MALL:
Creating a Theoretical Framework
for the Development of Episodic,
Live-action, Dialog-style Mobile Content
for Adult EFL Learners

MATSUZAKI, Takeshi

Abstract
The last ten years have seen the rapid development of mobile technologies, and there are a number of research studies that are exploring the possibilities of using these technologies to supplement L2 classroom instruction, which are collectively referred to as MALL (Mobile Assisted Language Learning). On the grounds that language use and learning depend heavily on memorized chunks, episodic memory, and retrieval structures, and that post-critical-period L2 learners have difficulty acquiring those memory bases through limited exposure to typical communication, the writer proposes the development of MALL materials that are episodic, live-action, and dialog-style to support such learners.

Introduction
Since the introduction of Microsoft Windows 95 on August 25, 1995, personal computers have drastically transformed our lives. By the mid-2000's, Internet connection speeds had become fast enough to keep us from being frustrated when surfing the Internet as long as the content being viewed or downloaded did not contain heavy files. At present, in Japan and other technologically advanced societies, WiFi has become the norm, and there is even a faster wireless connection called WiMax (Worldwide Interoperability for Microwave Access). While Internet technology has been advancing, mobile technology has also
been evolving at an amazing pace and now we are seeing the proliferation of smartphones.

The IT revolution has also created an exponential increase in the use of computer-related materials in the foreign language classroom. In educational institutions in Japan and many other industrialized countries, digital technology is utilized for all mediums of instruction. Even small-scale schools now have classrooms with projectors. Practitioners who are computer-literate make use of cutting-edge digital technology such as DVD movies, Microsoft Office software, and the vast, free information available on the Internet in their classrooms. It is not as easy as it looks, however, to utilize these technological advancements; there appears to be an emerging gap, a huge one, between computer savvy language instructors and those who are not.

Since the 1960s, under the acronym CALL which represents Computer-Assisted Language Learning, computer technology has been applied to foreign language learning as an ancillary or complementary tool for classroom instruction. The past 15 years of the IT revolution have seen commercial programs such as Rosetta Stone and learning management systems (LMS) such as Moodle have become integral parts of the classroom. The results of studies concerning CALL have not been so promising, however. While positive results have been reported in the CALL literature in terms of developing listening and reading skills, no such results can be found in terms of speaking (with the exception of pronunciation), writing, and real-time oral communication. Currently, CALL research appears to be seeking possibilities around interactiveness with the use of the Internet, one example being the interaction-oriented virtual reality program Second Life. Classroom practitioners are also making use of computer-mediated communication (CMC) tools such as Blogs and Skype to improve the quality of their teaching. While computer-based research and practices have been growing, so has the potential for utilizing mobile devices in language programs.

Given the dramatic progress of mobile computer technologies, this paper will focus on the emerging field of MALL (Mobile-Assisted
One huge advantage of MALL is the increase in continuity and spontaneity (Kukulska-Hulme & Shield, 2008); that is, with mobile devices, learners can learn at their own pace. The history of MALL is, no doubt, much shorter than that of CALL, and studies involving it have just started to be instituted since the iPod and smartphones came onto the market. Thus far, only the overall facilitative learning effects from the use of mobile devices seem to have been studied (Ali & Irvine, 2009).

With the current physical limitations of mobile devices, that is, the difficulty of typing on the tiny device on the move, as well as the facilitative learning effects of utilizing CALL already reported on comprehension skills, the narrower focus of this paper is on the oral communication aspects of language use and learning. This paper will propose developing episodic, live-action, dialog-style materials for mobile learning to help adult English-as-a-foreign-language (EFL) learners develop basic, threshold skills for face-to-face oral communication. The target population is Japanese university students and business persons who struggle to orally communicate in real time despite receiving at least six years of English education before graduating high school.

It has to be kept in mind that it is the EFL learning context that is being discussed. The EFL setting needs to be considered in comparison to an ESL context. The former is a context where the English language is not the means of everyday communication and access to it is quite limited, while the latter, standing for English-as-a-second-language, is another type of context where English is the language used daily and is accessible everywhere. This dichotomy is crucial in creating pedagogical strategies: how to teach and guide a learner are quite different, unless the learning environment is devised in such a way that English is the sole medium of communication.

