THE TESOLANZ JOURNAL

VOLUME 25, 2017

ISSN 1172-9694

The TESOLANZ Journal is published once a year by TESOLANZ, Teachers of English to Speakers of Other Languages, Aotearoa/New Zealand (Incorporated), Te Ropu Kaiwhakaako Reo Ingarihi ki Iwi Reo Ke. The journal is double blind peer reviewed. A subscription to the journal is included in the annual membership fee of the Association. Enquiries about membership and subscription to the journal should be addressed to:

TESOLANZ Linguistics and Applied Language Studies Victoria University of Wellington PO Box 600 Tel: (04) 463 5600 Fax: (04) 463 5604

Those interested in submitting an article, report or book review should consult the Notes for Contributors printed at the back of this volume and be a current member of TESOLANZ.

EDITOR

Dr Angela Joe Linguistics & Applied Language Studies Victoria University of Wellington

EDITORIAL ASSISTANT

Dr Patrick Coelho Linguistics & Applied Language Studies Victoria University of Wellington

EDITORIAL BOARD

Professor Gary Barkhuizen Applied Language Studies & Linguistics University of Auckland

Dr Margaret Kitchen Faculty of Education University of Auckland

Professor John Read Applied Language Studies & Linguistics University of Auckland

Dr Gillian Skyrme School of Languages Massey University

REVIEWS EDITOR

Dr Katherine Quigley Linguistics & Applied Language Studies Victoria University of Wellington

Associate Professor David Crabbe Faculty of Education Victoria University of Wellington

Associate Professor Jonathan Newton Linguistics & Applied Language Studies Victoria University of Wellington

Dr Corinne Seals Linguistics & Applied Language Studies Victoria University of Wellington

Professor Cynthia White School of Languages Massey University

THE TESOLANZ JOURNAL

CONTENTS

ARTICLES

Repeating a monologue under increasing time pressure: A replication of Thai and Boers (2016)	1
Frank Boers & Chau Thai	
Language learning principles and MALL: Reflections of an adult learner 1 John Macalister	2
Can image schemata help L2 learners to differentiate between semantically similar words? The case of basic auditory verbs <i>listen</i> and <i>hear</i>	25
Can adult beginner English language learners manifest signs of a reading habit? 3	39

Can adult beginner English language learners manifest signs of a reading habit?..... 39 Jenny Field

ARTICLES

REPEATING A MONOLOGUE UNDER INCREASING TIME PRESSURE: A REPLICATION OF THAI AND BOERS (2016)

Frank Boers and Chau Thai Victoria University of Wellington and University of Danang

Abstract

In the 4/3/2 activity, learners repeat a monologue under increasing time pressure. The increasing time pressure is intended to push the speaker to render the same content faster and more efficiently, which is why this activity has been recommended for fluency practice. However, a recent small-scale experiment (Thai & Boers, 2016) cast doubt on the proclaimed merits of the 4/3/2 procedure, because (a) the increasing time pressure was found to negatively affect linguistic accuracy and syntactic complexity, and (b) fluency gains were found also in a comparison condition where learners repeated their talks without increasing time pressure. The present article reports a replication of Thai and Boers (2016), the only difference being the less familiar topic used to elicit learner speech. Overall, the new results parallel those of the earlier study, with the exception that this time no noticeable fluency gains occurred in the absence of time pressure. The influence of topic selection on the way learners tackle 4/3/2 type tasks is discussed.

Introduction

In the 4/3/2 activity (Maurice, 1983), learners talk to a partner about a given topic for four minutes, repeat their talk to another partner but in a reduced time span of three minutes, and then try to deliver the same content to a third listener in just two minutes. Each delivery of the monologue is addressed to a new listener, so that the speaker does not feel the need to makes changes to the content in order keep the listener interested. The advantage of repeating the same content, so the argument goes, is that little effort needs to be invested in content generation as the talk is repeated. As a result, the speaker can pay more attention instead to the packaging or formulation of that content (e.g., Fukuta, 2016; *Cf.* Levelt, 1989). When learners are given the opportunity to repeat a speaking task, the second performance is often more fluent (Bygate, 2001; Wang, 2014) thanks to the affordances of rehearsal. However, fluency can be enhanced further by the element of increasing time pressure, which distinguishes the 4/3/2 activity from other repeated speaking tasks (e.g., Arevart & Nation, 1991; de Jong & Perfetti, 2011; Wood, 2009).

Apart from the benefits for fluency development, there have been suggestions (e.g., Nation, 1989) that the 4/3/2 activity also brings about improvements in other dimensions of speaking performance, such as grammatical accuracy and syntactic complexity. However, the results of two recent studies, Boers (2014) and Thai and Boers (2016), cast doubt on this. While speech rate was found in these studies to increase dramatically from the first to the third iteration of the participants'

monologues, no improvements were attested for accuracy and complexity. Comparison conditions, where participants repeated their talks *without* increasing time pressure, also yielded noticeable gains in speech rate, and in addition seemed to create better opportunities for enhancing accuracy and complexity. This led the authors to question whether adding time pressure to task repetition was actually a judicious choice of task condition. These were small-scale studies, however, and—like all studies—they require replication (Porte, 2012). The present article reports such a replication.

A synopsis of Thai and Boers (2016) and grounds for replication

In Thai and Boers (2016), two groups of ten intermediate EFL students in Vietnam were given the task to deliver a monologue three times, either with or without increasing time pressure. Because piloting had revealed students at this level of proficiency found it too challenging to talk in English about a topic for four minutes, the 4/3/2 procedure was modified to a 3/2/1 procedure. In the condition without increasing time pressure, the students were given two minutes for each delivery of their talks (i.e., 2/2/2). The topic the students were asked to talk about was their favourite film. This prompt was chosen because the students had previously indicated on a list of possible topics (compiled from IELTS practice materials) that this was familiar subject matter, which they felt comfortable talking about.

The recorded monologues were analysed for changes in speech rate (as a measure of fluency), error-free clause ratios (as a measure of accuracy), and mean number of clauses per AS (analysis-of-speech) unit (as a measure of syntactic complexity). In the task condition with increasing time pressure (i.e., 3/2/1), the speech rate was found to increase dramatically (by almost 65%) from the first to the third delivery of the tasks. However, no improvements were observed for the measures of accuracy and complexity. In the task condition without increasing time pressure (i.e., 2/2/2), the speech rate was also found to increase, albeit less spectacularly so (by about 16%). Analyses of the transcripts revealed that, as the students in the 2/2/2 condition conveyed content more efficiently when they repeated their talks, they also managed to add content, such as more detail about their favourite film's plot and further evaluative comments. Despite the much faster delivery, the third rendering of the talks under the 3/2/1 condition typically showed reduced content. The time pressure in that condition also appeared to compromise any chances of improvements at the level of grammatical accuracy and syntactic complexity. The 2/2/2 condition, by contrast, did generate improvements-albeit very modest ones-regarding the grammatical accuracy and the syntactic complexity of the talks. Together, these findings led the authors to suggest that "the 4/3/2 technique is recommendable if the sole aim of the activity is to push fluency. However, if the objective is to foster other speech qualities as well, it would seem that task repetition without increasing time pressure is the more judicious option" (Thai & Boers, 2016, p. 386).

This recommendation was based on a small-scale experiment, in which just one prompt, i.e., "my favourite movie", was used to elicit monologues from the students. On the one hand, this made Thai and Boers (2016) a more tightly controlled experiment than Boers (2014), where a diversity of prompts was used with a heterogeneous group of participants. On the other hand, it may have been a prompt that lent itself particularly well to a fluency activity, because familiarity with the topic meant that the students needed to invest little effort in the generation of new ideas, and so they could readily allocate their cognitive resources to the expression of known content (Bui & Huang, 2016). Topic familiarity is a matter of degree, however, and teachers wishing to implement the 4/3/2 activity (or 3/2/1 activity) on a regular basis probably need various prompts, some of which may be more novel to their students than others. That regular incorporation of 4/3/2 activities in a language programme can generate long-term benefits for fluency development was demonstrated by de Jong and Perfetti (2011). Such regular implementation inevitably requires a diversity of topics for learners to talk about, though, and so the question becomes how much the choice of topic matters for students to reap the rewards of 4/3/2 practice. In what follows, we report an experiment which very closely mimics Thai and Boers (2016), apart from the prompt chosen to elicit the students' monologues. The new prompt was "advantages and disadvantages of advertisements".

The question we address is essentially whether the results of this approximate replication will parallel those of the earlier study in that (a) repeating a monologue under increasing time pressure brings about fluency gains over and above task repetition alone, but this at the cost of accuracy and complexity, and that (b) noticeable fluency gains are obtained also in a comparison condition where learners repeat a monologue without increasing time pressure.

Method

Participants and choice of topic

The same two groups of ten Vietnamese EFL students (N = 20) who participated in Thai and Boers (2016) also participated in this replication, during the same school year as when the data for the earlier study were collected. These were all 15-year-old high school students who had all followed the same English curriculum at school for the previous four years, at a rate of three 45-minute classes per week.

As preparation for the earlier study, a list of prompts collected from IELTS practice materials had been presented to the students, who had been asked to indicate on a fourpoint scale the degree to which they were familiar with each topic. While a topic indicated by the students as highly familiar ("my favourite movie") was selected for the Thai and Boers (2016) experiment, we chose for this replication study a topic which all the participating students gave a lower familiarity rating to: "advantages and disadvantages of advertisements". This does not mean that advertising was an alien concept for these learners. Advertisements are, of course, part of everyday life in Vietnam.

Procedure

The experimental procedures mimicked those of Thai and Boers (2016). Because piloting had revealed that speaking non-stop in English for as long as four minutes was too daunting for these students, the 4/3/2 activity was modified to a 3/2/1 sequence, where the students were first asked to talk about the given topic for three minutes, then for two minutes, and finally for one minute. Ten students repeated their talks in this shrinking time condition. Ten other students repeated their talks in a constant time condition—with a window of two minutes per delivery (2/2/2). On receiving the prompt, students were given three minutes for (unassisted) pre-task planning. No additional planning time was allocated between the deliveries of the talks. As stipulated by the guidelines for implementing the 4/3/2, the students were assigned a new listener per delivery of their talks (using a rotating system with fellow students who were not at this time expected to perform the output activity). All the participants' monologues were audio-recorded and transcribed.

