Language Teaching Research Quarterly

2019, Vol.12, 16–36

# Listening to Teachers' Voices: Evaluating Language Learning Apps with a Task-based Language Teaching Framework

Joshua Matthews[1]*, Rachel Burke[2]

[1]University of New England, Australia

[2]University of Newcastle, Australia

***Received*** *08 May 2019* ***Accepted*** *18 July 2019*

**Abstract**

Language learning applications (LL Apps), especially those that can be downloaded to a mobile device, are popular, widely available and represent a potentially useful tool for language teachers and their students. However, concerns exist that many LL Apps are inadequately aligned with second language acquisition (SLA) theory, and/or poorly suited to the expectations and classroom requirements of language teachers. The current study examines these concerns by providing six language teachers with opportunities to use two popular LL Apps over the period of six weeks. During and after this period, qualitative data were gathered from the teachers by way of questionnaires and interviews. Teachers' elaborations of the perceived affordances of the LL Apps for their language teaching were then coded against the methodological principles of a Task-based Language Teaching (TBLT) framework (Doughty & Long, 2003). Analysis indicated that although some methodological principles central to TBLT were evident in the affordances offered by the LL Apps, many were not, or only minimally so. Concrete suggestions about how to more strongly align LL App design with the principles of TBLT are provided, and the important role of the teacher in the LL Apps design process is emphasized.

**Keywords:** *Mobile Assisted Language Learning, Language Learning Applications, Task-based Language Teaching, Language teachers*

## Introduction

The ubiquitous nature and complex functionality of mobile devices have positioned them as potentially valuable tools for second language learning. The targeted use of mobile devices and mobile applications in the pursuit of language learning outcomes, so called Mobile Assisted

jmatth28@une.edu.au

Language Learning (MALL), is a burgeoning field of research (Burston, 2015). A mobile application or *app* is software that has been designed to run on a mobile device, and has scope to apply the range of functionalities offered by the device. Modern mobile technology is programmable and adaptable and can deliver data in multiple modalities (audio, text, & video). Mobile devices also have sophisticated touch screen interfaces, rapid internet connectivity, and can facilitate various forms of interpersonal communication (text chat, live audio & video chat, etc.). Furthermore, there is a large and ever expanding number of language learning applications (LL Apps) that can be easily downloaded (often for free), installed, and readily used. Taken at face value, the potential of harnessing and delivering mobile technology's complex functions for the purposes of language learning by way of LL Apps is an exciting prospect. Despite this apparent potential, however, the reality of LL Appuse in second language instruction has been largely disappointing (Reinders & Pegrum, 2015), and their use remains on the periphery of foreign language teaching (Burston, 2014a). As will be detailed below, in explaining this lack of integration, our broad thesis has two parts: a) that LL App design is not adequately underpinned by second language acquisition (SLA) theory, and b) that the affordances of LL Apps are typically poorly aligned with the requirements of language teachers.

**Missing Links in LL App Design: SLA Theory and the Teacher's Voice**

A recurring criticism of LL Apps is that their design often lacks an explicit foundation in SLA theory (Burston, 2014b; Kukulska-Hume & Shield, 2008; Reindeers & Pegrum, 2015). Aligning LL App design with principles emergent from SLA theory is especially important because as Levy (2015) points out "…learners are increasingly utilizing their own powerful, personal technologies in their language learning independently outside of class without a teacher present" (p. 557). If the design of LL Apps is based on imperatives other than sound pedagogical principles (e.g., convenience in program coding), then this is a serious concern, as such LL Apps may well contribute to the perpetuation of ineffective language learning practices. Research addressing LL Apps adds weight to these concerns, with many shown to possess designs that facilitate a behaviourist type transmission model of instruction, with a predominant focus on quizzes targeting vocabulary and grammar (Burston, 2014a, 2014b). This unsatisfactory situation would appear to be largely avoidable, as contemporary mobile devices do have the functional capacities to facilitate language instruction that is well aligned with the broadly accepted principles of SLA that have emerged from research over the past 40 years (Burston, 2014b; Kukulska-Hume & Shield, 2008). Clearly, pushing LL App design to more strongly align with SLA theory is an important future challenge. Here, we assert that a useful response to this challenge is the systematic assessment of LL Apps with evaluative frameworks (Reinders & Pegrum, 2015; Rosell-Aguilar, 2017), and application of these assessments in an effort to guide future design directions.

Other than a noted lack of SLA theory underpinning LL App design, an additional concern is that LL App design often takes place in the absence of substantive consultation with teachers. This is a significant issue, not only because language teachers have a central role in facilitating

the use of technology among their students (Ushioda, 2013), but also because, in our experience, teachers often possess a wealth of firsthand experience on how to do so effectively. In planning for daily teaching duties, active teachers are engaged in a continuous iterative cycle of lesson design, and redesign. This iterative process is acutely attuned to *what* and *who* needs to be taught, and is fundamentally based on the teacher's perception of what '*works'*. This contextualised knowledge means that teachers are uniquely positioned to assess the affordances of tools such as LL Apps, and thus have a pivotal role in integrating appropriate technology into the language classroom (Guichon & Hauck, 2011).

Here, we follow Haines (2015) in defining technological affordance as "the potential that teachers perceive in a particular technology tool that will support learning and teaching activities in their educational contexts" (p. 166). Thus, although the affordance of a given LL App is necessarily linked to its design features and resultant functions, the teacher's perception about how any given LL Apps may be usefully applied in their language classroom is absolutely crucial. For this reason, the involvement of language teachers in the evaluation process, and ultimately the design, of MALL technology is important for its successful implementation. However, as language learning technology continues to rapidly develop and change, research examining the interface between pedagogy and new technology inevitably lags behind (Guichon & Hauck, 2011). For this reason, more research that taps into practicing language teachers' perceptions of the affordances of technology, such as LL Apps is warranted. Further, analysing teachers' perceptions against pedagogical frameworks of best practice represents a potentially useful way to synthesise the missing links of LL App design (viz., SLA theory & teachers' voices), and may provide a body of knowledge that can be usefully applied in future LL App design.