Fundamental knowledge bases for real-time L2 communication

To start, what linguistic abilities are required for a novice L2 learner to communicate face-to-face and accomplish the minimum com-
municative goals of understanding the interlocutor and getting her/his message(s) across? Among those are memorized lexical chunks that are readily retrievable from the long-term memory for listening comprehension, and a (much) narrower set of memorized chunks for immediate speech production.

How can one communicate with another if she or he cannot comprehend most of or at least part of what her or his interlocutor is talking about? When a Japanese tourist visiting the U.S. orders a lunch set at a McDonald's by pointing to some hamburger set on the menu sheet with the minimal utterance "This, please," what happens if the tourist does not understand the server's response "For here or to go?" The tourist looks puzzled, and the server elaborates, saying "Do you want to eat it in here or take it out?" The tourist may still not be able to understand the server's message. Finally, the server uses his index finger pointing to the floor first and then outside while tilting his head with a questioning face, and at last the tourist understands. This is just one example, and in normal communication between friends, classmates, co-workers, etc., the length and complexity of the communication extends much beyond this. If constant repetitions or elaborations are needed, a breakdown in communication is likely. It is evident that one has to have at least basic listening comprehension skills in order to orally communicate with others.

While the ability to keep up with the fast speech of a native-speaker is vital, an L2 speaker also needs to have the capability to produce language with at least some degree of ease and facility. More often than not, however, many adult Japanese learners struggle to compose English sentences during oral communication. Beginning level L2 communicators simply lack the basic speaking skills necessary to accomplish this. They are slow and make many syntactic and morphological mistakes. Their paralanguage is often unnatural or awkward and their word choices often lead to misunderstandings. It is not easy for interlocutors, especially if not experienced with novice speakers, to follow such a dialog.

How, then, can EFL adult learners develop their listening and
speaking skills for oral communication? This paper argues that such learners will benefit enormously from storing a number of lexical and episodic items in their long-term memory as readily accessible chunks while or even before engaging in a series of authentic oral communication tasks. Before we can put forward a practical proposal for utilizing MALL based on the grounds that memorized chunks play a significant role in language processing, it is essential to outline the mechanisms of human language processing.

Cognitive psychology: an overview of language use and learning

The two systems in the brain responsible for information processing are the working memory and long-term memory, though there is an argument that these two components of memory are part of one large system (Cowan, 1995, 2005). Working memory holds information for a brief period of time. The information briefly kept undergoes cognitive processes such as monitoring and manipulation for reasoning, input comprehension, speech production, and learning. Processes for the manipulation of information include decomposing the incoming data, disregarding part of it in order to focus on other parts of it, retrieving relevant data from the long-term memory, and integrating the data (including the temporarily activated long-term memory).

Thus far, four major components of the working memory have been identified. The main one is the central executive which coordinates cognitive processes (Baddeley & Hitch, 1974). Two other components are the phonological loop and the visuo-spatial sketch pad (Baddeley & Hitch, 1974). The phonological loop maintains the incoming aural data as it is for a brief period of time. It has been argued that individual differences exist in the amount of aural data that individuals can temporarily hold (Skehan, 1998) and presumably in the accuracy of it as well. The visuo-spatial sketch pad, on the other hand, functions as if a camera, but our brain does not process all parts of a given mental image; instead, only focused bits go through cognitive processes. Indi-
Individual differences probably exist in the ability of this particular component, too. The last component recently proposed by Baddeley (Baddeley, 2000) is the episodic buffer. This buffer binds various kinds of information into one united episodic mental representation, which is, to be clear, distinct from episodic memory (see below) already formed in the long-term memory. Types of information to be encoded in the buffer include not only phonological, visual and spatial data but also semantic data, and it is claimed that they are combined into one set of data, an episode, which presumably has the chance of being sustainably stored as is in the long-term memory. It is probably the case that what comes in through the other senses (i.e., touch, taste, and smell) also goes through episodic unification.