Analysis

Fluency is a multifaceted construct (Segalowitz, 2010), but-like Thai and Boers (2016)—we will focus here on speech rate or, more specifically, trimmed speech rate. Speech rate refers to the mean number of syllables produced in a given time span. 'Raw' speech rate includes the syllables that make up filled pauses, false starts, repairs and repetitions, which are actually phenomena considered to be signs of disfluency rather than fluency. The more meaningful temporal measure of speech fluency, therefore, is 'trimmed' (or 'pruned') speech rate (Lennon, 1990), which is computed by counting the mean number of syllables in a given time span excluding said disfluency phenomena. Trimmed speech rate per minute was the fluency measure focused on most in Thai and Boers (2016) as well. Also analogous to that earlier study, the degree of accuracy of the language used in each talk was gauged by calculating the ratio of error-free clauses per total number of clauses, and syntactic complexity was gauged by counting the mean number of clauses per AS unit (Foster, Tonkyn, & Wigglesworth, 2000)-which is essentially a measure of the proportion of subordinate clauses in text. In Thai and Boers (2016), the repeated monologues were also examined for changes at the level of lexis, but no trends were discerned. No such trends were discerned in the speech samples examined in the present study either, and so we will exclude that performance aspect from this report.

Results and Discussion

The first point worth mentioning is that few of the students who followed the 3/2/1 procedure actually delivered a first talk that was a full three minutes long (mean length in seconds: 164; range 140–180; SD 14). This suggests that some of the students ran out of ideas regarding the advantages and disadvantages of advertisements before the

three minutes were over. Apart from the first talk in the 3/2/1 condition, all the other deliveries of the talks (in both the 3/2/1 and 2/2/2/ task conditions) did meet the task requirements in terms of length.

Table 1 presents the results of the analyses of the two groups' repeated talks in terms of the aforementioned measures used to gauge fluency, accuracy and complexity. The first thing to notice is that in the condition without increasing time pressure (i.e., 2/2/2) there were virtually no changes in mean speech rate. This is very different from Thai and Boers (2016), where fluency gains did occur when students repeated their monologues in the 2/2/2 condition. The virtual absence of change in the 2/2/2 condition in the current experiment also stands in stark contrast to the 3/2/1 procedure, where the number of syllables produced per minute increased by close to 75% from the first to the third delivery. Figure 1 illustrates the contrasting trends under the two task conditions (i.e., 3/2/1 vs. 2/2/2). A two-way ANOVA confirms that these contrasting trends reflect a significant interaction between task repetition (talk 1 vs. talk 2 vs. talk 3) and task condition (3/2/1 vs. 2/2/2): F 27.9; p < .0001.





Trimmed speech rate across the repeated monologues

It is worth noting, however, that the mean speech rate in the 3/2/1 condition was much lower in the first delivery than it was in the 2/2/2 condition, and consequently there was more room for gain in subsequent deliveries. The participants may have experienced the three-minute window as ample time to convey the content they had prepared, and this may account for their slow delivery of that content. It actually took until the third delivery (where the time pressure was greatest) for the mean speech rate of the students in the 3/2/1 condition to rise above that of the students in the 2/2/2 task condition. While the two-way ANOVA reveals a close-to-significant between group difference (*F* 3.46; *p* = .07), subsequent Tukey HSD tests confirm the impression that the 3/2/1 group's mean speech rate was significantly *lower* in the first delivery (*p* < .01) and only became significantly higher than that of the 2/2/2 group in the third delivery of the talks (p < .01).

Means (and standard deviations) per performance measure								
	2/2/2			3/2/1				
	1 st talk	2^{nd}	3 rd	1 st talk	2^{nd}	3 rd talk		
		talk	talk		talk			
Trimmed speech	123	126	124	101	119	175		
rate / minute	(14)	(18)	(14)	(12)	(10)	(24)		
Error-free clause	.19	.21	.21	.33	.28	.27		
ratio	(.12)	(.15)	(.14)	(.15)	(.13)	(.15)		
Clauses/ AS-unit	1.96	2.02	2.18	2.39	2.21	2.08		
	(0.34)	(0.34)	(0.52)	(0.52)	(0.38)	(0.38)		

 Table 1:

 Means (and standard deviations) per performance measure

Turning now to the other measures of speech performance, task repetition as such appears to have had a modestly positive effect in the 2/2/2 condition (see Table 1), which is consistent with earlier studies on task repetition (e.g., Lynch & Maclean, 2000), including Thai and Boers (2016). Statistical significance is reached only for the increase in clauses/AS-unit ratio, however: F 4.31; p = .03. In the 3/2/1 condition, by contrast, there was a slight downward trend in the clauses/AS-unit ratio, but this trend was non-significant. Figure 2 illustrates the contrasting trends.



Figure 2:

Clauses/AS-unit ratios across the repeated monologues

A downward trend was observed in the 3/2/1 condition for the accuracy measure as well, but also this trend fell short of significance, albeit narrowly so: F 2.89; p = .08. In sum, contrary to the modest gains in the 2/2/2 conditions, no gains regarding accuracy and syntactic complexity were observed in the 3/2/1 condition, suggesting

that the increasing time pressure compromised the potential benefits of task repetition on these dimensions. This is consistent with Skehan's (e.g., 2009) trade-off model, according to which the performance aspects of fluency, accuracy and complexity compete for the speaker's limited processing capacity. Further support for this is lent by the computation of correlation coefficients to evaluate the association between the students' (N = 20) gains in speech rate from their first to their third talk on the one hand, and their gains in error-free clause ratios and clauses/AS unit ratios on the other. These correlations with fluency gains are both negative, with r = -.440 and r = -.315for accuracy gains and complexity gains, respectively. It is also worth noting in this regard that it was the talk delivered under the least time pressure (i.e., when students were given a three-minute window) that tended to exhibit the most accuracy and complexity, which suggests the students monitored their output best when they took their time to say what they had planned to convey.

Conclusions, Implications and Perspectives

Overall, the results of this replication study parallel those found in the earlier study: adding time pressure to task repetition led to a dramatic increase in speech rate, but it also militated against the affordances of task repetition when it comes to the dimensions of accuracy and complexity.

Contrary to what was found in the earlier study, however, the comparison condition *without* increasing time pressure did not bring about a noticeable increase in speech rate. As the only design difference with the earlier study was the choice of prompt, it is likely that it was this prompt which affected performance differently in the 2/2/2 condition. In Thai and Boers (2016), the topic which the students were asked to talk about was their favourite movie. This was a topic which they could say a lot about, as evidenced by the observation that they added content to their monologues as these were repeated. The increase in speech rate (i.e., saying more in the same short time span) was thus associated with the greater amount of content conveyed. In the present replication, however, the topic—advertising—was one the students had less to say about, and so the students probably felt no inclination to increase their speech rate with a view to adding content. Scrutiny of the transcripts indeed shows that hardly any content was added as the monologues were repeated.

Turning back to the condition *with* increasing time pressure, the dramatic rise in speech rate between the first and the third performance suggests that also relatively novel prompts can serve the purpose of 4/3/2-type activities—at least as far as reaping the benefits regarding fluency are concerned—although it is worth bearing in mind that the size of the dramatic jump in speech rate may to some extent be an artefact of the very slow first delivery.

On the downside, the results of this replication lend support to the assertion that grammatical accuracy and syntactic complexity are less likely to benefit from the

increasing time pressure which characterises the 4/3/2 procedure. To be fair, however, the 4/3/2 activity was originally proposed solely as a fluency development activity. In that sense, then, the finding that this activity also 'works' with prompts that are relatively novel to students is encouraging for teachers who wish to implement this activity on a more regular basis. It nevertheless stands to reason that topics which students have a lot to say about will lend themselves best to the activity. Teachers may therefore find it useful to consult their students when selecting prompts. It is also important to bear in mind that familiarity with a topic can easily be enhanced prior to a 4/3/2 activity. It is not difficult to envisage implementation of a 4/3/2 activity toward the end of a thematic course module through which the students have become knowledgeable about a topic that was relatively novel to them at the start.

The study reported here demonstrates the usefulness of (approximate) replications. All too often are conclusions drawn and recommendations made in our discipline on the basis of one-off experiments. Pure replication studies help us to evaluate whether earlier findings were perhaps 'accidental'. Approximate replication studies, such as the one reported here, can help us to evaluate to what extent observed trends are susceptible to relatively small variations in context and study design. Here, we have considered just one small variation—topic familiarity—but there are several more variables that invite further exploration. One is the profile of the learners. The experiment reported here was run with young EFL students, but it is not inconceivable that adult learners in an ESL setting may go about a 4/3/2 task quite differently.

Another important variable is the way the time pressure in a 4/3/2 type activity is operationalized. In the experiment reported here it was run as a 3/2/1 sequence, which entails an extremely drastic decrease in available time (where the length of the final talk is only one third of the first, whereas in the 4/3/2 sequence it is half the length of the first). In future explorations of the activity, teachers may want to experiment further with ways of imposing time limits (e.g., 3/2.5/2 minutes). Alternatively, they may simply encourage their students to re-deliver a talk more smoothly without setting time limits, and then unobtrusively time them for feedback purposes. Although the 4/3/2 activity has attracted a fair amount of attention from pedagogy-minded applied linguists, it is clear that there is still ample room for exploring the ways in which teachers adapt the activity and make optimal use of it in their own classrooms.

References

- Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal*, 22, 84–94. doi:10.1177/003368829102200106.
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal, 45*, 221–235. doi:10.1177/0033688214546964.
- Bui, G., & Huang, Z. (2016). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research* (Online First). doi:10.1177/1362168816656650.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 23–48). London, England: Longman.
- de Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, *61*, 533–568. doi:10.1111/j.1467-9922.2010.00620.x.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354–375. doi:10.1093/applin/ 21.3.354.
- Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. Language Teaching Research, 20: 321–340. doi:10.1177/1362168815570142.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387–417. doi:10.1111/j.1467-1770.1990.tb00669.x.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, *4*: 221–250. doi:10.1177/136216880000400303.
- Maurice, K. (1983). The fluency workshop. TESOL Newsletter, 17, 429.
- Nation, I. S. P. (1989). Improving speaking fluency. *System*, 17, 377–384. doi:10.1016/0346-251X(89)90010-9.
- Porte, G. (Ed.) (2012). *Replication research in applied linguistics*. Cambridge: Cambridge University Press
- Segalowitz, N. (2010). Cognitive bases of second language fluency. New York: Routledge.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*, 510–532. doi:10.1093/applin/amp047.
- Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, 50, 369–393. doi:10.10002/tesq.232.
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12, 39–57.

Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Philadelphia, PA: John Benjamins.

Appendix 1: Examples of a 1^{st} and a 3^{rd} delivery in the 3/2/1 condition (trimmed speech)

1st delivery

I will talk about the advantages and disadvantages of advertisement.

