiteboard", as is the case with *Lyceum* (Hampel & Hauck, 2006: 11). What may happen, however, is that learners seek to use text chat or voice chat for purposes similar to those which the additional tools of *Lyceum* are used for, for instance brainstorming. But this only serves to highlight that the relationship between audio and text chat will not be the same in the two environments.

When comparing text and voice chat in *Lyceum* and in *Voice Direct*, the following similarities and differences come to the fore. First, both technologies allow oral interactions to be recorded digitally in real time, and text chat content to be saved. Second, both use a set of icons to manage turn-taking in audio. Specifically, both have icons for requesting the floor and gaining access to the floor as well as icons, or equivalent devices, showing who is holding the floor, who has requested the floor, and the order in which people are queuing. However, *Lyceum* also has 'voting' buttons as well as an 'away' button: clicking on the voting buttons enables participants to answer yes or no to a question without taking the floor, while clicking on the away button enables them to indicate that they are leaving their computer

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

temporarily and that no response should be expected from them (Hampel *et al.*, 2005: 6; Lamy, 2004: 531–532). In *Voice Direct*, the provision of such information requires verbal communication through voice chat or text chat. A second area of difference is that *Voice Direct* features a half duplex system so that the current speaker has control of the floor (but see below). *Lyceum*, on the other hand, allows several people to talk and be heard, even if, in order to avoid interruptions, participants are expected to request the floor by clicking on a 'hand' icon instead of clicking immediately on the 'talk' icon that opens their microphone (Lamy, 2004: 531). As already mentioned, *Lyceum* does not grant special privileges to tutors. *Voice Direct* is less democratic in this respect. Whereas tutors can intervene (although they do not have to) to take over the microphone or to give it to someone who has requested the floor, other participants cannot. Finally, *Lyceum* allows text chat to be turned on and off, whereas the text chat facility of *Voice Direct* cannot be disabled. This means that it is not possible to stop students from adopting avoidance strategies. Since tutors can only pass the microphone to someone if they have requested it (see *Wimba tools* website), students can choose to bypass voice chat altogether and communicate solely in writing, although tutors can try and dissuade them.

Clearly, the differences between audio and text chat in the two systems have interactional implications (for example, with respect to who can intervene, and how) that may prove vital in some cases. At the same time, the chatting facilities that they offer have much in common, especially in relation to the management of turn-taking. This stands out when *Lyceum* and *Voice Direct* are compared with face-to-face and telephone communication. Provided *Lyceum* users adhere to the set protocol, and *Voice Direct* tutors refrain from interfering, both systems share with most text-only environments the much claimed advantage of there being no turn-taking competition (Blake, 2005: 499; Kitade, 2000: 149). This raises the question of how many of the claimed benefits of written synchronous CMC apply to voice-based chat technology.

5 Benefits of synchronous CMC revisited

Most of the claims that have been made about synchronous text-based CMC have to do with its impact on learner participation. CMC is primarily credited with

1. increasing learner participation
2. promoting more equal participation patterns
3. lessening teacher dominance

Learners are reported (e.g., Chun, 1994; Kern, 1995) to produce more output than in traditional classroom activities, to use a wider variety of discourse functions, and to receive more input from their peers, with a concomitant reduction in the teacher's input. There is also a tendency towards more equal participation (Beauvois, 1998; Warschauer, 1996), with shy students benefiting the most.

Several interrelated factors have been invoked to account for the changed participation patterns listed above. One self-evident explanation for increases in learner participation is the novel opportunities that CMC provides to engage in 'talk', including the possibility of setting up small group discussions. However, in order to trigger the increases in *genuine, meaningful* talk involving the production of comprehensible output

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

and the processing of comprehensible input that are believed to lead to acquisition, opportunities for talking must be embedded in appropriate tasks. This gives pride of place to pedagogy, including expertise in exploiting technology, rather than to functionality, and makes the advantage potential rather than automatic.