As outlined above, the working memory is conceptualized as multi-functional. However, it has one crucial limitation. That is to say, the central executive can monitor and manipulate only a limited number of pieces of information at one time. George Miller (1956) first developed the idea of the magical number seven, stating that the maximum number of pieces of information that a young adult can store at a time in the working memory is around seven. Later research suggested that the number varies depending on the type of information: for words, it is around five. If we phrase this in another way, attentional resources are so limited that when the data to be processed is excessive, the working memory becomes overloaded (cf. Tomlin & Villa, 1994). Now we see one fundamental question. If attentional resources are so limited, how can we explain the significant differences between proficient language users (including bi- or multi-linguals) and poor or mediocre L1 users and unsuccessful L2 learners? The answer lies in the quality of the information stored in the long-term memory.

The capacity and functionality of the long-term memory is enormous. Some of its characteristics are crucial for the discussion of this paper’s theme. First, memory for a word is stored in multiple representations (cf. Bolinger, 1975; Pawley & Syder, 1983; Peters, 1983). For example, in the brain of a native speaker of English, the word ‘school’ is represented in instances like ‘go to school,’ ‘high school,’ ‘at school,’ ‘in
school,' 'business school,' etc. When the language user with these chunks memorized in her or his long-term memory encounters one of these phrases, she or he directly accesses the equivalent chunk memory as a single unit rather than, say, first accessing the memory for 'school', then searching for the memory for the concomitant word(s), and finally understanding the phrase by connecting all of the retrieved memory data. The latter manner would take up far more attentional resources, making it unlikely for the language user to manage language tasks in real time. The same information may be represented in multiple forms by episodic memory (cf. Tulving, 1983) that contains numerous types of information such as phonological, visual, and semantic, which are critical resources for language processing. Also, numerous connections between representations can be formed as retrieval structures, which will help language users process incoming data by allowing them to immediately access the relevant information stored in their long-term memory (Ericsson & Kintsch, 1995; Gobet, 2000).

The most important characteristic of the long-term memory, however, is that there is no limit to the activation of representations, referred to as long-term working memory (Cowan, 1995, 2005) and that when connected to the activated representations and retrieval structures, the data processed in the working memory can be held over a much longer period of time for further cognitive processes. This concept is extremely useful in explaining why some of us can so adeptly deal with long, complex language tasks in real time, while others cannot. That is, the former type of individuals, while suffering the same limitation to attentional resources as everyone else, can manage such tasks better by means of reserving part or most of the relevant data in their working memory as readily reusable references through linking them to long-term working memory without using up attentional resources.

In ending this brief summary of what cognitive psychology tells us about how we handle language tasks, another reason why memorized chunks, episodic memory, and retrieval structures can be important for language learning is offered by the SLA theorist Richard Schmidt.
Schmidt (Schmidt, 1990, 1994, 1995; Schmidt & Frota, 1986) claims that noticing must precede any type of learning. Noticing can be defined as the mental process of rehearsing some detected information from the input in the working memory. While humans can choose to process the incoming data at a deeper level of awareness (i.e., notice things), a vast majority of such data go unnoticed though detected (cf. Tomlin & Villa, 1994). Noticing, in other words, is a cognitive process of arriving at or discovering something in the working memory that operates cognitively at a deeper level than does mere perceptual detection. It then follows that when supported by long-term working memory, the working memory can set aside attentional resources for the task of noticing, which is a precondition for learning.

In conclusion, for language comprehension, output production, and learning, the working memory plays a central, orchestrating role. However, as the number of pieces of information that the working memory can process at a time is so limited, language users need support from the long-term memory, specifically in terms of chunks and long-term working memory, in order to successfully manage language tasks at hand. In essence, the more rigorous the chunk and episodic memory is, the more attentional resources will be saved for cognitive processes (including noticing for learning) and thus more can be managed. It can be concluded then that for successful L2 learning, the L2 learner should engage in building memorized chunks or episodic memory in order to deal with the hurly-burly of language processing in real-time communication. However, the difficulty facing EFL learners is that they seriously lack opportunities to form such chunks or episodic memory. Moreover, with adult learners, there are other special considerations to take into account.