First, advertisement is good because as it appear on media such as TV it inform viewer about different product and service of different company. Viewer watch it and know like what and where and when and how they can buy something. They can see what is better and go buy it and what is cheaper too. By that company can introduct their stuff and sell them while customers can buy what they want because they know about them and buy them. Beside advertisement can make life more interesting because company have to make interesting ad to attract viewer and compete with other company.

But I think second, advertisement is not good too. When company overuse it and spend a lot of money on it and make it appear so long on TV, to do so they have to reduce their product quality. So I don't want to buy a thing that appear too many on TV.

Sometime I hate advertisement because it encourage people to buy thing they don't need. It make people want the product but not needing it, and the more they watch advertisement customers become greedy and want everything.

3rd delivery

I will talk about the advantages and disadvantages of advertisement.

First, advertisement is good because it appear on media so it inform viewer about different kind of product and service of different company. Viewer watch it and know how they can buy something. By that company can introduce their product to sell them when customers can buy what they want.

Second, I think advertisement is not quite good, especially when company overuse it and spend lot of money. It appear so long. To do so they have to reduce product quality. I hate advertisement because it encourages people to buy thing they don't need. It make people want the product but not needing it.

LANGUAGE LEARNING PRINCIPLES AND MALL: REFLECTIONS OF AN ADULT LEARNER

John Macalister Victoria University of Wellington

Abstract

Mobile assisted language learning (MALL) offers new possibilities for language learning, and particularly perhaps for learning beyond the classroom. This paper reports on one adult learner's experience with two apps, Duolingo and Quizlet, and considers their effectiveness in terms of language learning principles. While the experience was positive, the apps do have some of the same shortcomings—in terms of applying principles—as coursebooks. To some extent this reinforces the important role that teachers can play in the classroom, but to fulfil that role teachers do require a commitment to on-going professional development, both formal and informal.

Introduction

The field of language learning has long been a field of change. Succeeding waves of new approaches have taken us from grammar translation through audiolingualism and communicative language teaching to task-based language teaching, to name some of the better known. New approaches have been informed by new thinking about input (Krashen, 1985), output (Swain, 1988), interaction (Long, 1996), and noticing, again to name a few. And then there have been the technological innovations, with Banda machines giving way to duplicators, duplicators to photocopiers, audio cassettes to CDs, CDs to digital sound files, filmstrips to videos, videos to DVDs, as examples. Many of these new technologies accompanied the advent of the computer and digital ages, and in turn triggered a raft of new acronyms such as CALL (computer-assisted language learning), TELL (technology-enhanced language learning), and, more recently, MALL (mobile-assisted language learning). MALL, the focus of this article, is explored more fully in the next section.

Language teachers may not, of course, always welcome technological innovation. Some may be seduced by technology's affordances, but by no means all. Hubbard (2008), for instance, catalogues seven arguments deployed against CALL adoption, arguments that include inertia and ignorance as well as insufficiency of the infrastructure and the lack of an established methodology. At the same time, and not least because learners come to expect a link between the technology they use out of class and the activities they are expected to do in class, total resistance to technological innovation can seem Canute-like. Uncritical embracing, however, may not be the correct response either. The four major questions posed by Salaberry (2001, p. 51) in a reflection on the use of technology remain as relevant now as they were when first asked:

- 1. Is increased technological sophistication correlated to increased pedagogical effectiveness?
- 2. Which technical attributes specific to new technologies can be profitably exploited for pedagogical purposes?
- 3. How can new technologies be successfully integrated into the curriculum?
- 4. Do new technologies provide for an efficient use of human and material resources?

The first three of these questions guide the reflection that is at the core of this paper, although it is worth emphasising that the interest here is on complementing the classroom rather than integrating into the classroom. There exists already a substantial literature on curricular integration (e.g. Kukulska-Hulme, Norris, & Donohue, 2015).

Mobile-assisted language learning

Considering the second of Salaberry's questions, mobile phones and the development of apps are the new technology that can be turned to pedagogical purposes, giving rise to what has been variously described as mobile learning, m-learning, and MALL. Pegrum (2014, p. 16) sees MALL, like mobile learning itself, as "a fuzzy concept, referring to language learning scenarios in which varying degrees of mobility pertain to the devices, the learners and the learning experience itself". The three potentially mobile elements in this definition can be viewed as the 'technical attributes' of the new technology. Drawing on these elements, Pegrum goes on to suggest a useful three-level taxonomy, from only the devices being mobile, through devices and learners being mobile, to all three components being mobile. Another way of making this spectrum of mobility clear is to remind ourselves that "[t]he word 'mobile'...is not simply synonymous with a mobile phone" (Kukulska-Hulme et al., 2015, p. 5).

Although of relatively recent genesis, MALL has received considerable research attention already, and is at times spoken of as being potentially transformative of language learning and teaching. Kukulska-Hulme et al. (2015, p. 3), for example, claim that "[t]he adoption of mobile devices has potentially far-reaching consequences for learners, learning design and how learning is supported by teachers and advisors" while Godwin-Jones (2017, p. 3) sees smartphone use as having the potential for fundamental disruption of learning and teaching. However, while much of the research attention does extol the benefits of MALL over more traditional approaches in a defined area of learning, issues have also been flagged. One such is learner uptake. Stockwell (2010), for example, reported on the experiences of three cohorts of English language learners in Japan, 175 in total, over a three-year period. Learners could choose to complete vocabulary activities on a mobile phone or on a desktop computer. He found "the number of learners who did not attempt to use the mobile platform at all was consistently high" although argued for a trend towards increased use of mobile phones, at the same time recognising that "numbers [of users] remained quite low"

(Stockwell, 2010, p. 106). Among the reasons proposed for this lack of uptake was that it took longer to complete the activities on a phone than on a PC.

Stockwell's study is typical of a fair number of other reported uses of MALL (e.g. Liu, 2016) in that it appears teacher-centred and teacher-directed. These uses can be classified as instances of 'tutorial MALL', where students "complete low level consolidation exercises" outside the classroom thus allowing students to "work at their own pace and according to their own needs" (Pegrum, 2014, p. 104). This teacher control of the learning experience may not perhaps be seen as much of an issue, but it certainly deserves to be considered when discussing questions of, for instance, learner uptake. After all, if teachers are not exploiting the potential of mobile devices, if learners are not motivated by teachers' use of MALL, the technology may not deliver pedagogical rewards. Furthermore, it should not be assumed that because students are digital natives they do not "need and value the advice and guidance from teachers" about the uses of new technology (Godwin-Jones, 2017, p. 10). This, however, raises questions about teacher willingness to learn about, and to make properly informed decisions about adopting, these new technologies. This question is returned to in the pedagogical implications section of this article.

One area of investigation that does not yet seem to have received much attention is the potential of MALL to contribute to learning beyond the classroom. Like MALL, learning beyond the classroom is a similarly new area of research interest. It is, indeed, an area still seeking to define itself. Benson (2011), for instance, proposes four dimensions for the field: location, formality, pedagogy, and locus of control. These would appear to fit well with MALL: 'location' fits with Pegrum's (2014) mobility of learners and devices, and Benson's other three dimensions with his 'learning experience'. Further, as these are dimensions along which variable positions can be taken, learning beyond the classroom enjoys similar conceptual 'fuzziness' to that suggested by Pegrum for MALL. What is striking, however, in two recent books focussing on this new area of interest (Benson & Reinders, 2011; Nunan & Richards, 2015) is the absence of any explicit attention to MALL in either. One of the contributions of this article, then, is to consider MALL from the perspective of a learner well and truly beyond the target language classroom.

Principles in language learning

Attention has been paid to the application of principles in many areas of language learning including, as examples, instructed second language acquisition (Ellis, 2005), learner autonomy (Cotterall, 2000), extensive reading (Day & Bamford, 2002), and intercultural communication (Newton, Yates, Shearn, & Nowitzki, 2010). No matter what the explicit focus of such lists, however, there is often a degree of commonality; the role of meaningful input often receives mention, for instance. As a result of the shared elements, it is also possible to create a list of general principles for language

learning, such as that provided by Nation and Macalister (2010). This is a list of twenty principles broadly applicable to any language programme.

The importance of principles is, simply, that they are drawn from research and theory about language learning, about *effective* language learning, and thus their application is expected to lead to better language learning outcomes for the learner.

Despite their importance, however, awareness of principles is not always evident in places where such awareness might be anticipated. One example is in reports on research studies. To illustrate with reference to MALL, Liu (2016) compared vocabulary learning in two conditions, learner-constructed concept-mapping and text-only strategy groups and found "that applying the concept-mapping strategy had a direct influence on learners' English learning ability" (p. 136). While her results suggested that it certainly did, and the results are worthwhile in themselves, they are not explained in terms of concept-mapping requiring greater *depth of processing* (Nation & Macalister, 2010, pp. 60-62), and that this in turn meant that learners under that condition spent more *time on task* (Nation & Macalister, 2010, pp. 58-60). Indeed, consideration of those two principles could have led to a hypothesis predicting the benefits of concept-mapping. Further, by not drawing attention to the principles at work, opportunities to generalise beyond this particular study may be limited.

Absence of awareness of principles in research reports is, however, much less likely to have an impact on language learning outcomes than it is when identified in coursebooks for these have the potential to influence the development of large numbers of learners. In a recent study looking at the application of language learning principles in coursebooks (Macalister, 2016), four of the twenty principles suggested in Nation and Macalister (2010) were identified as often absent. These will be discussed in more detail below but for now it suffices to note that they were principles affecting:

- 1. The four strands
- 2. Fluency
- 3. Frequency
- 4. Interference

Consideration of principles returns us to the first and third of Salaberry's (2001) questions, which are reflected in the research questions below. The first considers the pedagogical effectiveness of new technology; in particular, whether the new technology enhances learning outcomes. The third considers ways in which the new technology can be integrated into—or, in this case, complement—the classroom. The role of principles here is, first, that it is through the application of research-based principles that pedagogical effectiveness is achieved, and, second, that new technology may be a way of introducing principles that may otherwise be lacking, or under-

applied, in a course (as seen in Macalister, 2016). This in turn gives rise to the research questions that this paper addresses directly:

- 1. To what extent are language learning principles evident in selected apps?
- 2. In what ways could selected apps complement in-classroom language learning?

The study

This is an account of and reflection on my experience using two mobile apps for French language learning, autonomously and out-of-class. As many readers are likely to be involved in TESOL, it is worth emphasising that the focus is on the *experience* of learning, not the *object* of learning. That experience, it may also be worth pointing out, is seen through the eyes of a learner with a certain degree of metacognitive knowledge of language learning and teaching.

Language learning background

I began learning French at intermediate school. The first four years of French were mandatory, but I continued with French for a further three years, passing national examinations each year. Until the final year, the teachers were English-speakers and the focus was traditional, in keeping with the grammar-translation nature of exams. In the seventh year of formal French instruction, however, the teacher was a French speaker who, unsurprisingly but—for me at least—shockingly, spoke French in class. My attendance record that year was very poor.