Another explanation put forward in the literature is that CMC creates a congenial environment. Participants are said to have time to reflect and to be shielded from immediate exposure to other people's reactions, since they cannot see those they interact with. They can work at their own pace, "without the dangers of being interrupted, making interlocutors become bored or impatient, receiving physical or verbal evaluative sighs from the audience, or forgetting one's own ideas while waiting for an opportunity to take the floor" (Ortega, 1997: 91). There is undoubtedly some truth to these putative advantages, but they are far from being established facts, and there is another side to some of the points. In the first place, text chat frequently contains emoticons, acronyms, and punctuation marks designed to compensate for the lack of 'normal' paralinguistic cues. Second, there is insufficient information on how learners process text chat (as distinct from other kinds of written CMC), including how much scrolling up and down they perform, how much time they spend planning their interventions, and how far they consider other participants' possible reactions. One must also beware of overplaying the opportunity for reflection that CMC is supposed to provide, since spending too long reading or composing messages may cause other participants to become bored and can lead to being sidelined in multiparty exchanges. Users might not receive sighs, but they may receive negative comments. Moreover, feedback from others can be positive (nods, smiles, etc.) as well as negative. Finally, as indicated above, editing opportunities depend on the mechanisms through which text is displayed on other participants' screens. In short, there are a number of question marks over the congeniality of text-based CMC.

On the other hand there is little doubt that text-based CMC creates a less pressured environment than the 'on the fly' processing and sense of urgency that are commonly associated with oral communication. As noted above, text chat allows students to formulate their ideas as they occur to them. In contrast, students communicating through oral virtual environments must wait for their turn, either obligatorily with a system such as *Voice Direct*, or in order not to break the protocol in *Lyceum*. Consequently, they run the risk, as they do in face-to-face, of forgetting their ideas in the meantime, unless they also have access to text chat and use it to jot things down. Above all, written CMC does not require students to speak, when it is well known that speaking, for many learners, remains the most challenging aspect of learning a foreign language and a major source of communication apprehension (Arnold, 2007: 471). Any environment that enables learners to communicate without having to speak is likely to be perceived as less inhibiting, and as such may encourage production. Conversely, having to speak in front of others or with them is likely to provoke anxiety, as supported by evidence showing that, whether in classroom-based learning or in a distance setting, learners fear not being understood and worry about their accent (Hauck & Hurd, 2005). There are grounds to believe, therefore, that voice chat and voice plus text chat may be less conducive to interaction than text chat. It may be that the loss of embodiment that occurs in an online environment without webcams has a mitigating influence but, to the author's knowledge, no study

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

has yet demonstrated that the levels of anxiety associated with a virtual speaking environment are lower than those experienced by the same learners in face-to-face, let alone accounted for such a decrease.

To the impact of the modality must be added the influence of the locus of the activity. In a critical analysis of Blake's (2000) study, in which students sitting in a lab interacted with each other online in writing, Lamy & Hampel (2007: 118) rightly ask: "Would the outcomes have been different if the learners had not been co-located in a lab, but had participated at home from their own computers?" In the case of voice chat, evidence gathered at the author's university suggests that co-location has a negative effect on students' willingness to communicate orally online, and that better results are achieved by placing students on their own in different rooms. The physical presence of others, including the teacher, appears to have a dampening effect. It makes the learner feel conspicuous, on stage, a temporary feeling which diminishes when interlocutors can no longer be seen. Thus, commenting on why she had been less reluctant to speak online when alone in a room, one student wrote: "It was the only option, making it seem more appealing as you couldn't just turn around and talk to the person. It forced the issue but also made it a lot easier knowing no-one could hear you other than who you were talking to". The partner the student was interacting with was someone she was friendly with, which adds personal rapport to the mix, in line with the response of the student cited in Hampel *et al*. (2005: 14) who stated: "I felt more at ease in the text chat especially if people are around that I don't know well".

One factor that might be assumed to work against equal participation in any kind of CMC, but particularly in voice CMC, is language proficiency. There is empirical evidence (e.g., Lee, 2004: 94–95), that NS-NNS written networked interactions can sometimes be dominated by the native speakers. Given that differences in proficiency levels tend to be more striking in oral communication, one would expect less fluent speakers to participate less. Yet, Vetter and Chanier (2006: 14) found "no apparent correlation between perceived linguistic level and degree of audio participation". The reason for this counterintuitive finding is unclear, but may have something to do with the distinctive way in which turn-taking operates in audio conferencing. It may be that the visibility of the 'hand up' mechanism of systems like *Lyceum* and *Voice Direct* incites the more reticent learners to take part (for fear of being shown up as mere 'passengers'), and at the same time restricts the length of interventions when there is a queue, so that the floor is shared more equally. Both the finding itself and the reasons behind it need further investigation. Like Sauro (2004), Vetter and Chanier also found evidence that text chat can serve to rebalance inequalities in oral contributions, with some learners using text chat as a support, or a substitute, for oral communication. There is some doubt, however, over the effectiveness of this compensatory effect and, by implication, over the potentially equalizing effect of a dual system. Significantly, the change in position of Sauro's (2004) less dominant speaker (referred to above) did not happen within the confines of the dyad; it followed an intervention by the researcher. While the student was able subsequently to use text chat to *maintain* a foothold in the conversation, she might not have been able to *gain* a foothold and reposition herself on her own, if the researcher had not intervened.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