SLA: an overview of adult L2 learners

At least three premises need to be kept in mind for adult L2 learners. First, neuroscientifically speaking, with adults, there is a so-called sensitive or critical period in second language learning (Birdsong, 1999;
Long, 1990). In other words, for adults learning a second language, the qualitatively different predisposition to language learning that young learners possess is no longer available (Bley-Vroman, 1989). After the sensitive or critical period, natural language use will not result in nativelike proficiency. This account of adult SLA has received substantial empirical support (not only for foreign language contexts but also second language contexts). This in turn means that for adult learners, whether in an EFL or ESL setting, neither normal language use nor traditional classroom teaching will result in successful L2 learning.

This reality for adult learners can be explained from another theoretical perspective as well. That is, adult L2 learners are so adept at extracting meaning (far better than children) using the general schematic knowledge they have acquired through L1 learning and life experience that linguistic features with little communicative load are less likely to be processed linguistically (to the extent that they are noticed, or even detected) than loaded ones (Skehan, 1998). Their interlocutors, in most cases other adults, will also comprehend or only guess the gist of non-nativelike output (by means of their general schematic knowledge), which then may serve as positive feedback to L2 learners, thus not giving them enough motivation to improve.

A similar line of explanation has been offered by Bill VanPatten (1990, 1993; Skehan, 1998). VanPatten has argued that meaning takes priority in language processing, and meaning distracts attention from form. According to this account, the priority is toward meaning and not toward form, and thus during real-time language use, L2 learners’ focal attention is unlikely to be directed toward formal linguistic features, especially when they carry little communicative load.

Taken together, a special kind of support, even an unnatural one in terms of communicative authenticity, is called for that would help adult L2 learners develop chunk memory, episodic memory, and retrieval structures. For classroom instruction, some SLA theorists have expressed concern over the inadequacy of purely communication oriented approaches, claiming that not only is focus-on-form (FonFS) a non-ideal approach, but also focus-on-meaning (FonM) has not been particu-
larly successful, and *focus-on-form (FonF)* should be implemented in the L2 classroom (Doughty & Williams, 1998a, 1998b; Long, 1989). Long and Robinson (1998) define focus on form as "an occasional shift of attention to linguistic code features – by the teacher and/or one or more students – triggered by perceived problems with comprehension or production (p. 23)." Doughty and Williams (1998a) distinguish focus-on-form, focus-on-meaning, and focus-on-formS in the following manner: whereas focus on form "entails" a focus on formal elements of language, focus on formS "is limited to" such a focus and focus on meaning "excludes" it (p. 4). While not disregarding the importance and potential of FonF approaches at all, this paper also claims a role for an FonFS approach (cf. DeKeyser, 1998; Matsuzaki, 2011) in supporting adult L2 learners. Now, we turn to the central proposal of this paper — MALL.

**Development of MALL materials as one practical solution**

For MALL, in terms of assisting adult EFL learners in developing threshold oral skills for real-time communication, the preceding accounts on language use and learning can be orchestrated in the following way (see Figure 1 below). Language use is conducted in the working memory, and noticing is a prerequisite for learning. Though multifunctional, working memory is limited in its capacity. While the number of pieces of information that the working memory can store and process at a given moment is so limited, there is no such limitation to so-called long-term working memory, that is, long-term memory in the state of high activation through retrieval. Therefore, what is crucial for successful language use and learning is how much chunk and episode memory a language user has stored in her or his long-term memory. This type of memory contains a substantial amount of information that can be connected, and depending on how complexly her or his retrieval structures have been formed, can be used for immediate activation. While, for first language acquisition, the building of chunk memory, episodic memory (including linguistic forms), and retrieval
structures naturally occurs through one's everyday interactions, that is not the case with adult L2 learning. Adults have far less capacity to pick up language than children, and since they are equipped with general schematic knowledge, adults (both learners and interlocutors) prioritize meaning over form. In EFL settings, the story is worse, as there is the problem of paucity of input and interaction in the L2. Taken together, fundamentally, if adult EFL learners are left alone, there is little chance for them to successfully address the triple task of comprehension, speech production, and learning through the course of natural oral interaction, especially when they are not equipped with a minimal chunk and episodic memory base. Thus, one reasonable solution to the problem adult L2 learners are facing is to increase the chunk and episode memory bases even if it means the utilization of an unnatural, FonFS approach.