**A Task-based Language Teaching Framework**

Although a universally accepted account of optimal language instruction remains elusive (Ellis, 2005, 2006), Task-based Language Teaching (TBLT) is a well-established, empirically supported movement (Ellis, 2003, Nunan, 2004; Skehan, 1996) that continues to attract attention (Ellis, 2010, 2017). In the context of TBLT, Ellis defines a *task* as a language teaching activity that places a primary focus on meaning, and requires the language learner to comprehend and convey communicative messages in the target language. Ellis further asserts that a task entails a language learner applying their own linguistic and non-linguistic resources to negotiate an information gap in order to achieve a communicative outcome (TESOLacademic, 2014). Although, the principles of TBLT have been clearly articulated in the academic domain for some time (Long, 2016; Ellis, 2003, 2017), various alternative models of TBLT exist (Ellis 2017). However, Doughty and Long's (2003) framework for TBLT is particularly relevant here as it explicitly addresses the realisation of TBLT in computer-mediated instruction. The framework provides ten methodological principles of TBLT, which are presented under the categories of *activities*, *input*, *learning processes*, and *learners*. The principles outlined by Doughty and Long

(2013) are overviewed below as these provide the analytical framework used for the current study.

**Activities**

*Using tasks as the unit of analysis (principle 1)* discourages teaching activities and syllabus design which are primarily focussed upon decontextualized linguistic elements, such as vocabulary items or grammatical structures. Holding the task as the unit of analysis encourages a strong move away from so-called synthetic syllabuses that deconstruct the linguistic elements of the language and present them in sequence (Long & Crookes, 1992). At the level of the learner, the focus is not on the metalinguistic analysis of text, but rather on the learner's own formulation of socially-contextualized language artefacts; that is, text as evidence of task completion.

*Emphasising learning by doing (principle 2)* aims to assist learners to meet their present or future real-world communicative needs by way of encouraging relatively authentic hands-on, problem-solving activities. It is important to apply graded tasks set at appropriate levels of complexity, as task complexity can have a measurable impact on communicative outcomes associated with completing the task (Robinson, 1995). Simulations may provide a useful context within which learning by doing can be undertaken.

**Input**

*The use of elaborated input (principle 3)* is evident in teaching when, while engaged in the use of language to undertake a task, modifications such as repetition, comprehension checks, lexical switching, use of redundancy, and other scaffolding occur. A key objective of elaborated input is to enable the meaning-focused use of authentic texts, and effectively negate the need to provide excessive metalinguistic explanation. A key objective is to avoid the use of simplified texts which are devoid of the language features with which language learners are required to become proficient. Here we define an authentic text as one that has been "created to fulfil some social purpose in the language community in which it was produced" (Little, Devitt, & Singleton, 1988, p. 27).

**Learning Processes**

*Providing rich input (principle 4)* involves affording the learner with a diversity of text types, and further involves ensuring input has not been impoverished in terms of vocabulary, grammatical complexity, sentence length and so forth. Additionally, it involves the provision of task-specific and domain-specific language typical of that used by native speakers when engaged in task completion. Generally, the provision of rich input involves ensuring learners complete tasks by way of engaging with a broad range of texts which are aligned with the language learning needs of the learners (Doughty & Long, 2003, p. 61).

*Emphasising inductive learning (principle 5)* involves encouraging opportunities to develop implicit stores of knowledge. Rather than studying language as an object through explicit analysis of language samples, opportunities are provided for learners to gain knowledge of the

underlying structures of the language (grammatical form, collocations, idiomatic language, etc.) through induction. In a practical sense, this involves the engagement with rich and elaborated input as components of pedagogic tasks which encourage the learner to store and retrieve language in larger units ("chunks of meaning").

*Facilitating focus on form (principle 6)* provides an opportunity for the TBLT model to incorporate explicit teaching of linguistic forms, while at the same time maintaining the task as the unit of analysis of the instructional program. In contrast to so-called *focus on forms*, which involves learners spending a significant proportion of their time engaged in metalinguistic analysis of linguistic features, *focus on form* involves learners' attention only being strategically directed to linguistic items to induce "noticing" while being predominantlyengaged in meaning-focused interactive tasks. Such instances of incidental focus on form are associated with positive second language learning outcomes (Loewen, 2005).

*Providing negative feedback (principle 7)* entails the provision of explicit (e.g., metalinguistic explanation) or implicit (e.g., corrective recast) negative feedback in response to the production of an erroneous linguistic form (Ellis, Loewen, & Erlam, 2006). The provision of feedback is associated with learning of linguistic forms (Mackey, 2006), and can be effectively facilitated with computers (Matthews, Cheng, & O'Toole, 2015). Of additional assumed importance is a relatively short time lag between the triggering event and its provision of feedback (Ellis, 2017).

*Respecting developmental processes (principle 8)* involves an understanding that some stages of the language learning process are immutable in regards to the chronological sequence of their learnability (Ellis, 1989). The learning objectives of the tasks which the learners are asked to complete, need to be formulated by way of ongoing analysis of the current processing capacity of the learners. Understanding the learner's current developmental stage will also inform the need to intervene in relation to the provision of feedback and explicit focus on form.

*Promoting collaborative learning (principle 9)* involves facilitating discourse between learners as they collaboratively and cooperatively complete clearly defined tasks. Such interaction among groups of language learners facilitates opportunities for the provision of input, output and feedback (Pica, Lincoln-Porter, Paninos, & Linnell, 1996) each important elements associated with productive language learning contexts (Gass & Mackey, 2006).

**Learners**

*Individualizing instruction (principle 10)* involves ensuring learning is tailored to individual differences (e.g., different goals, motivation, cognitive style, and learning strategies), and this can be achieved by way of needs analyses (Long, 2016). The course designer directs the underlying learning plan, but provides individualized options (especially relating to content) within that plan which cater to individual difference. The principle of individualization can be facilitated by the affordances of technology and adaptive design (Doughty & Long, 2003; Matthews, O'Toole, & Chen, 2017).