That marked the end of my formal French language learning. Around eight years later, and occasionally over the following four years, I found myself in France on holiday and needing to use French for communicative purposes. A decade or more later I was working in a Francophone institute in Cambodia, which certainly led to considerable amounts of French language use, although in social situations it was often challenging and in formal, work-related settings at times I was in the position, familiar to many language learners, of knowing what was being talked about, but not what was being said.

Since Cambodia there had been occasional opportunities to use French, but in 2016, with the promise of a month's holiday in France ahead, I resolved to try and improve, to undertake some attempt at structured learning of the language.

The apps

For various reasons, classes were out of the question but I had a smartphone and was professionally aware (including through conference presentations, such as Stringer, 2016) of a small number of language learning apps, including Duolingo and Quizlet. Duolingo is a freely downloadable app for learning a large number of languages. Currently it offers 21 languages for English speakers, i.e. using English as the

interface, and provides languages to learn for speakers of 22 other languages, not limited to languages that employ Roman orthography. However, thirteen of these other languages offer English language learning only. While this is helpful for TESOL it does reflect what has been called the 'English problem' (Sauro, 2016); for example, a speaker of Polish or Bahasa Indonesia wanting to learn French or Vietnamese using Duolingo would need to do so through the English language interface.

In its online promotion Duolingo claims that 34 hours on the app is as effective as a whole university semester of study (https://youtu.be/8OebgtUjLg4). This claim is based on a study conducted by Vesselinov and Grego (2012), although aspects of the study have subsequently been questioned (Krashen, 2014). It is not, however, the case that one form of language learning is being presented as 'better' or 'worse' than the other. The interest is in effectiveness rather than superiority. With Duolingo, a language is arranged in a graded series of topics, both semantic (e.g. Clothing) and grammatical (e.g. Prepositions), with each topic containing a variable number of short lessons. Within any one lesson there are a range of activity types, with common examples being:

- Read and listen
- Select right word (word + picture)
- Jigsaw sentences
- Listen and write
- Vocabulary matching
- Read and speak
- Repetition

As Duolingo is graded, and given my pre-existing French language knowledge, the early lessons seemed easy, revision rather than new learning. However, as time went on, I began to encounter lexical items and phrases that were new to me and that I wanted to remember. I thus decided to download a second free language learning app, Quizlet, which is essentially an app for creating and using word cards, an activity promoted by Nation (1990, pp. 126-127).

With Quizlet, learners can either download pre-existing word lists or create their own. When creating their own, learners enter the word or phrase in the target language and a synonym, definition or explanation on the flip side. Entries are automatically linked to sound files so that learners can hear as well as read the entries. As well as the ability to run through the flashcards, the app provides opportunities to write the target words, to be tested, and to match words and meanings. Because of its ability to provide repeated exposure, Quizlet is an example of Spaced Repetition Software (SRS) and regarded as one of the most popular of its type (Godwin-Jones, 2010, p. 8).

In terms of Pegrum's (2014) taxonomy, Duolingo sits at the second level; the device and the learner are mobile, but the learning experience is pre-determined. Quizlet, on

the other hand, allows the learner to determine the content of the learning, and thus has elements at least of the third level.

Principles and the apps

In order to consider how selected apps could complement in-classroom language learning, the primary focus of this section will be Duolingo and the four principles that were found lacking in coursebooks (Macalister, 2016). Other principles will be discussed in the following section, where the experience of learning with the two apps is evaluated.

As Duolingo markets itself as a language course, it seems appropriate therefore to begin with what could be described as the unifying principle, that of the four strands, which proposes that "a course should include a roughly even balance of the four strands of meaning-focused input, language-focused learning, meaning-focused output and fluency activities" (Nation & Macalister, 2010, pp. 51-52). Of these four strands, language focused learning and fluency development are much more evident than the other two in Duolingo; while there are certainly opportunities for input and output, it would be hard to describe these as meaning focused, given the artificial nature of many of the language prompts. Translating into written French a sentence such as *My mother knows how to support me in difficult times* or speaking *Il va casser le robot* are language-rather than message-driven activities. They serve to teach a language point.

Fluency development is not about teaching new language but, rather, about giving learners opportunities to use what they already know (Nation & Macalister, 2010, p. 54). This may be achieved receptively, through seeing the form and retrieving the meaning, or productively, through retrieving the form in order to convey an intended meaning. Duolingo does provide such opportunities, for all four macro-skills, in a variety of ways, such as:

- Multiple exposures to target lexicon or structures within a mini-lesson
- Repeated opportunity within a mini-lesson if the wrong response is given the first time
- Repeated exposure to target lexicon and structures in later mini-lessons and later topics; this does not appear systematic, however
- Opportunities to do topic-based revision mini-lessons, and for personalised 'strengthen your skills' mini-lessons

The other two principles that were found lacking in coursebooks were both vocabulary based, and addressed the issues of frequency and interference. The frequency principle states that "A language course should provide the best possible coverage of language in use through the inclusion of items that occur frequently in the language, so that learners get the best return for their learning effort" (Nation & Macalister, 2010, p. 40). In other words, higher frequency—more common—language items should be

learned before lower frequency ones. To test the impression that learners were being presented with a selection of mixed frequency lexical items in semantic topics at least, I analysed the content words from the first three mini-lessons of the Food topic (Table 1); this was the fourth topic a learner would encounter, and thus could be assumed to be at an elementary level. Yet, after selecting a French language corpus to use with VocabProfile on the LexTutor website (<u>http://www.lextutor.ca</u>), I found only one type (*eau*) at the first thousand word family level, while one (*fraise*) was at the ten thousand level, one (*bière*) was apparently off-list, and the remainder were from the second to the eighth thousand level.

Table 1:

Selected food words from Duoling	0
----------------------------------	---

Mini-lesson 1	Mini-lesson 2	Mini-lesson 3		
viande, tomate, sandwich,	bois, eau, alcool, oeuf,	boivent, riz, pain, vin,		
baguette, soupe, fraise	oignon, salade	lait, bière		

Interference, on the other hand, means that "The items in a language course should be sequenced so that items which are learned together have a positive effect on each other for learning, and so that interference effects are avoided" (Nation & Macalister, 2010, p. 48) and to check this the target items from a grammatical topic, Conjunctions 2, were examined. As shown in Table 2, the apparent similarity of meaning of some of the target items—notably *lorsque, alors que, pendant que*—has the potential to make learning, and distinguishing, the items more difficult than if they were presented separately.

Table 2:

Conjunctions from Duolingo					
Target items	English				
	translation				
sinon	if not				
parce que	because				
lorsque	when				
alors que	while				
pendant que	while				
dès que	as soon as				
autant que	as much as				

As mentioned earlier, language-focused learning was clearly evident in Duolingo, and was the only one of the four strands relevant to Quizlet. As with all the strands, it exists when certain conditions are met (Nation, 2007) and these, for language-focused learning are:

- It is learner-centred
- It is meaningful
- There are opportunities to notice and compare
- There are opportunities for receptive retrieval
- There are opportunities for productive retrieval

One strength of Quizlet over Duolingo is that it more obviously satisfies the 'meaningful' condition, at least when learners create their own sets of word cards; then they are choosing words and phrases that seem worth learning to them, that are, in other words, personally meaningful. As examples, one word I chose to enter into the app was *aleatoire*, because I liked the sound of it and hoped to find an opportunity to use it; the phrase *il semble avoir* I entered for the structure as much as for the meaning.

Evaluation of the learning experience

As a language learner I found the experience of using these two apps to be a positive one. The goal was to improve my French ahead of a holiday in France, and I felt this was achieved. Existing knowledge of the language had been activated prior to departure from New Zealand, some new learning had occurred, and I felt that I had made gains in fluency, particularly listening fluency, and particularly through Duolingo. Very occasionally during the holiday there were moments when I realised an immediate language need was being met through learning from one or other of the apps, but to quantify such learning would be very difficult.

What was it about the apps that contributed to this positive experience? What principles were being operationalised that contributed to what I viewed as a successful learning experience?

The first would be motivation. While I clearly brought my own motivation, both intrinsic and integrative, to the language learning, the gamification features of both apps certainly contributed. In Duolingo these included such features as being able to set daily targets and receiving daily reminders to maintain your learning streak; I admit I did feel disappointed when international travel broke an 83-day streak. In Quizlet the different ways in which you engage with the cards you create, such as timed matching exercises, are the most obvious motivating feature.

Through use of the apps, other principles are triggered including time on task—at even five minutes a day, an 83-day streak equates to 415 minutes of focused language learning activity—and spaced retrieval through, for example, returning to the Quizlet words cards on a regular or irregular basis. These are principles that result from use of the apps, rather than being inherent in their design.

The other key principle at work is that of deliberate learning, the idea that "[t]he course should include language-focused learning in the sound system, vocabulary, grammar and discourse areas" (Nation & Macalister, 2010, pp. 57-58). With Duolingo in particular, this links to the third of the conditions for language-focused learning, where noticing target features is achieved through bolding.

Table 3 summarises the language learning principles that are present (shown by a tick) and contributed to the positive learning experience with one of or both the apps, and provides a response to the first of the questions this article seeks to address. Two of the principles in the Table have not been discussed in any detail, but suffice to say that both apps provide immediate feedback, and that Duolingo, through its design as a course, allows the user to keep progressing. Principles that result from learner use, such as time on task, are not included here, and a question mark has been used to indicate cases where it was either unclear or dependent on learner use for the activation of a principle included here.

Tabl	e 3:
1 401	v <i>J</i> .

Principles identified in Duolingo and Quizlet

	Four strands	Fluency	Frequency	Interference	Motivation	Deliberate learning	Spaced retrieval	Keep moving forward	Feedback
Duolingo	?		X	X			?		
Quizlet	n/a	\checkmark	?	?	\checkmark	\checkmark	\checkmark	Х	\checkmark

Pedagogical implications

The second question this inquiry sought to answer was the ways in which selected apps—Duolingo and Quizlet—could be used to complement in-class language learning. Based on my experience, they certainly have potential to support what happens in the classroom. In particular, they provide opportunities to develop fluency with the target language, and this has been a lack noted in coursebooks. They are, therefore, tools that teachers should make themselves familiar with so that they can then advocate their use to learners in an informed fashion. As Godwin-Jones (2017, p. 10) points out, this is "the kind of information students will need for language maintenance and future language study". Not only should teachers advocate, but they can also encourage learners to try using different apps and to share their experiences.