With regard to teacher dominance, voice CMC resembles text-based CMC in offering an environment in which learners are able to take charge of an activity. At the same time, as illustrated by the comparison between *Lyceum* and *Voice Direct*, environments differ in the level of control given to the teacher, the teacher's behaviour introducing further variations. The same is true of written CMC. Whether or not the teacher exploiting *Daedalus Interchange* chooses "to participate actively in the discussion or to observe students' exchange" (Kern, 1995: 458) will affect learner control. As with other issues, whether learners can and do avail themselves of the opportunity to take over discussion management, for example, partly depends on the attributes of the technology, but it is also a function of many other factors, including the teacher's approach, the learners' level of proficiency and their ability to become autonomous. This, together with what was said about equal participation, supports Yates' (2001: 29) view that "CMC is not the open democratic medium that some educationalists and technologists initially hoped for".

In fact, it is difficult to escape the conclusion that the benefits of CMC have been overstated. On closer examination, the alleged benefits of written CMC appear rather spurious and not as closely connected with functionality and affordances as they are often taken to be. They are, at best, potential benefits that require optimal conditions on a number of dimensions to fully materialize. The same would seem to apply to voice CMC and dual systems. In so far as voice CMC and dual systems add to the range of opportunities learners have to communicate in the target language, there is clearly some potential to increase participation in oral interactions. However, there is a need for studies illustrating how this can be achieved, as well as for investigations of the equal participation issue, over which there remains considerable doubt. Lastly, as with text-based CMC, whether voice CMC entails a reduction in teacher dominance will depend on many factors and should not be regarded as a given.

6 Conclusion

This paper has shown that functionalities and affordances cannot be adequately discussed without reference to a particular context made up of a motley collection of factors, most of which have manifold differential implications. To carry out an exhaustive inventory of the factors, even if it were feasible, would be outside the scope of the paper. However, following O'Dowd and Ritter (2006), it is helpful to recognize three interconnected levels:

- the individual level
- the methodological level
- the socio-institutional level.

The individual level comprises all the variables brought by the learner to the learning process, such as age; proficiency (including both linguistic and sociocultural expertise); affective factors (e.g., self-confidence), motivation and expectations; familiarity with technology. The methodological level covers tuition related aspects such as task design; student preparation; learner allocation (size and configuration of groups, including matching procedures in relation to proficiency level, gender, personal rapport); group characteristics (degree of homogeneity, number of shared

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

languages). Variables belonging to the socio-institutional level, a level that has attracted much attention in the literature on telecollaboration, include functionality and affordances, locus of the activity; distance or classroom learning; relation to other aspects of the course and everyday life.

With reference to voice chat and combinations of voice and text chat, the paper pinpointed the following as especially relevant to discourse patterns:

- level of integration with other devices and tools
- tutor privileges
- number and function of icons
- type of transmission (full versus half duplex).

However, it must be emphasized that while one can make general hypotheses about the consequences of a given feature, outcomes are sensitive to too many influences for the nature and significance of the effects of a particular functionality to be specified. Even with empirical data, it is likely to be difficult to isolate the effects of individual factors. The purpose of the framework and lists above is not to provide a formula for assessing influences, but rather to act as a reminder of the range of variables that should be taken into account.

Case studies with controlled variables would undoubtedly be useful, but require facilities and resources beyond most researchers. However, even outside strictly comparable contexts, it should be possible, with access to a body of properly documented studies, to illuminate the degree of contingency of findings by examining their levels of recurrence. This, in turn, would shed some light on which variables are most significant and which differences actually make a difference.

References

Arnold, N. (2007) Reducing foreign language apprehension with computer-mediated communication: a preliminary study. *System*, **35**(4): 469–486.