**The Current Study**

The body of research addressing the nexus between TBLT and technology is emerging (Ellis, 2010; Thomas & Reinders, 2010), but with the ongoing release of LL Apps, and the noted absence of an SLA theoretical foundation for many of these, more research in this space is required. As Burston (2014b) asserts, "The future of MALL lies in the exploitation of the communication and multimedia affordances of mobile devices in ways that support collaborative, task-based learning both within and outside of the classroom" (p. 344), and such conviction provides impetus for the current study. This study seeks to tap into the perceptions of language teachers in relation to the affordances offered by two popular LL Apps. Further we seek to interrogate the degree to which the content of teachers' elaborations of these perceived affordances can be aligned with the methodological principles of Doughty and Long's (2003) TBLT framework. A final objective is to draw on a synthesis of these findings to offer concrete suggestions about the future direction of LL App design.

## Methods

### Participants and their Context

Participants were recruited through alocal language teachers' association. Inclusion criteria included a willingness to engage for six weeks with language learning technology, the possession of a formal language teaching qualification, and firsthand experience in language teaching. Six teachers in total were involved in the study. Each was working as a Japanese language teacher in different secondary high schools in the state of New South Wales, Australia. All six were native English speakers and had a range of teaching experience from 1 to 23 years ($M$ = 10.33). Table 1 presents general information about the teachers, their assigned pseudonyms, and demonstrates how they use technology in their own lives as well as in their language learning classrooms. Table 1 also shows the teachers' responses to statements about technology asked of them at the point of recruitment, and indicates a positive attitude towards technology use in language teaching.

Initial interviews also provide insight into the teachers' attitudes toward technology use in the language learning classroom. For example, Faye's comments present a clear position:

*I think there's no use denying that technology is there, we need to look at how we can benefit our classrooms instead because I think, more than anything, language teachers can benefit from technology.*

Similarly, when Deena was asked if she had any general concerns about technology in the language learning classroom, her response was decisive:

*I don't have any concerns at all. I think we should encourage [students] to explore it in whatever way we can*

Table 1
*General information about the participating teachers*

| Participant teacher | Anne | Bella | Carol | Deena | Ellen | Faye |
|---|---|---|---|---|---|---|
| Number of years teaching experience | 10 | 1 | 1 | 23 | 2 | 7 |
| Do you use technology (e.g. mobile phones, laptop, etc.) on a daily basis? | Yes | Yes | Yes | Yes | Yes | Yes |
| Do you use technology in the language classroom? | Yes | Yes | Yes | Yes | Yes | Yes |
| Incorporating technology into the language classroom increases learner interest in the lesson | Strongly agree | Strongly agree | Strongly agree | Strongly agree | Agree | Strongly agree |
| Technology can be used to differentiate language learning activities and help learners work at their own pace and level | Agree | Strongly agree | Agree | Strongly agree | Agree | Agree |
| Using technology in the language classroom allows for more learner collaboration | Agree | Strongly agree | Strongly agree | Strongly agree | Agree | Agree |
| Technology can be used to provide additional scaffolding and support for learners | Strongly agree | Strongly agree | Strongly agree | Agree | Strongly agree | Strongly agree |

Note: pseudonyms used

Teachers also provided insight into their students' attitudes towards technology in the language classroom. Ellen notes that her students seem to engage more in class when using technologies such as LL Apps:

*I've found when I can get them on applications, they do pick things up a lot faster because it's just the way that they like to engage. It's just giving them that access and getting them to … overcoming those barriers*

Faye also overviews her students' enthusiasm for technology in the classroom:

*I don't think I'd be able to think of a student over the last three or four years who hasn't been a hundred percent on-board with the technology, which is surprising. So, yeah, I have a hundred percent buy-in from [the students]*

In sum, the teachers involved in this study were practising, 'techno-competent' language educators, who had a positive disposition towards the use of technology in the language classroom.

**Mobile Hardware and LL Apps**

A standard high-end mobile device was provided to each of the participants. These were WiFi compatible, touchscreen (9.7 inch) tablets, with 32 GB memory. Each mobile device was pre-loaded with two commercially available language learning applications *(App 1 and App 2).*

The two applications were selected based on their popularity (measured in estimated number of registered users), accessibility, and degree of alignment with general principles of language teaching. The company that owns the first language learning application featured in this study, estimates that as of January 2019, there are some 90 million registered users. The program is described by the designer as intended to provide interval-based language practice, in which the users work through a pre-set, topic-based sequence of written and spoken input based on their

level of proficiency. While there is no capacity for users to practice their pronunciation when using application 1, there is a built-in interactive component in which users can invite others to correct their written output.

The second language learning application follows a similarly topic-based structure as application 1 and is also described by the designers as being intended for short, frequent revision and application of language content. In addition to the focus on reading, writing, and listening, users are able to benefit from in-built voice recognition technology to practice their pronunciation. There is no interactive component to application 2, however in 2018 there were an estimated 300 million registered users.

The features and affordances of these two applications are presented in the following table, which is based on a similar overview used in Rosell-Aguilar (2017):

Table 2
*General information about the participants in the group*

| Features and Affordances | App 1 | App 2 |
|---|---|---|
| Grammar (focus on the structures of language) | ✓ | ✓ |
| Vocabulary (introduction to content and grammatical terms) | ✓ | ✓ |
| Culture (cultural information regarding the language and its speakers) | X | X |
| Interaction (capacity to engage with other users) | ✓ | X |
| Phonology (capacity to practice pronunciation using voice recognition technology) | X | ✓ |
| Listening (receptive audio capabilities) | ✓ | ✓ |
| Writing (capacity to compose written text beyond the sentence level) | ✓ | X |
| Script (capacity to practice stroke order, character or letter formation) | X | X |
| Reading (presentation of text beyond the sentence level) | ✓ | ✓ |
| Chat (capacity to use video chat, or voice over the internet features) | X | X |

## Instruments

Two instruments were used in this research: an online questionnaire (Appendix 1), and a semi-structured interview (Appendix 2). The online questionnaire presented a range of questions and statements that teachers had the opportunity to respond to in written form after use of each of the two LL Apps used as part of this study. The statements were presented in banks framed against the ten methodological principles of TBLT overviewed above. It is important to note that the details of the TBLT framework were not explicitly mentioned in the questionnaire.