These two apps do, however, serve different purposes. Quizlet is obviously a vocabulary learning tool that can be linked to teacher-determined lexical items that learners can then learn in their own time, but it also has the advantage of allowing learners to select words and phrases that seem important, thus enhancing the meaningfulness of the learning. A risk, though, is that learners create such long lists

that they become unmanageable, or daunting. The teacher, therefore, needs to be able to provide advice on ways in which the vocabulary can be organised.

Duolingo is rather different in that it forms a course. Learners may not wish to start from the beginning (although starting there provides opportunity for revision) and they do have the option to 'test out' of different levels so that they can begin at a level more suited to their proficiency. For the teacher, however, perhaps the most exciting potential is in the ability to register a whole class, monitor their use of the app, and set targets. This may be easier with a class that shares a common first language, but as Duolingo caters to English language learners from many language backgrounds, this should be possible when multiple first languages are represented in a class. While making Duolingo use a whole class activity does shift the app use away from learner control and towards teacher direction to some extent, it is a way of promoting out of class use and the possibility at least that learners will continue their language learning after the formal learning period ends. Munday (2016), for instance, reports favourably on the use of Duolingo with university level Spanish language learners.

However, it must be stressed that neither of the apps discussed here provide an effective counterbalance for the principles found lacking, or under-represented, in coursebooks. This, therefore, re-emphasises the important role the teacher has to play in applying principles that promote effective language learning in the classroom. The teacher can do this through adapting, replacing, omitting, and adding to what is presented in the coursebook.

Caveats and limitations

My experience of these apps was as a language learner with some proficiency in the target language, and thus, with Duolingo, I was not unduly fazed by meeting new lexical items or structures. It is an open question how a learner embarking on learning a new language without classroom support would cope. This is a question worthy of investigation.

My experience was also that of a foreign language learner—French in largely Anglophone New Zealand—rather than as a second language learner—English in New Zealand. I did note that my use of Duolingo quickly ended once in France, although I continued to add to the Quizlet wordlist. There are various possible explanations as to why my engagement with Duolingo ended at that point: it had achieved its purpose; my proficiency met my needs; there was ample opportunity to use and practise the language on a daily basis; I had changed my approach to learning, and was practising extensive reading with newly-purchased books. However, the extent to which an app such as Duolingo works with second language learners in a target language setting is also a question worthy of investigation.

Concluding thoughts

Technological innovations keep coming, and just as we learn to incorporate them into our personal lives so must we think about how they can affect our professional lives. As part of their ongoing professional development language teachers should, then, develop familiarity with new technology in order to guide learners in its use. Apps and their contribution to out-of-class learning are certainly an area in which language teachers should develop such familiarity and expertise. MALL may not be *the* future, but it does offer an additional way of assisting learners achieve their language learning goals, whatever they may be.

Language learning apps do, however, share some of the shortcomings in terms of the application of language learning principles that characterise coursebooks, which only serves to reinforce the important role that the language teacher can play in the classroom. Teachers should be familiar with research about effective language teaching so that they can ensure learners' time in class is spent optimally, and that learners engage fruitfully in language learning activities beyond the classroom. To be able to fulfil these responsibilities requires a commitment to on-going professional development, both formal and informal.

References

- Benson, P. (2011). Language learning and teaching beyond the classroom: An introduction to the field. In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 7 16): Palgrave Macmillan
- Benson, P., & Reinders, H. (Eds.). (2011). *Beyond the language classroom*: Palgrave Macmillan.
- Cotterall, S. (2000). Promoting learner autonomy through the curriculum: principles for designing language courses *ELT Journal*, *54*(2), 109 117.
- Day, R. R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, 14(2), 136-141.
- Ellis, R. (2005). Principles of instructed language learning. System, 33, 209-224.
- Godwin-Jones, R. (2010). From memory palaces to spacing algorithms: Approaches to second-language vocabulary learning. *Language Learning & Technology*, 14(2), 4-11.
- Godwin-Jones, R. (2017). Smartphones and language learning. *Language Learning & Technology*, 21(2), 3-17.
- Hubbard, P. (2008). CALL and the future of language teacher education. *CALICO Journal*, 25(2), 175-188.
- Krashen, S. (1985). The Input Hypothesis: Issues and Implications. London: Longman.
- Krashen, S. (2014). Does Duolingo 'trump' university-level language learning? *The International Journal of Foreign Language Teaching*, 9(1), 13-15.
- Kukulska-Hulme, A., Norris, L., & Donohue, J. (2015). *Mobile pedagogy for English language teaching: A guide for teachers*. London: British Council.

- Liu, P.-L. (2016). Mobile English vocabulary learning based on concept-mapping strategy. *Language Learning & Technology*, 20(3), 128-141.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Macalister, J. (2016). Applying language learning principles to coursebooks. In W. A. Renandya & H. P. Widodo (Eds.), *English language teaching today: Linking theory and practice* (pp. 41-51). Switzerland: Springer.
- Munday, P. (2016). The case for using Duolingo as part of the language classroom experience. *RIED. Revista Iberoamericana de Educación a Distancia, 19*(1), 83-101.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Boston, Mass.: Heinle and Heinle.
- Nation, I. S. P., & Macalister, J. (2010). *Language Curriculum Design*. New York and London: Routledge/Taylor & Francis.
- Nation, P. (2007). The Four Strands. *Innovation in Language Learning and Teaching*, *1*(1), 1-12.
- Newton, J., Yates, E., Shearn, S., & Nowitzki, W. (2010). *Intercultural communicative language teaching: Implications for effective teaching and learning*. Wellington, New Zealand.
- Nunan, D., & Richards, J. C. (Eds.). (2015). *Language learning beyond the classroom*. New York: Taylor & Francis/Routledge.
- Pegrum, M. (2014). Mobile Learning. London: Palgrave Macmillan UK.
- Salaberry, M. R. (2001). The use of technology for second language learning and teaching: A retrospective. *The Modern Language Journal*, 85(1), 39-56.
- Sauro, S. (2016). Does CALL have an English problem? Language Learning & Technology, 20(3), 1-8.
- Stockwell, G. (2010). Using mobile phones for vocabulary activities: examining the effect of the platform. *Language, Learning & Technology, 14*(2), 95-110.
- Stringer, L. (2016). English outside the classroom: Using Duolingo to practice skills. Paper presented at the 12th Annual CamTESOL Conference, Phnom Penh, Cambodia.
- Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal*, *6*(1), 68-83.
- Vesselinov, R., & Grego, J. (2012). Duolingo effectiveness study. http://static.duolingo.com/s3/DuolingoReport_Final.pdf.

CAN IMAGE SCHEMATA HELP L2 LEARNERS TO DIFFERENTIATE BETWEEN SEMANTICALLY SIMILAR WORDS? THE CASE OF BASIC AUDITORY VERBS *LISTEN* AND *HEAR*

Masanobu Sato Shobi University, Kawagoe, Japan

Abstract

The difference between listen and hear is not clear-cut to many Japanese university students, often causing lexical confusion (Sato & Tanaka, 2015). These verbs are part of the basic lexicon, or a set of basic words that learners are exposed to at the early stages of L2 language development, and yet, the ability to use those words fully and differentially is hard to obtain. This study examines the possibility of using image schemata for disambiguating the already learned words in the minds of Japanese university students. More specifically, this study compared a traditional way of lexical teaching, translation-based instruction (TBI), with a new way, schema-based instruction (SBI), with respect to the pedagogical power of differentiating listen and hear. Two groups of university students (N=28) in the same school participated in this study: the SBI group (n=13) and the TBI group (n=15). The study was conducted within a pretest-posttest design. The results showed that the SBI group performed as well as the TBI group, and positive comments from the SBI group were obtained. Overall, this study showed that the use of image schemata has a strong potential for L2 lexical teaching.

Introduction

As Nation (2013) suggests, lexical learning is "not a goal itself, but without it, learners will not be able to listen, speak, read or write in the target language" (p. 362). Lexical competence has an important role in almost all areas of language learning (Schmitt, 2000). However, many Japanese language learners feel that their lexical competence is 'not good enough' (Shirai, 1990; Sato & Tanaka, 2015). Tanaka (2012) makes a distinction between basic lexical competence and extended lexical competence, and suggests that students need to develop their basic lexical competence, which should be the foundation of L2 lexical competence. In this paper, I focus on what is called "basic words," which appear in the earliest stages of L2 development and are used most frequently. More specifically, I am concerned with the usages of *listen* and *hear*.

Many Japanese university students admit that they are familiar with *listen* and *hear*, and consider them among the "easiest words." However, when they carry out sentence completion tasks they soon realize that distinguishing between these words is far from easy. Sato and Tanaka (2015) informally asked 97 university students taking a

Language and Communication class at Keio University to select the correct word or phrase in the following sentences.

- 1. Bill (listened/heard) carefully, but didn't (hear/ listened to) anything.
- 2. Can you (hear / listen to) the music coming from outside?
- 3. You (heard / listened to) me wrong.

Of the 97 university students, 47 students gave wrong answers for the first sentence; 85 students chose the wrong answer 'listen to') for sentence (2); and 67 students chose the wrong answer ('listen to') for sentence (3). For the sentences in (2) and (3), one may argue that the errors are due to the exemplar effect. In other words, 'listen to (the) music' and 'listen to me' are two major exemplars in the minds of Japanese students. In other words, given "(the) music" or "me," the verb 'listen to' is automatically triggered. The exemplar effect is, however, weak in sentence (1), where about half of the students filled in the blanks wrongly. During post-task interviews, many students admitted that they were not able to distinguish *hear* from *listen*. Thus, we assumed that the lexical differentiation between *hear* and *listen* is not clearly made in their mental lexicon.

Adult second language learners tend to use their first language (L1) to learn the meaning of an L2 word. The use of this strategy, sometimes called "search-translation-equivalent strategy" (Tanaka & Abe, 1985), is particularly dominant when second language learning takes place in input-poor contexts. Japanese students generally take it that *hear* and *listen* mean *kikoeru* and *kiku* at the lexeme level, respectively. However, at the usage level, the lexical correspondence between *hear* and *kikoeru* often does not work; 'I heard the story from John' is translated into 'watashi wa John kara sono hanashi o kiita' (the word *kiita* is the past form of *kiku*). Thus, Japanese *kiku* corresponds to both *listen* and *hear* which is partly responsible for the lexical confusion as observed in the above fill-in-the-blank test.