![Diagram: Long-Term Memory](image)

**Figure 1** mechanisms of L2 use and learning for adult EFL learners

As one possible course of action for developing the skills needed for face-to-face communication, the writer proposes developing materials to follow and copy that are specially designed for mobile devices. There are some criteria to optimize MALL materials for developing these skills. Above all, the more episodic the materials are, the better; that is, the content should be designed in such a way that learners will store it in the episodic buffer and (through repetition) register it as an episode in the long-term memory which then can be activated in future lan-
guage use and learning. Linguistic candidates for such materials are general phrases and sentences with broad use, ones that are considered to be used in as many communicative contexts as possible. The more general, the better, although technical language can be a candidate also as long as the realm of technicality matches the learner's interests or needs.

Most importantly, the best learning effects will be realized if the content is live-action video. Live-action content provides input to the visuo-spatial sketch pad, stimulating the central executive, and is likely to facilitate the learning of face-to-face communication skills as it will provide models for learners to follow and copy in terms of body movements, gestures, and facial expressions. If the working memory has an episodic buffer, then the importance of live-action content becomes even more essential; that is, episodic memory resulting from unified data in the episodic buffer is an extremely powerful tool which will allow language users to instantly retrieve various kinds of information such as the pronunciation of words, body language, facial expressions, etc., and to execute whichever ones they see are most appropriate for the language processing task at hand, without draining attentional resources. Lastly, in relation to live-action, the style of discourse should be primarily dialog-style, as most communication takes the form of dialog. With dialog-style materials, the actual interactive cues, including paralanguage, between two speakers can also be learned. The properties and their expected learning effects are summarized in Figure 2 below.

<table>
<thead>
<tr>
<th>properties</th>
<th>effect(s) for learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>mobile</td>
<td>facilitates (repetitive) use and chunk learning</td>
</tr>
<tr>
<td>episodic</td>
<td>stimulates episodic buffer and facilitates episode learning</td>
</tr>
<tr>
<td>live-action</td>
<td>stimulates episodic buffer and facilitates episode learning</td>
</tr>
<tr>
<td>dialog-style</td>
<td>provides models to follow for actual communication</td>
</tr>
<tr>
<td>general use</td>
<td>increases the chance of chunks learned being applied to actual use</td>
</tr>
</tbody>
</table>

Figure 2  A breakdown of the MALL proposal for adult EFL learners
Conclusion

While there will be educators who question the idea of engaging learners in rather unnatural memorization processes, the theoretical accounts above certainly support the practical advantages of implementing this pedagogical strategy, especially in EFL settings with adult learners. The writer is currently embarking on an experimental study in which the kind of MALL content described above is developed and will be tested in terms of its learning effects on a number of subjects over a period of time. It should be noted, however, that developing instructional oriented MALL materials is not easy. In Japan, not enough funding is available for developing such materials for academic research purposes, a challenge for MALL researchers working in Japan.

There is also an issue of how closely learners are encouraged to copy models by native speakers. Which accent should they follow? Should they be encouraged to mimic what some Japanese speakers might consider as exaggerated body languages? These are questions that touch on the issue of national or cultural identity. While acknowledging the existence and importance of such an issue, the writer still recommends copying and memorizing native speakers' models. Learners can override whatever body language or accents they see as inappropriate later on, with whatever ones they regard appropriate. Such further learning processes or fine tuning may be possible only when learners can appreciate such delicacies during ongoing communication, which, in fact, is only possible through the use of chunk and episodic memory.

As a concluding remark, there is no doubt about the significance of engaging in actual, authentic communicative activities in order to fully develop one's oral communication skills. However, successful learning may take place only when such communicative moments are readily available to L2 learners. The writer does question their availability in EFL settings, at least in Japan. The central government or the educational institute a learner belongs to might attempt to rather deliberately
create such communicative arenas, but such an attempt is no easy thing. While this paper does not take any stance as to whether Japan should raise the priority for enriching communicative conditions and spend the necessary funds for them, it does, as an interim solution, suggest developing mobile content that can be relatively easily accessed for L2 learners.

Note
1 The length of each dialog may be an important consideration, too. Given the short amount of time a learner would spend at a given time on the content with their mobile device on the move, I feel that one string of content should be short enough for the learner to choose to repeat for (complete) copying.

References
DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learn-


University Press.


（まつざき・たけし 政治経済学部特任准教授）