Bax, S. (2003) CALL – past, present and future. *System*, **31**(1): 13–28.

Beauvois, M. H. (1998) Conversations in slow motion: computer-mediated communication in the foreign language classroom. *The Canadian Modern Language Review*, **54**(2): 198–217.

Belz, J. A. (2002) Social dimensions of telecollaborative foreign language study. *Language Learning & Technology*, **6**(1): 60–81.

Blake, R. (2000) Computer mediated communication: a window on L2 Spanish interlanguage. *Language Learning and Technology*, **4**(1): 120–136.

Blake, R. (2005) Bimodal CMC: the glue of language learning at a distance. *CALICO Journal*, **22**(3): 497–511.

Buckingham Shum, S., Marshall, S., Brier, J. and Evans, T. (2001) Lyceum: Internet Voice Groupware for Distance Learning. In: *Proceedings of Euro-CSCL 2001: 1st European Conference on Computer-Supported Collaborative Learning*, Maastricht, The Netherlands http://www.ll.unimaas.nl/euro-cscl/Papers/24.pdf

Chun, D. M. (1994) Using computer networking to facilitate the acquisition of interactive competence. *System*, **22**(1): 17–31.

Ciekanski, M. and Chanier, T. (2008) Developing online multimodal verbal communication to enhance the writing process in an audio-graphic conferencing environment. *ReCALL*, **20**(2): 162–182.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

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

Crystal, D. (2001) *Language and the Internet*. Cambridge: Cambridge University Press.

Egbert, J. L. (2005) Conducting research on CALL. In: Egbert, J.L. and Petrie, G.M. (eds.), *CALL research perspectives*. Mahwah, New Jersey: Lawrence Erlbaum, 3–8.

Erben, T. (1999) Constructing learning in a virtual immersion bath: LOTE teacher education through audiographics. In: Debski, R. and Levy, M. (eds.), *WORLDCALL: global perspectives on computer-assisted language learning*. Lisse, The Netherlands: Swets & Zeitlinger, 229–248.

Felix, U. (2008) The unreasonable effectiveness of CALL: what have we learned in two decades of research? *ReCALL*, **20**(2): 141–161.

Fernández-García, M. and Martínez Arbelaiz, A. (2003) Learners' interactions: a comparison of oral and computer-assisted written conversation. *ReCALL*, **15**(1): 113–136.

Gaver, W. W. (1991) Technology affordances. In: *Proceedings of the CHI 1991 Conference*. New York, NY: Association for Computing Machinery, 79–84.

Hampel, R. (2003) Theoretical perspectives and new practices in audio-graphic conferencing for language learning. *ReCALL*, **15**(1): 21–36.

Hampel, R. (2006) Rethinking task design for the digital age: a framework for language teaching and learning in a synchronous online environment. *ReCALL*, **18**(1): 105–121.

Hampel, R., Felix, U., Hauck, M. and Coleman, J. (2005) Complexities of learning and teaching languages in a real-time audiographic environment. *German as a Foreign Language*, **3**, http://www.gfl-journal.de/3-2005/hampel_felix_hauck_coleman.pdf

Hampel, R. and Hauck, M. (2006) Computer-mediated language learning: making meaning in multimodal virtual learning spaces. *JALT CALL Journal*, **12**(2): 3–18.

Hassan, X., Hauger, D., Nye, G. and Smith, P. (2005) The use and effectiveness of synchronous audiographic conferencing in modern language teaching and learning (online language tuition): a systematic review of available research. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Hauck, M. (2007) Critical success factors in a TRIDEM exchange. *ReCALL*, **19**(2): 202–223.

Hauck, M. and Hurd, S. (2005) Exploring the link between language anxiety and learner self-management in open language learning contexts. *European Journal of Open Distance and E-Learning*, http://www.eurodl.org/materials/contrib/2005/Mirjam_Hauck.htm

Hubbard, P. (2005) A review of subject characteristics in CALL research. *Computer Assisted Language Learning*, **18**(5): 351–368.

Huh, K. and Hu, W.-C. (2005) Criteria for Effective CALL Research. In: Egbert, J.L. and Petrie, G.M. (eds.), *CALL Research Perspectives*. Mahwah, New Jersey: Lawrence Erlbaum, 9–21.

Jepson, K. (2005) Conversations – and negotiated interaction – in text and voice chat rooms. *Language Learning and Technology*, **9**(3): 79–98.