The second instrument used was a semi-structured interview. The interview was designed to tap participant teachers' views about each of the individual apps, as well as their views about possible improvements that could be made to the LL Apps based on their first-hand user experience. The semi-structured interview aimed to gather data that added richness to the data gathered with the online questionnaire.

## Procedures

After receiving participant information and signing consent forms, participants were provided with the mobile device upon which the two LL Apps at the centre of the study had been installed.

As all participants were teachers of Japanese, each LL App was set to Japanese as the language being taught. The participants were asked to personally use the LL Apps as much as was practical within the six week period allocated for the study. It was explained to the teachers that the main objective of the study was to capture their views on the usefulness of the LL Apps in their own language teaching contexts. The participants were free to use the LL Apps in any way they wished, but were asked to ensure that they had engaged substantively with the first LL App within the first three weeks, and engaged substantively with the second LL App before the end of the sixth. During the period that teachers engaged with the LL Apps, the researchers made themselves available to answer any questions and troubleshoot any technical issues that were reported by the participants. After approximately three and six weeks of engagement, a questionnaire was sent to the teachers. The two questionnaires were identical other than each was targeted at exploring the teachers' experiences with LL App 1 and 2 respectively. Texts provided by the participants in the questionnaires were collated.

At the conclusion of the engagement period, as per the details in the participant information documentation, a telephone interview was undertaken with each participant. The format of the interview followed a semi-structured form based on the guiding questions presented (see Appendix 2). Each interview was recorded, with both written and contemporaneous spoken consent provided by each participant. Interviews were conducted individually with participant teachers and were approximately 40 minutes in duration. Recorded sound files were transcribed, and yielded approximately 36,200 words in total, with an average of approximately 6000 words per participant (*Min* = 4200 words & *Max* = 8000 words).

Texts sourced from the questionnaires and interviews were collated and formed the primary data set for content analysis. Recurring lexical items and phrases were systematically identified to establish the presence and frequency of key themes within the transcripts (Berg, 2007). In keeping with our intention to consider the synergies between MALL applications and the TBLT framework, we undertook a more deductive approach to content analysis. This method allowed for detailed investigation of the research objectives and consideration of the teachers' appraisals of the applications according to the principles of TBLT. However, we augmented this deductive approach with the inductive analysis of any emergent themes that did not fit with the expected categories. This was a reiterative process of increasingly refined analysis until all data was categorised according to the final key themes. In the concluding stage, we reviewed the definition and naming of key themes, and then crosschecked for consistency of categorisation.

As an additional point, it is important to make clear that throughout the research period, TBLT was not explicitly addressed with the participant teachers. This approach was intentional, as our primary goal was to facilitate teacher engagement with the LL Apps, from which we sought to draw out candid perceptions of their affordances. We did not wish to potentially influence or confound those perceptions by explicitly introducing the TBLT framework to the teachers. It was only after the data had been gathered that the researchers themselves applied the TBLT principles as a framework for analysis.

## Results

Firstly, to provide a broad overview, the frequency of references mapped to each TBLT principle is provided in Table 3.

Table 3
*Overview of content mapped against the principles of TBLT*

| Principle | No. of references mapped to principle |
|---|---|
| Individualizing instruction (principle 10) | 51 |
| Respecting developmental processes (principle 8) | 31 |
| The use of elaborated input (principle 3) | 20 |
| Providing rich input (principle 4) | 20 |
| Promoting collaborative learning (principle 9) | 12 |
| Facilitating focus on form (principle 6) | 11 |
| Emphasising learning by doing (principle 2) | 10 |
| Using tasks as the unit of analysis (principle 1) | 9 |
| Providing negative feedback (principle 7) | 9 |
| Emphasising inductive learning (principle 5) | 1 |

We next seek to delve deeper into the sentiments behind, and details around, the teachers' perception of the LL Apps. An analysis of these in descending order of frequency, along with vignettes and elaboration is provided below.

### Individualizing Instruction (Principle 10) and Respecting Developmental Processes (Principle 8)

As many of the same references were coded to both principles 8 and 10, these are considered together. Firstly, a recurring theme evident in the teachers' discourse in relation to individualisation was a lack of measured progression in the sequencing of complexity of the learning activities of the LL Apps. Almost all teachers described an initial level of complexity that they viewed as appropriate for their students, but then noted an excessive 'jump' in complexity:

Carol: *... all of a sudden it just jumped to, like, a really difficult kind of content*
Anne:*... started off reasonable, but then jumped. I found that one jumped too quickly*
Bella: *… But I did find them a little bit harder than the lesson that you'd just done. So it was quite a big jump up*

Another recurring theme was a perceived lack of alignment between the contexts of language learning the LL Apps presented and those contexts the teachers assumed their students would enjoy engaging with. Deena notes this lack of alignment:

*I know that they're targeted at different … like, for all age groups, but I'd like to see something targeted at teenagers. … like, we get the same standard do this, do that but, you know, there's nothing for those boys out there that love their surfing or they love their cricket, sports, so nothing that's meaty enough to engage them that way. So I would love to see a bit more of that*

*... I just think the scenarios for teenagers might be a little different to the ones that are targeting adults and any language learner.*

Ellen also notes the LL Apps lack the affordance of individualisation and offers a suggestion approximating that which could be viewed as a needs analysis:

*But even if when they first enrol in the* [the LL App] *or something there's a survey, I guess, of what are your interests. They select those and, through an algorithm or something, it would generate examples using those interests*

Faye also notes the future importance of some kind of needs analysis as part of the design imperatives of a LL App:

*… students' input into what they're learning, like that survey aspect of it, okay, I'm interested in sports, I'm interested in, I don't know, martial arts and then to try and get those kids that would normally disengage because every sentence is just I like swimming - but I don't like swimming*

Participants noted that one of the LL Apps did have a placement test at the beginning of the learning cycle, which could advance students to higher levels of complexity in the learning sequence; however, this did not negate the predetermined sequencing and content, but rather just allowed users to move further along this sequence.