Thus, the use of search-translation-equivalent strategy is problematic when it comes to the task of lexical differentiation. In recent years, we have witnessed the rise of cognitive linguistics, which highlights the semantically-motivated aspects of a language, and we have theoretical constructs that explain a language in a meaningful way, including embodiment, perceptual salience, a usage-based schema, and a prototype. Researchers in second language learning and teaching can benefit from cognitive linguistics, especially because semantic motivation is a concept directly linked with meaningful learning (Tyler, 2008). This paper focuses on the notion of image schema (Johnson, 1987; Lakoff, 1987) and discusses its role in helping students re-adjust their L1-based understandings of L2 lexicon.

According to Johnson (1987), an image schema has experiential bases and emerges as a generalized image. The preposition *in*, for example, is used in a variety of situations

and, on the basis of different usages, the container-like image schema emerges within an individual. A sentence like "John is in the kitchen" is a good example of the image schema "container." The kitchen refers to a physical space. When someone says, "John is in the army," the army is not a container in a clearly delineated fashion, but we can consider it a social space. Even a sentence like "John is in love" can be conceived as John being in a psychological space named "love." In addition to "John is in love," we have expressions like "John fell in love" and "John fell out of love," which suggest that "love" is conceptualized as a container or psychological space. The usage "in the kitchen" is the prototype exemplar of the image schema of *in*; the same schema applies to "in the army" and "in love." An image schema is a usage-based concept (Tomasello, 2003; Taylor, 2012). The idea of usage-based schema or schema abstraction applies not only to prepositions but also to basic verbs.

An image schema, or a usage-based image schema, has been introduced into the field of second language teaching (Tyler, 2008; Ellis, 2013). As an alternative to translation-based instruction, we can turn to the concept of image schema, and use it as a means of showing the meaning of a word. This alternative approach is called "schema-based instruction (SBI)" (Morimoto & Loewen, 2007).

This study compares two types of instruction: schema-based instruction (SBI) and translation-based instruction (TBI). The question here is which type is more effective in disambiguating the selective use of *listen* and *hear* in the learner's mind.

Schema-based instruction versus Translation-based instruction

The effect of SBI has been investigated in some studies, and yet the results are not conclusive. Some studies have reported the comparative effect of SBI over TBI on L2 lexical acquisition (Morimoto & Loewen, 2007; Tyler, Mueller & Ho, 2011; Akamatsu, 2010; Cho & Kawase, 2011; Mitsugi 2010). In a study dealing with Japanese high school students, Morimoto & Loewen (2007) focused on the usages of *break* and *over*, and reported that SBI was as effective as TBI both in acceptability judgment and production tests in the case of *break*, and SBI was more effective than TBI in the case of *over* (but not *break*). Tyler, Mueller and Ho (2011) showed in a study with Italian learners of English at advanced stages that a cognitive approach using image schemata is effective in teaching English prepositions such as "to," "for," and "at." Cho and Kawase (2011) also used image schemata to teach English articles to Japanese college students, and reported that the SBI group out-performed the TBI group in a delayed posttest (8 weeks after the treatment).

However, as Akamatsu (2010) points out, the results are not conclusive. We should also point out that these studies were concerned with the superiority of SBI over TBI, and yet, they did not look at schema-based instruction vs. traditional instruction over a period of time. Instead, the two forms of instruction were compared after a single session of instruction involving the learning of two or three words.

In this study, we assume that the two words in focus (i.e., *listen* and *hear*) were introduced to university students in their early stages of language learning. In other words, university learners have had considerable previous experience with the two words including attempts by teachers to clarify their meanings in Japanese. It is therefore doubtful whether a single session of instruction can have much effect on vocabulary learning in comparison with the weight of all the previous experience.

Over a series of instructional sessions, this study investigated Japanese university students' re-adjusted understandings of two words by addressing the following research question: Does the use of image schemata help students use *hear* and *listen* in a different way?

More specifically, on the basis of the previous studies, this study tested the following hypotheses:

Hypothesis (1): In the short term, SBI is as effective as TBI in terms of raising students' awareness of the difference between *listen* and *hear*.

Hypothesis (2): SBI is better than TBI in the long-term.

As for hypothesis (1), note that the participants of this study were more familiar with TBI, and SBI was completely new to them. In order to show that SBI is also a viable approach, we felt it necessary to demonstrate that it is at least equally as effective as a conventional method (Hypothesis 1).

Hypothesis 2 was formulated because SBI presumably encourages students to understand the usages of a word through the image schema, and the schematic processing presumably makes deeper information processing possible (Gairns & Redman, 1986). In other words, SBI may lead to better memory retention.

Method

Participants

Two groups of university students (N=28) in the same school participated in this study over four weeks (four 90-minute sessions). The SBI group (n=13) and the TBI group (n=15) were considered at the same level of English proficiency (the lower intermediate level within the university) on the basis of the TOEIC scores that they had taken according to a school policy. The average score of the students belonging to this level was about 410 (or the A1 level on the CEFR scale). The author taught the two classes and considered that both were comparable in terms of their level of proficiency.

Tests

Three kinds of tests were devised: a pretest, posttest-1, and posttest-2. To identify the multiple senses or semantic types of *hear* and *listen*, I referred to the Oxford Advanced Learner's Dictionary and Merriam-Webster Dictionary, and used four semantic types for *listen* and five semantic types for *hear*. The pretest consisted of 14 items in total, of which seven items required *hear* as the correct answer, five items required *listen*, and the remaining two (requiring *sound*) were distractors. An example is given in Figure 1:

EXAMPLE **Context**: You're talking to Tommy on the mobile phone. You'd like to know if your voice reaches him all right. Tommy, can you (1. listen to 2. hear 3. sound) me? **The level of confidence** weak 1 ------ 2 ------ 3 ------ 5 strong

Figure 1:

An example of an item from the pretest

The participants were also asked to rate their level of confidence on a five-point Likert scale in choosing the answer for each item, with 1 being the weakest and 5 being the strongest.

Posttest-1, consisting of 14 items in total, was designed exactly in the same format as the pretest, but the test sentences for each verb (*listen* or *hear*) were novel and produced within the same semantic types. Posttest-2 (i.e., the delayed posttest) was prepared in the same format and had 28 items, containing all the items presented in the pretest (14 items) and posttest-1 (14 items). The 28 items consisted of 14 items for *hear*, 10 items for *listen*, and 4 distractors for *sound*. The numbers were not even between "listen" and "hear" because the degree of polysemy was different between the two verbs, with *hear* being slightly more polysemous. Posttest-2 intended to measure the memory retention of the two previous tests.

Image schemata of listen and hear

This study dealt with the verbs of auditory perception: *listen*, and *hear*. To practice SBI in class, we used the following image schemata of *listen* and *hear* based on the *E*-*Gate English-Japanese dictionary* (Benesse Corporation):



Figure 2:

The image schemata of *listen* and *hear* (Permission to reproduce these has been obtained from the copyright holder).

These schematic representations are verbally explained. That is, the verb *listen* involves "paying auditory attention (to capture a sound)", whereas the verb *hear* implies "capturing something with the hearing organ," according to E-Gate English-Japanese dictionary. Characterizing the meaning of *listen* this way helps to explain why *listen* can be used as an attention-getter or why it often requires the preposition *to*, which indicates the target to which auditory attention is directed. On the other hand, the verb *hear* means "capturing something with the hearing organ". Thus, a sentence like 'You heard me wrong' requires *hear*, not *listen*. The difference of the two verbs is well illustrated in the following sentence: "I listened hard, but didn't hear anything".

Procedure

In the pretest, the participants were asked to choose either *listen* or *hear* or *sound* depending on the context given in Japanese. The same pretest was given to the SBI and TBI groups.

A week after the pretest, the SBI group received the treatment of schema-based instruction, and the TBI group, the treatment of translation-based instruction. Both groups were taught the same items from the pretest. Each session was 90-minutes long. Each group was asked to think why a given test item requires "listen" or "hear." With the TBI group, the instructor (the researcher) explained how to use "listen" or "hear" in each context using the translation method, and encouraged the students to memorize the examples along with their Japanese translations. With the SBI group, the instructor (the researcher) introduced the image schemata of "listen" and "hear" as described above, and encouraged the students to establish a mental link between the schematic image and usage. Establishing a mental link between an example sentence and the image schema was the main task for the students. For example, every time *listen* was used, the students gave their own ideas about how the particular use of *listen* could be linked with the *listen* schema, as shown below in Figure 3:

The Listen Schema



The Usage of Listen

Figure 3: The usage of *listen* and the *Listen* Schema

With this exercise, the students were expected to mediate the English sentence and its Japanese counterpart through the *listen* schema. The students in the TBI group did not have any pictorial devices. Instead, they were asked to memorize how to translate a Japanese sentence into English.

The students in both groups were encouraged to ask questions when they found it difficult to give the right answer. In the SBI group, the teacher answered the question with reference to the schemata, while in the TBI group, the teacher answered it in Japanese as follows: In a situation as described in the Japanese sentence, we use *listen to* and say, "Don't listen to such a guy." Just memorize this expression.

One week after the treatment, posttest-1 was given to the two groups. Finally, one week after posttest-1, posttest-2 (a delayed posttest) was given to the two groups.

The comparison between the pretest and posttest-1 answers intended to demonstrate that in the short term, the schema-based instruction is as effective as the translation-based instruction in terms of raising students' awareness about the difference between *listen* and *hear*. We had a "stronger" hypothesis that the effect of the schema-based instruction lasts longer than that of the translation-based instruction. To test this stronger hypothesis, we administered posttest-2 one week after the posttest-1, with no prior announcement.

Results

The overall results (means and SDs) are shown in Table 1. The items relating to *sound* were excluded from the analysis since they were introduced as distractors. As a result, the pretest and posttest-1 had 12 items, and posttest-2 had 24 items. Because of the difference in the number of test items, Table 1 shows the percentages of correct answers as well as the raw scores. The mean scores of confidence level were also included in the table.

in three tests								
		pretest	posttest-1	posttest-2	pretest	posttest-1	posttest-2	
Treatme	ent	correct %	correct %	correct %	confidence	confidence	confidence	
SBI	Means	.368	.604	.650	2 60	2 21	2.26	
(n=13)		(4.4)	(7.2)	(15.6)	2.09	3.21	5.50	
	SDs	.120	.139	.169	.407	.752	.596	
TBI	Means	.400	.588	.586	2.54	2.07	2.02	
(n=15)		(4.8)	(7.1)	(14.1)	2.34	2.97	5.05	
	SDs	.111	.223	.224	.345	.598	.722	

Table 1:Descriptive statistics: Percentage of correct answers, confidence scores and their SDsin three tests

Note: Numbers in parentheses are raw scores.