Kenning, M.-M. (2007) *ICT and language learning: from print to the mobile phone*. Basingstoke: Palgrave.

Kern, R. G. (1995) Restructuring classroom interaction with networked computers: effects on quantity and characteristics of language production. *The Modern Language Journal*, **79**(4): 457–476.

Kern, R., Ware, P. and Warschauer, M. (2004) Crossing frontiers: new directions in online pedagogy and research. *Annual Review of Applied Linguistics*, **24**: 243–260.

Kitade, K. (2000) L2 learners' discourse and SLA theories in CMC: collaborative interaction in Internet chat. *Computer Assisted Language Learning*, **13**(2): 146–166.

Kress, G. (2003) *Literacy in the new media age*. London: Routledge.

Kress, G. and Van Leeuwen, T. (2001) *Multimodal discourse*. London: Arnold.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press

Lai, C. and Zhao, Y. (2006) Noticing and text-based chat. *Language Learning and Technology*, **10**(3): 102–120.

Lamy, M.-N. (2004) Oral conversations online: redefining oral competence in synchronous environments. *ReCALL*, **16**(2): 520–538.

Lamy, M.-N. and Hampel, R. (2007) *Online communication in language learning and teaching*. Basingstoke: Palgrave MacMillan.

Lee, L. (2004) Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning and Technology*, **8**(1): 83–100.

Littlejohn, A., (ed.) (2003) *Reusing online resources: a sustainable approach to e-learning*. London: Kogan Page.

Liu, M., Moore, Z., Graham, L. and 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.

O'Dowd, R. and Ritter, M. (2006) Understanding and working with 'failed communication' in telecollaborative exchanges. *CALICO*, **23**(3): 623–642.

Ong, W. J. (1982) *Orality and literacy: the technologizing of the word*. London: Routledge.

Ortega, L. (1997) Processes and outcomes in networked classroom interaction: defining the research agenda for L2 computer-assisted classroom discussion. *Language Learning and Technology*, **1**(1): 82–93.

Pea, R. D. (1993) Practices of distributed intelligence and designs for education. In: Salomon, G. (ed.), *Distributed cognitions – psychological and educational considerations*. Cambridge: Cambridge University Press, 47–87.

Pellettieri, J. (2000) Negotiation in cyberspace: the role of c*hatting* in the development of grammatical competence. In: Warschauer, M. and Kern, R. (eds.), *Network-based language teaching: concepts and practice*. Cambridge: Cambridge University Press, 59–86.

Ponterotto, J. G. (2006) Brief notes on the origins, evolution, and meaning of the qualitative research concept 'thick description'. *The Qualitative Report*, **11**(3): 538–549. http://www.nova.edu/ssss/QR/QR11-3/ponterotto.pdf

Rosell-Aguilar, F. (2007) Changing tutor roles in online tutorial support for open distance learning through audio-graphic SCMC. *The JALT CALL Journal*, **3**(1–2): 81–94.

Sanders, R. (2006) A comparison of chatroom productivity: in-class versus out-of-class. *CALICO Journal*, **24**(1): 59–76.

Sauro, S. (2004) Cyberdiscursive tug-of-war: learner repositioning in a multimodal CMC environment. *Working Papers in Educational Linguistics*, **19**(2): 55–72.

Smith, B. (2003) Computer-mediated negotiated interaction: an expanded model. *Modern Language Journal*, **87**(1): 38–57.

Smith, B. (2004) Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition*, **26**(3): 365–398.

Vetter, A. (2004) Les spécificités du tutorat à distance à l' Open university: enseigner les langues avec Lyceum. *ALSIC*, **7**: 107–129.

Vetter, A. and Chanier, T. (2006) Supporting oral production for professional purposes in synchronous communication with heterogeneous learners. *ReCALL*, **18**(1): 5–23.

Wang, Y. (2004) Distance language learning: interactivity and fourth-generation Internet-based videoconferencing. *CALICO Journal*, **21**(2): 373–395.

Warschauer, M. (1996) Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, **13**(2/3): 7–26.

Wimba tools: http://www.wimba.com/support/docs/WVTM5_UserGuide.pdf

Yates, S. J. (2001) Gender, language and CMC for education. *Learning and Instruction*, **11**: 21–34.

https://doi.org/10.1017/S0958344009990164 Published online by Cambridge University Press