Another noted theme was the desire for the LL Apps to enable the teacher themselves to control and attune the instruction for their students. This was not only seemingly for the purposes of respecting the 'learner's developmental syllabus', but also, importantly, in order to reflect the requirements of the national curriculum. Faye notes that this lack of control posed problems both in terms of the learners' developmental processes, and in terms of integration into mandated programs:

*… because it doesn't follow … like, I couldn't pick and choose what got taught first, second, it kind of generated its own order … And I just find it difficult to align that with what our syllabus says, with what we're teaching and the order that we* [teach] *it in. It just … it pushes the kids either too far ahead or teaches them something completely off the track that doesn't end up being useful to their study at school*

**The Use of Elaborated Input (Principle 3)**

Teachers noted what they referred to as a lack of authenticity in the LL Apps' instructional material. It was apparent that the new version of the national curriculum that the teachers were working with demanded an emphasis on authentic texts and communication. Faye notes that the LL Apps did not cater to this requirement:

*… so we have the new syllabus which is really pushing for a lot of authentic communication and the way that they present that information in those apps, I find, is not a particular authentic way of teaching it, the chunks of things that we're teaching … And I found this in each of them actually, the chunks of skills that they're teaching aren't authentic skills, they're not skills that students are going to use right away either*

Further, there were repeated comments from teachers to the effect that there was 'not enough scaffolding' and several remarks spoke to a lack of elaboration around the input that was

available through the LL Apps. The comments below present a picture of the learners being asked to inductively engage with input with a decided lack of scaffolded elaboration which rendered the input incomprehensible for their proficiency level:

*Ellen: The app throws the learner into a large body of text with no explanation of how to approach the task … with minimal support and no contextualising beforehand*

Bella describes a dialogue that she perceived as being too difficult for her students as it lacked adequate levels of elaboration and contextualisation:

*Bella: But there's no way most kids would be able to access that.*

*Interviewer: Too much of a stretch.*

*Bella: Without too much help. Some of them would definitely ... they need to be simplified or you'd have to do a lot of teaching, pre-teaching for it*

Faye notes that the mode of instruction of the LL App ran counter to her own teaching approach:

*…it was introducing language without explaining it either. That kind of goes against the way I teach, which is trying to give context to the new information that students are getting instead of overloading them with a whole heap of new skills and sentence structures …*

**Providing Rich Input (Principle 4)**

Teachers such as Faye acknowledged the value of the multimodal input delivered by the LL Apps:

*I also liked the fact that there were the different aspects of things, so you could listen to it, you could read it*

But, then also goes on to point out that the LL Apps offered rather impoverished input that lacked real-world applicability:

*Faye: I would love the inclusion of authentic texts for students to read … So it's not just the computer-generated sentence that, kind of, they're never going to be using in real life*

Deena also remarked that there is insufficient richness in the listening material provided:

*And one of the big things I find that we're really missing is a lot of listening skills variety … it's hard to find ones that are a school level*

**Promoting Collaborative Learning (Principle 9)**

Teachers noted the affordance of interaction, particularly in App 1, was evident, but also noted that the activities facilitated by the LL Apps were predominantly individual. As Deena notes:

*… it does offer interaction with others to a degree which is good but is mostly individual*

Faye alludes to the absence of sufficient opportunities to facilitate interaction through use of the LL Apps, and further highlights the discrepancy between the types of interaction targeted by the teachers in class and those which are facilitated through the LL Apps:

*So the sort of things that we're asking students to talk about and the authentic interactions that they'd be having are probably not the same as what our apps are going for. For example, penpals relationships and things like that, which is probably how our students are going to get*

*their first authentic interactions with Japanese students, those sorts of skills were things that they'd be discussing in those penpals or those Skype sessions are not things that are covered by those apps. So, no, I don't even think the content really matches up*

Although the opportunities for collaboration were very limited, Carol commented that the opportunity to send a text sample written in LL App 1, which was then responded to by another user was favourable:

*... I don't know, the only thing I really liked about it was the fact that you could talk to, like, another person, you can write something or whatever and then they can kind of correct it for you*

However, the ad hoc nature of interaction with other LL App users, even after a request for interaction had been initiated, was a frustration for Bella and potentially also for her students:

*I am still waiting for a response from other learners on my activities … there'd need to be assurances [students are] going to get an answer, because if a kid puts their stuff out there to be corrected and is expecting someone to come back, and then they don't get feedback so it's a disappointment. They get very frustrated*

Obviously, this type of lag in a response speaks to the limited capacity of the LL Apps to facilitate real-time collaboration between groups of language learners. Bella goes on to advocate for the expanded capacity to collaborate through the affordances of the LL Apps. In particular, by way of formalising a collaborative network that could be accessed through the affordances of LL App 1:

*I was thinking, like, could you ... if you linked it, like, with a sister school where the students actually are going to be personally motivated to answer each other, that would work. Yeah, like, if you could actually say, "Well, actually, I'm doing this, I'm interested in this aspect. Do you want to co-work together on something?" That would be useful, yeah*

It was surprising to note that there was no mention of the affordances of instant messaging, video chat, or voice over the internet associated with these LL Apps, seemingly an opportunity lost in the design of the LL Apps being investigated. It was noted that discussion forums were available in association with quiz questions, but these were almost exclusively written in English (despite being a Japanese learning environment), and were used to primarily engage in metalinguistic analysis of quiz answers.