A two-way ANOVA was used to test differences between the two groups (betweensubject factor), and between the three tests (within-subject factor). As for the use of an ANOVA with small sample sizes, Norman (2010, p. 629) notes that "empirical studies of robustness of ANOVA date all the way back to Pearson (1931) who found ANOVA was robust for highly skewed non-normal distributions and sample sizes of 4, 5 and 10" (for a similar view, see Gordon & Johanson, 2011). For a two-way ANOVA, the effect size was calculated using a measure of η^2 (eta squared). To estimate the required number of subjects, we also used G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) on the basis of the effect size ($\eta^2 = 0.463$) in this study, and obtained the minimum number 6. The Bonferroni correction method was used to counteract the problems deriving from multiple comparisons of the within-factor and the between-factor.

Overall, the main effect of the within-subject factor (i.e., pretest, posttest-1, posttest-2) was statistically significant (F(2, 52)=22.34, p < .001, $\eta^2 = .463$), while the main effect of the between-subject factor (i.e., SBI, TBI) did not reach the level of significance (F(1, 26)=.114, p > .05, n.s., $\eta^2 = .004$). More specifically, the comparison of pretest means indicates that the two groups were essentially at the same level of English proficiency at the beginning of this study (F (1, 26)=.536, p > .05, n.s., $\eta^2 = .020$). A week after the 90-minute treatment, the posttest-1 was administered to the two groups. The results from comparing the pretest and the posttest-1 showed that both groups gained from their treatment considerably (p < .05, Bonferroni correction). There was no statistically significant difference between the two groups in terms of the scores on the posttest-1 (F (1, 26)=.053, p > .05, n.s., $\eta^2 = .002$).

This supported the weak hypothesis that both SBI and TBI are effective ways of teaching the difference between *listen* and *hear* in the short-term. In the short-term (that is, one week after the treatment), the students in the SBI group did equally well as those in the TBI group.

One week after posttest-1, the delayed posttest (posttest-2) was administered to test the second hypothesis that SBI is better than TBI in the long-term. There was no statistically significant difference between the two groups (F (1, 26)=.702, p > .05, n.s., $\eta^2 = .026$). Thus, the stronger hypothesis was not supported in this study. However, we must note that the SBI group showed an increase from 60.4 percent to 65 percent (i.e., 4.6 percent increase), while the TBI group leveled off between 58.8 percent and 58.6 percent. The previous studies showed a considerable decrease in the delayed posttest, as compared with the posttest-1 (Mitsugi, 2010; Morimoto & Loewen, 2007). In this study, I observed an increase in the delayed test, which suggests that the use of image schemata contributes to the learning effect on lexical memory retention. I speculated that increased test performance was obtained in the delayed test because 90 minutes was spent on instruction (much longer than the previous studies). Considering the small number of participants in this study, however, further research should be carried out with a larger sample to show this learning effect of image schemata.

As regards the level of confidence in answering each question on the Likert scale, the SBI group (pretest: 2.7; posttest-1: 3.2; posttest-2: 3.4) generally showed a higher confidence level than the TBI group (pretest: 2.5; posttest-1: 3.0; posttest-2: 3.0), as shown in Figure 4.



Figure 4: Levels of confidence in answering questions

The main effect of the within-subject factor was statistically significant (F (2, 52) =14.62, p<.001, $\eta^2=.463$). Thus, both SBI and TBI had a significant effect on improving the learner's level of confidence when answering questions. The main effect of the between-subject (SBI vs. TBI) factor did not, however, reach the level of significance (F (1, 26)=1.82, p>.05, $\eta^2=.065$). The comparison between the two groups was not statistically significant either in the pretest (F (1, 26)=1.18, p>.05, n.s., $\eta^2=.043$), posttest-1 (F (1, 26)=0.95, p>.05, n.s., $\eta^2=.035$) or posttest-2 (F (1, 26)=1.67, p>.05, n.s., $\eta^2=.060$). However, considering that the use of image schemata is new to

the students, their responses to the confidence scale seems to suggest that if the students become adept in using schemata as their learning strategy, then they could feel more confident in using *listen* and *hear* selectively.

Discussion

In this section, let us first take a look at some of the test items. The same 12 items were included both in the pretest and in the posttest-2. Of the five *listen*-related items, the following showed fairly big differences between the two tests within the SBI group.

- Don't listen to such a guy. (SBI: $31.0\% \rightarrow 85\%$; TBI: $31\% \rightarrow 63\%$)
- Listen, I got something to tell you. (SBI: $23\% \rightarrow 77\%$; TBI: $13\% \rightarrow 56\%$)
- She listened for the steps of her husband. (SBI: $15\% \rightarrow 61\%$; TBI: $24\% \rightarrow 56\%$)

It is likely that the students' knowledge about the use of *listen* was limited initially, and yet, after the treatment of SBI, they became capable of applying the schema to more *listen*-related situations. When asked to produce sentences using *listen*, most Japanese produce the collocation "listen to music," which is their prototype exemplar. Presumably, the participants in this study were not familiar with the examples mentioned here. Though the TBI group also did better on the posttest-1, the amount of increase was much greater with the SBI group for the three test items above. When the schema of *listen* was presented in class along with these examples, the students in the SBI group listened to my explanation with a big nod, which may suggest that they understood the link between the schema and the examples.

Of the seven *hear*-related items, the SBI group did considerably better on the following three items than the TBI group.

I've heard the news from her. (SBI: $30\% \rightarrow 77\%$; TBI: $40\% \rightarrow 63\%$) Ned heard who was speaking at the meeting (SBI: $34\% \rightarrow 78\%$; TBI: $46\% \rightarrow 69\%$)

I heard my father snoring in the other room. (SBI:16% \rightarrow 92%; TBI:25% \rightarrow 75%)

The schema of *hear* suggests capturing something with one's ears. In relation to this, I explained in class that the target of hearing is not syntactically restricted to a noun phrase, but includes clauses and constructions such as NP + present participle / bare infinitive, and that the core schema applies to these cases as well. In dealing with "I heard my father snoring in the other room," students understood that the target of hearing was 'my father snoring in the other room'.

There were test items which showed decreases in the scores between the pretest and the posttest(s). The item: "Belle heard about the accident the next morning" was a case in point. In the pretest, the SBI group scored high (89%), and yet, two weeks after the treatment of SBI, their performance declined (64% in posttest-2). The performance of the TBI group was stable between the two tests. The author speculates that the image schema of *hear* had a negative effect because the image schema only applies to the transitive usages of *hear* and there is no good fit between the schema and the intransitive "hear about." This study dealt with one more example of the intransitive use of hear: "She heard from her mother." This example was used in posttest-1 and posttest-2, with the results of 53 percent in posttest-1 and 39 percent in posttest-2. Here again, the image schema presumably operated negatively in these cases. This suggests that the generalizing power of the schema can be constrained by the transparency of schema projection, or the ease of applying the schema to exemplars.

I did a follow-up interview with the participants in the SBI group. All the participants agreed that the schematic comparison of *hear* and *listen* helps them differentiate the two verbs in question. All of them reported that SBI made them feel more comfortable and confident as far as the difference between the two verbs is concerned. In addition, the participants gave their open-ended comments in Japanese including the following (translated by the author):

"I wanted to learn English this way while I was a junior high school."

"For the first time, I was able to see the difference between *listen* and *hear*."

"Pictures [i.e., image schemata] come to me more naturally than words."

"I feel I can use these words more confidently from now on."

"Prior to the instruction, I didn't know that *listen* and *hear* are totally different words."

As the instructor in this study, I felt that the participants in the SBI group were more engaged than those in the TBI group. I also felt that more learning took place in the SBI group. The TBI group did not comment on the instructional method, but gave more free opinions such as the following:

"These words are much more difficult than I thought."

"It is difficult to learn even these simple words."

"You have to memorize a lot of examples to use these words."

All in all, the results seem to suggest that the use of image schemata has the potential to be a viable alternative to translation in teaching and learning L2 vocabulary, although this study was not able to demonstrate empirical support for this suggestion. As both researcher and instructor, one limitation of this research is instructor bias, particularly as I favoured the schema-based method.

Verspoor and Lowie (2003) did a study with Dutch students and suggested that the use of the core sense promotes both better guessing and long-term retention of figurative senses of polysemous words. They claimed that the core sense helps students develop a "precise elaboration." An image schema is an L1-independent device, which intuitively captures the core sense of a basic word. If so, it is possible that an image schema functions as a "device" for cognitive re-adjustment or "precise elaboration" (Vespoor & Lowie, 2003). As explained above, both translation-based instruction and the search-translation-equivalent strategy lead to the "listen=kiku" formula, which encourages the learner to understand the meaning of *listen* through the L1 grid. However, kiku is the translation equivalent of hear as well. Thus, interlingual matching does not help learners use *listen* and *hear* differentially. If the learner's *"listen=kiku"* formula is readjusted through the image schema of *listen*, then there is a possibility that the image schema functions in such a way as to minimize the negative effects of semantic discontinuity and semantic circularity triggered by a translationbased learning. Vocabulary instruction that stimulates cognitive engagement with the word in question promotes retention. In this sense, SBI is a strong possibility when we teach vocabulary in an input-poor environment.

Conclusion

Both *listen* and *hear* are usually taught at junior high school level. Presumably all students who took part in the study have been taught these two words in the past by TBI. So what I did in this study was not the teaching of new vocabulary but an attempt to disambiguate already taught vocabulary in the minds of the learners.

The comparison between SBI and TBI did not show a clearly significant difference. However, the overall results of this study seem to suggest that the use of image schemata is a promising way of re-adjusting the learner's mental representation of already taught vocabulary. To support this suggestion, we need to further study how to introduce image schemata in classroom teaching. One possibility is that the teacher encourages students to attempt to find a schema underlying different usages of a word in a collaborative, active learning context. Another possibility is that the teacher deductively gives the students the image schema of a word when the word is introduced in order to discourage students from formulating a one-to-one interlingual correspondence as their initial hypothesis. The big issue here is not only what to present to students, but also how to present it optimally.

References

- Akamatsu, N. (2010). Difficulty in restructuring foreign language vocabulary knowledge: polysemous verbs. *JACET Kansai Journal*, 12, 68-79.
- Cho, K., & Kawase, Y. (2011). Effects of a cognitive linguistic approach to teaching countable and uncountable English nouns to Japanese learners of English. *ARELE*, 22, 201-215.