It is important to note that the teachers did not mention the LL App offering the affordance to speak or communicate directly to other language learners. This would again appear to be an opportunity missed in the LL Apps' design as verbal and written communicative discourse naturally facilitates opportunities for the provision of, and engagement with, rich and elaborated input.

**Facilitating Focus on Form (Principle 6)**

One of the major concerns about the capacity of LL Apps to facilitate focus on form, in the manner that it is described in TBLT frameworks, was that the LL Apps fundamentally focussed on the quiz-heavy behaviourist type transmission model as Burston (2014a, 2014b) has previously described. The LL Apps seemed more congruent with the *focus on forms* model.

Teachers' responses to the question 'What does the App ask the learner to do?' are revealing in this regard:

*Carol: It goes through some vocabulary drills, tests your knowledge on the words, puts words into dialogue for examples and allows a small amount of practice in using the language*

*Faye: Learn new words through listening, listening to dialogue and recognise where new language fits in cloze passages*

The desire for a more meaning-based set of contexts within which focus on form could be achieved was verbalised, for example, by Faye:

*I'd love it if, you know, you had units that were designed based on specific language skills … So, for example, it might be you want to teach about verbs but you want it to be in the context of getting around town or giving directions, something like that, or verbs but in a classroom context. So you could both pick the skill but also pick the context that they're learning that skill in …*

**Using Tasks as the Unit of Analysis (Principle 1) and Emphasising Learning by Doing (Principle 2)**

As almost all of the same references were coded to both principles 1 and 2, these are considered together. As with the previous sub-section, it must be noted that the LL Apps appeared fundamentally not to maintain the task as the central unit of instruction. A majority of the instructional approaches described by the teachers suggested that the LL Apps did not revolve around meaning-focussed, achievement of tasks. Ellen's assessment of one of the LL Apps was telling in this regard:

*I feel this application uses rote learning in an online format. It then goes on to test this knowledge immediately after teaching. While this would be good to maybe brush up skills, I feel most learners would get bored having to go back and do the same activity in order to memorise content*

Participants also acknowledged that there was little emphasis on problem solving, as Carol noted:

*There are no problems to solve. It simply tests vocabulary knowledge*

**Providing Negative Feedback (Principle 7)**

It appeared that the provision of automated feedback was something that the LL Apps did quite well; for example, when an incorrect or correct option was chosen the LL App provided an automated response. The automated feedback element of the LL Apps was viewed as positive by the teachers:

*Ellen: Good amount of repetition as well as the in-built function to correct the student*

*Deena: I liked … I just liked the activities that it had, the feedback that it had*

Although, feedback in its automated form worked well, and was well received by teachers, forms of feedback that were potentially more amenable to communicative, meaning focussed task completion seem to be problematic. As mentioned previously, the feedback, or lack of

feedback, on spoken utterances submitted to the LL Apps was often problematic, as Bella describes:

*I used the record voice function but am yet to get any feedback. I suspect that this may not occur a great deal as it relies on the generosity of other students. … I would love to have seen the conversation feedback thing really work, but I didn't really get anything …*

In short, when it came to feedback that lent itself most strongly to communicative purpose and task completion, both applications were found to be lacking.

**Emphasising Inductive Learning (Principle 5)**

Only one sample of teacher discourse was mapped against principle 5. This may reflect the teachers' perceived relevance of inductive approaches to language learning for their instructional contexts. Deena describes how answers to matching activities were provided to LL App users far too explicitly.

*Some of the things I found, though … that frustrated me, and I don't know if it's because I'm an adult or whether the kids would find it comforting, when you had to match pictures to words, they'd have a picture of … like, they'd have the word at the top and then they'd have the word in the language … like, sometimes it gave the answer away too easily*

By far the majority of the instruction evident in the teachers' discourse describes explicit analysis and focus upon discrete linguistic elements. There was no clear evidence of a focus on formulaic language, collocations or other instructional approaches such as meaning focused extensive reading and listening, that are likely to build stores of implicit knowledge and that can facilitate spontaneous language use.

**Discussion**

An overarching finding was that many of the instructional activities delivered by the LL Apps appeared to be fundamentally of the behaviourist type transmission model of instruction as surmised by Burston (2014a, 2014b). It was evident from the teachers' descriptions of the applications that the instruction predominantly emphasised metalinguistic analysis and decontextualized focus on forms, and this runs counter to the basic premise of TBLT. This realisation should encourage LL App designers to move away from instructional design that focuses predominantly on linguistic structures, and rather emphasise tasks that demand learners to use language to negotiate meaning and achieve communicative goals. An important initial stage in this process is for LL App designers and teachers to envisage the broad array of real world tasks that LL Appusers will find relevant. Fundamental to this process would be to conduct extensive surveys (needs analyses) with key stake holders to formulate a list of useful task types. The challenging next step would be to conceive of ways that these tasks could be contextualised and realised with the affordances of mobile devices.

Teachers were concerned that the instructional contexts of the LL Apps were poorly aligned with their students' interests. This speaks again to the importance of an in-depth needs analysis in the design of effective LL Apps. Providing learners with the opportunity to complete a task

centred on asking for directions, for example, could just as well entail directions to the university, the martial arts studio, the museum, the beach, the temple, and so on. Ideally learners would have the option to select themes or contexts of interest, and also the particular task types that were most relevant to their real world needs.

Another key concern articulated by teachers was that the LL Apps did not facilitate instruction that aligned with their learners' current developmental level. In particular, it was noted that the sequencing of the complexity of instruction was inadequate. Indeed, sequencing of tasks from less to more cognitively demanding is important as it enables learners to engage in tasks while at the same time having adequate cognitive resources available for focusing on form (Robinson, 2009). Although sequencing task complexity is challenging (Long, 2016), frameworks do exist that could be used by LL App developers to achieve this goal (Robinson, 2009, p. 304; Robinson & Gilabert, 2007, p. 164). We also encourage the involvement of teachers in the validation of the resultant sequence of tasks (see, Révész, Michel, & Gilabert, 2016). Adaptive pre-testing approaches that are based on valid and more fine-grained sequences of tasks of increasing complexity may also be useful in more precisely situating learners to an appropriate starting level in the instructional program.

Teachers' impressions also suggest that LL Apps would be improved by ensuring they can provide a greater variety and quantity of rich and elaborated input. In terms of unidirectional input, learners should be able to choose from a broad range of authentic reading or listening texts that learners must seek to understand in order to complete tasks. These authentic texts should be scaffolded with elaborations such as hyperlinked glosses, and multimedia elaborations. For example, we imagine a learner listening to an engaging authentic listening text in order to achieve a clearly defined goal, which the learner individually chooses. When the learner taps the screen, the text is paused, and an elaboration is made available. Perhaps this elaboration is a pre-recorded video file of a native speaker describing the approximate meaning of the section of the text at which the learner chose to pause, or perhaps a written elaboration, or perhaps both. Providing opportunities to listen and read a rich variety of elaborated texts with a communicative focus may also provide opportunities for inductive learning. Such input could also provide a practical approach to facilitating focus on form. For example while listening, if unsure of meaning, the learner could tap the screen and be presented with the subtitles (orthographical form) of the poorly understood section. Another click on the onscreen lexis or grammatical forms of the subtitles could take learners to a deeper level of elaboration; for example, metalinguistic descriptions of these forms and their functions in the learner's first language.

Of course, another way to facilitate meaningful engagement with rich and elaborated input is through communication with other language learners, and potentially first language/proficient speakers of the target language. In this regard, we see particular value in the affordances of real time chat, voice over the internet functionality, as well as discussion forums. Again, for example, imagine the hypothetical learner mentioned above that has just finished listening to an elaborated listening text. This learner could then proceed to communicate, perhaps by way of live chat, with another learner who has also just listened to the same text. Communicative interactions of this

nature also afford opportunities to promote collaborative learning, and also to receive real time feedback, albeit from other language learners.

In summary, we have drawn on the teachers perceptions' of the affordances of LL Apps and offered a number of suggestions on how to more strongly align their design with the principles of TBLT. It is acknowledged that the facilitation of such tasks as described above is technically and logistically far more complex that the provision of yes/no quizzes, vocabulary flash cards and fill-the-gap dialogues, but then so too is the acquisition of a second language. Despite the inevitable challenges, we view the effort to integrate the principles of TBLT into LL App design as very worthwhile. The principles of TBLT provide a systematic approach to bridging the gap between how language instruction has been traditional applied within LL Apps and what empirical evidence tells us about how language acquisition actually takes place.

**Funding Acknowledgement**

This research was funded by the Faculty of Education and Arts, University of Newcastle, Australia.

**References**

Berg, B. (2007). An Introduction to content analysis. In B. L. Berg (Ed.), *Qualitative Research methods for the social sciences* (pp. 238-267). Boston, MA: Allyn and Bacon.

Burston, J. (2014a). The reality of MALL: Still on the fringes. *Calico Journal*, *31*(1), 103-125. doi:10.11139/cj.31.1.103-125

Burston, J. (2014b). MALL: The pedagogical challenges. *Computer Assisted Language Learning*, *27*(4), 344-357. http://dx.doi.org/10.1080/09588221.2014.914539

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, *27*(1), 4-20. doi:10.1017/S0958344014000159

Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, *7*(3), 50-80. Retrieved from https://www.lltjournal.org//

Ellis, R. (1989). Are classroom and naturalistic acquisition the same?: A study of the classroom acquisition of German word order rules. *Studies in Second Language Acquisition*, *11*(3), 305-328.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, United Kingdom: Oxford University Press

Ellis, R. (2005). Principles of instructed language learning. *System*, *33*(2), 209-224. doi:10.1016/j.system.2004.12.006

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, *40*(1), 83-107. doi:10.2307/40264512

Ellis, R. (2010). Forward. In M. Thomas & H. Reinders (Eds.), *Task based language learning and teaching with technology* (xvi-xviii). London, United Kingdom: Continuum

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching*, *50*(4), 507-526. doi:10.1017/S0261444817000179

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*(2), 339-368. doi:10.10170S0272263106060141

Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, *19*(1), 3-17. doi:https://doi.org/10.1075/aila.19.03gas

Guichon, N., & Hauck, M. (2011). Teacher education research in CALL and CMC: More in demand than ever. *ReCALL*, *23*(3), 187-199. doi:10.1017/S0958344011000139

Haines, K. (2015). Learning to identify and actualize affordances in a new tool. *Language Learning & Technology*, *19*(1), 165-180. Retrieved from https://www.lltjournal.org//

Kukulska-Hulme, A., & Shield, L. (2008). Overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, *20*(3), 271-289. doi:10.1017/S0958344008000335

Levy, M. (2015). The role of qualitative approaches to research in CALL contexts: Closing in on the learner's experience. *Calico Journal*, *32*(3), 554-568. doi: 10.1558/cj.v32i3.26620

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, *27*(3), 361-386. https://doi.org/10.1017/S0272263105050163

Long, M. H., & Crookes, G. (1992). Three approaches to task□based syllabus design. *TESOL Quarterly*, *26*(1), 27-56. https://doi.org/10.2307/3587368

Long, M. H. (2016). In defence of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, *36*, 5–33. https://doi.org/10.1017/S0267190515000057

Little, D., Devitt, S., & Singleton, D. M. (1988). *Authentic texts in foreign language teaching: Theory and practice*. Dublin, Ireland: Authentic Language Learning Resources Ltd.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, *27*(3), 405-430. https://doi.org/10.1093/applin/ami051

Matthews, J., Cheng, J., & O'Toole, J. M. (2015). Computer-mediated input, output and feedback in the development of L2 word recognition from speech. *ReCALL*, *27*(3), 321-339. https://doi.org/10.1017/S0958344014000421

Matthews, J., O'Toole, J. M., & Chen, S. (2017). The impact of word recognition from speech (WRS) proficiency level on interaction, task success and word learning: Design implications for CALL to develop L2 WRS. *Computer Assisted Language Learning*, *30*(1-2), 22-43. https://doi.org/10.1080/09588221.2015.1129348

Nunan, D. (2004). *Designing tasks for the communicative classroom*. Cambridge, United Kingdom: Cambridge University Press.

Pica, T., Lincoln- Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly*, *30*(1), 59-84. https://doi.org/10.2307/3587607

Reinders, H., & Pegrum, M. (2015). Supporting language learning on the move: An evaluative framework for mobile language learning resources. In B., Tomlinson (Ed) *Second language acquisition research and materials development for language learning,* (pp.116-141). London, UK: Taylor & Francis.

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, *38*(4), 703-737. https://doi.org/10.1017/S0272263115000339

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, *45*(1), 99-140. https://doi.org/10.1111/j.1467-1770.1995.tb00964.x

Robinson, P. (2009). Syllabus design. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 294–310). Oxford, UK: Blackwell.

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching*, *45*(3), 161-176. doi: https://doi.org/10.1515/iral.2007.007

Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal*, *34*(2). https://doi.org/10.1558/cj.27623

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.

TESOLacademic. (2014, June 28). Prof. Ellis on task-based pedagogy: The what, why and how [Video file]. Retrieved from https://www.youtube.com/watch?v=zdRibzXW2TI

Thomas, M. & Reinders, H. (2010). Deconstructing Tasks and Technology. In M. Thomas & H. Reinders (Eds.), *Task based Language Learning and Teaching with Technology* (1-13). London, United Kingdom: Continuum

Ushioda, E. (2013). Motivation matters in mobile language learning: A brief commentary. *Language Learning & Technology*, *17*(3), 1–5. Retrieved from https://www.lltjournal.org//

## Appendix 1 - Online questionnaire

Note: Subcategory descriptors (i.e., those that are bolded and italicised below such as ***General***, ***Activities***, ***MP1***, ***MP2***, etc.) were not presented on the surveys made available to the teachers. A text field was made available after each of the questions/statements such that teachers could provide their written responses. These subcategories are drawn from Doughty and Long (2003, pp. 56-67).

A list of statements are presented. Please respond to these statements in the text field based on your experience with the LL App.

### *GENERAL*

- Which application are you reviewing in this survey?
- What did you like most about this application?
- What did you like least about this application?
- What does this application ask the learner to do?

### *ACTIVITIES*

#### *Using tasks as the unit of analysis (principle 1)*

- This application provides the learner with opportunities to communicate meaningful messages.
- This application requires the learner to apply his/her knowledge / understanding in order to achieve a task or goal.
- This application mostly focuses on the individual parts of the language not on the overall meaning.

#### *Emphasising learning by doing (principle 2)*

- The application involves learners using language to solve simple problems.
- The application involves learners using language to solve complex problems.
- The application involves language learning tasks that were contextualised within simulated or real-world scenarios.

### *INPUT*

#### *The use of elaborated input (principle 3)*

- The application provides sufficient opportunity to listen to and read target language input (listening materials and reading materials).
- The application involves engaging with simplified texts.
- The application involves engagement with authentic/genuine texts.
- The application provides a facility to communicate/negotiate meaning with another language learner (i.e., through chat functions or verbal discussion).
- The application provides useful scaffolding to help learners (for example, explanations in first language, pictures, etc.).

***Providing rich input (principle 4)***

- The application mainly presented text that contained basic grammatical forms.
- The application presented text that contained complex grammatical forms.
- The application presented vocabulary that was of a basic level.
- The application presented vocabulary that was of an advanced level.
- The application presented examples of realistic discourse between native and non-native speakers.

***LEARNING PROCESSES***

***Emphasising inductive learning (principle 5)***

- The application facilitates the use of chunks of language, such as idioms and collocations.
- The application facilitates explicit learning (e.g., activities which provide opportunities to explicitly analyse the target language).
- The application facilitates implicit learning (e.g., activities which provide opportunities to learn language inductively).

***Facilitating focus on form (principle 6)***

- The application delivers explicit instruction on isolated structures in a pre-determined manner.
- There is a good balance between focus on individual structures of the language and the overall meaning.

***Providing negative feedback (principle 7)***

- The application provides opportunities for the learner to notice errors that have been made.
- When an error is made an alternative correct form is provided.

***Respecting developmental processes (principle 8)***

- The application provides opportunities for the learner to select the order of content and learning activities.

- The application allows the learner to select content or learning activities according to their interest.
- The application often delivers content which is too easy or too difficult.

***Promoting collaborative learning (principle 9)***

- The application only involves the learner working individually with the application (no interaction with other learners using the application occurs).

***LEARNERS***

***Individualizing instruction (principle 10)***

- The application individualises content and activities based on the attributes of the learner.
- The application provides learners with the opportunity to select various modalities according to preference (audio, visual, written, combination of concurrent audio/visual/written modes).

**Appendix 2 - The foundation of semi-structured interviews**

Note: Semi-structured interviews entailed a guided, rather than a prescriptive discussion. The goal was to facilitate a forum within which teachers could present their candid thoughts and feelings about the LL Apps at the centre of the study.

- From your experiences with this LL App would you use these yourself either in your own language development or as a resource for the language classroom?
- What were some of the features of this LL App that stand out in your mind as being most useful?
- What were some of the key limitations of this LL App that stand out in your mind?
- If there are aspects of this application that you could change or add to what would they be?
- Is there anything else that you would like to say about this LL App?
- Could this LL Apps be used effectively alongside the current syllabus?'
- In your view what would make the application more compatible with the syllabus that you currently use in your language teaching?
- If you could design an ideal mobile phone application for the purposes of your language learning/teaching what would it do?
- What would be some of the key features of that application?
- I guess this is like a dream list for a language learning application for your teaching. Can you describe the perfect language learning app for your teaching purposes?
- Is there anything else you'd like to say about language learning applications generally?