- Ellis, N. (2013). Second language acquisition. In G. Trousdale & T. Hoffmann (Eds.), *Oxford Handbook of Construction Grammar* (pp. 365-378). Oxford: Oxford University Press.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.
- Gaims, R., & Redman, S. (1986). Working with words: A guide to teaching and learning vocabulary. Cambridge: Cambridge University Press.
- Gordon P. B., & Johanson, G. (2011). Sample size considerations for multiple comparison procedures in ANOVA, *Journal of Modern Applied Statistical Methods*, 10, 97-109. Johnson, M. (1987). The body in the mind: The bodily basis of meaning, imagination, and reason. Chicago: University of Chicago Press.
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination and reason.* Chicago: University of Chicago Press.
- Lakoff, G. (1987). Women, fire, and a dangerous thing: What categories reveal about the mind. Chicago: University of Chicago Press.
- Mitsugi, M. (2010). The effectiveness of core meaning based instruction on preposition choice. *Research Bulletin of English Teaching*, 10, 1-25.
- Morimoto, S., & Loewen, S. (2007). A comparison of the effects of image-schemabased instruction and translation-based instruction on the acquisition of L2 polysemous words. *Language Teaching Research*, 11, 347-372.
- Nation, P. (2013). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15, pp. 625-632.
- Pearson, K. (1931). Historical note on the distribution of the standard deviations of samples of any size drawn from an indefinitely large normal parent population. *Biometrika*, 23, 416-418.
- Sato, M., & Tanaka, S. (2015). L2 lexical development and the lexical network model: The case of basic verbs of perception. *International Journal of Applied Linguistics*, published online for early view
- Schmitt, N. (2000). Vocabulary in language teaching. Stuttgart: Ernst Klett Sprachen.
- Shirai, Y. (1990). Putting PUT to use: Prototype and metaphorical extension. *Issues in Applied Linguistics*, *1*, 78-97.
- Tanaka, S., & Abe, H. (1985). Conditions on interlingual semantic transfer. In P. Larson, E. Judd & D. Messerschmitt (Eds.), ON TESOL '84: A Brave New World of TESOL (pp.101-120. Washington, D.C.: TESOL.
- Tanaka, S. (2012). New directions in L2 lexical development. *Vocabulary Learning and Instruction*, *1*, pp. 1-9.
- Taylor, J. (2012). *The Mental corpus: How language is represented in the mind*. Oxford: Oxford University Press.

- Tomasello, M. (2003). *Constructing a language*. Boston, MA: Harvard University Press.
- Tyler, A. (2008). Cognitive linguistics and second language instruction. In P. Robinson & N. Ellis (Eds), *Handbook of cognitive linguistics and second language acquisition* (pp. 456-488). New York: Routledge.
- Tyler, A., Mueller, C., & Ho, V. (2011). Applying cognitive linguistics to learning the semantics of English to, for and at: An experimental investigation. *Vigo International Journal of Applied Linguistics*, *8*, 180–205.
- Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, 53, 547-586.

CAN ADULT BEGINNER ENGLISH LANGUAGE LEARNERS MANIFEST SIGNS OF A READING HABIT?

Jenny Field Waikato Institute of Technology

Abstract

Extensive reading is the reading of large amounts of text that is within the learner's current capability. Research into reading in a second language with adult beginner language learners and limited literacy learners is scarce. The study reported here occupies this space by focusing on a classroom-based extensive reading project which enquired about early L2 readers' behaviours and feelings about L2 reading. During the project students had a short period of silent reading on most days having selected their reading material. Eighty students then completed a checklist about what they had read in class and out of class. Pre-and post-interviews with nine of the students enquired further about their enjoyment of reading, and any constraints they were encountering when reading, in an effort to establish whether at these early stages learners were forming a reading habit. Results seem to support Day and Bamford's (1998) 'book strap' hypothesis whereby success in reading increases the motivation to read more.

Introduction

Adult second language learners bring a range of experiences and strategies to their learning to achieve their life goals. Reading is one of the platforms on which they will develop their learning, and teachers try to ensure that there is ample and level-appropriate material available to them to read both in paper-based form and through digital means. Teachers of beginner second language readers know that students will develop well in reading if they develop a habit of reading extensively, but may feel uncertain how to best foster this habit.

Scales and Rhee (2001) define the concept of a reading habit as "how often, how well and what adults like to read" (p. 178). More than a decade later Rodrigo, Greenberg and Segal (2014) conducted a study at Georgia State University with low literate adult beginner readers of English on an extensive reading programme. The study strongly suggested that "adult learners can benefit from extensive and pleasure reading and that a well-equipped library, easy access to books and encouragement and time to read are all key factors in the development of reading habits" (Rodrigo, Greenberg & Segal, 2014, p. 73). The present study explores the possibility that even beginner readers of a second language may start to develop a habit of reading after encountering some success with the scaffolded experience.

Literature Review

Literature focusing on early readers' fluency in the second language seems much more limited than studies about more advanced students. This may be because the prevailing attitude is that there is some sort of threshold or language competence ceiling to be attained before existing reading abilities in the reader's first language can begin to be transferred as suggested by Alderson (1984). Teaching reading is based on gaining meaning from the text (Freeman & Freeman, 2003) and, for readers, "successful initial experiences in extensive reading result in the discovery that they can read in a second language and that it is rewarding and pleasurable" (Day & Bamford, 1998, p. 20).

The idea of reading easily understandable material is based on Krashen's comprehensible input hypothesis which held that the language encountered by the learner should be understandable by the learner. Krashen (2004) claimed that "light reading is a bridge to heavier reading, as it supplies the competence necessary to understand challenging texts." However, the focus for most reading teachers in the second language classroom is a process of intensive reading which can lay a platform for extensive reading. Nation (2004) says that intensive reading involves the detailed reading of texts "with the two goals of understanding the text and learning language features through a deliberate focus on these items" (p. 20). The text chosen usually has some features which are unknown to students and they can be scaffolded to develop skills to learn the new vocabulary, syntax or discourse features.

According to Nation (2015), one benefit of including extensive reading (ER) into an L2 language learning programme is to consolidate and sustain vocabulary growth. He argues that vocabulary learning depends on the number of meetings with each word (quantity) and the quality of attention at each meeting. Nation (2015) claims that "the more meetings, the more likely learning is to occur" (p. 136). He contends that an ER course situated within a well-balanced programme will provide students opportunities to meet new words in different contexts. In relation to quality he recommends dictionary use so that students pay attention to each unfamiliar word. Maley (2009) also lists a range of benefits from extensive reading. They include learner autonomy, comprehensible input, extending general and world knowledge, sustaining the process to read more, enhancing general language competence, extending, consolidating and sustaining vocabulary growth, and helping to improve writing.

One recent study which sought to explore the benefits of ER was conducted by Rodrigo et al. (2014) with adult L2 students with limited literacy. The researchers enquired about the development of reading habits over a 100-hour programme. Interventions were carried out with two groups of students; one group implemented ER while the other group had direct reading instruction only. Results showed that "the ER group which had access to books, free choice of reading material and time to read was more motivated, developed a habit of reading, and experienced a change in reading behavior" (Rodrigo et al., 2014, p. 73).

The key approach that influenced the present project arose from Day and Bamford's (1998) ideas that ER is a means for students to gain fluency and identifies with the socio-psycholinguistic view. The central concept is that when students read whole texts silently and are familiar with most of the words, they will start to read fluently and confidently. Day and Bamford (1998) called this concept the 'extensive reading book strap hypothesis.'

Students' initial successful experiences in extensive reading result in the discovery that they can read in a second language and that it is rewarding and pleasurable. These positive beginning experiences then feed back into subsequent extensive reading experiences, resulting in greater gains in reading ability and positive attitude, and increases in motivation and enjoyment (Day & Bamford, 1998, p. 30.)

When considering other contextual factors that were contributing to learners' reading experiences, I assumed that online reading would be occurring more readily these days as web-based technology is becoming increasingly available, and students are becoming normalised towards reading online. Bax (2003) predicted that this normalisation of Computer Assisted Learning would be realised when it was "taken for granted in everyday life" (p. 23). With almost all students using smart phones as a tool for learning and using social media confidently, it seemed likely that quite a lot of offline as well as online reading would be occurring in both the students' L1 and L2.

The present study

This present ER study was conducted during one academic semester over ten weeks with 80 L2 students. The study enquired about what texts students were reading and sought to discover whether these L2 students' reading behaviours were becoming habitual at these early stages. The premise was that if students were starting to form positive reading habits in the early stages, this may assist them as they start to read texts with more challenging vocabulary and syntactical structures.

The participants were drawn from four English language classes; one class was a parttime course of six hours a week, and the other students were enrolled in three New Zealand Certificate in English Language programmes (Level 1a, Level 1b and Level 2). Teachers actively promoted the concept of Extensive Reading (ER), by explaining to students how to select material that they could read and the importance of regular reading. Guidelines for selecting appropriate reading material included choosing texts where the content was interesting, most of the vocabulary was familiar, and where the text could be read easily and reasonably quickly. The learners were asked what they were reading, and if they were enjoying the reading experience, whether in print form or online. Learners could read a range of graded readers which were short, easily readable stories or easy readers organised into chapters. In their computer hour they read material from English language websites as well as appropriately graded stories and associated tasks on the FLAX website (Witten, Wu & Yu, 2011). They were also encouraged to read magazines and newspapers out of class, and to join their local public library. The study was also interested in finding out what feelings students were experiencing when reading, whether they were changing attitudes toward the process and if there were any other positive factors that influenced their attitudes towards reading.

Taking an attitude of enquiry and also considering the context surrounding these learners, these research questions were formed:

- What kinds of texts do beginner English learners select?
- As they proceeded through the project did their attitudes towards reading change?
- Were any factors in their reading experiences associated with positive changes in attitudes towards reading?
- To what extent are these beginning English language learners starting to form reading habits?

Participants and programmes

The students studied at a polytechnic in Aotearoa New Zealand, and the classes were comprised of both domestic students living here permanently and international students. Students were drawn from a range of ethno-linguistic backgrounds and had varied backgrounds in terms of their educational backgrounds. Some students had graduated from secondary school in their countries of origin, and some with refugee backgrounds had had an interrupted education, resulting in less than three to five years of formal schooling. Most of the international students entered the institution having completed secondary school in their country of origin, and having learned English as a foreign language.



Figure 1:

Students' educational backgrounds in their countries of origin

Students were informed about the research and asked for their consent to conduct the research, and a total of 80 students volunteered to participate in the study. The four teachers ran their own ER programmes and all had access to graded readers that were brought to the ER sessions.

A selection of age appropriate reading material which consisted mainly of graded readers at A1 and A2 level was provided for all classes. Graded Readers were screened to ensure they consisted of high frequency words, visuals to support the story and themes that were of interest to students. The Readers provided were similar to those recommended in the Extensive Reading Foundation for beginner and elementary readers (The Extensive Reading Foundation, 2011).

The part-time course was benchmarked at pre-Level 1 and the Level 1 and 2 NZCEL courses were benchmarked at between A1 and A2 respectively on the European Framework of Reference. Descriptors written by the Institute for Test Research and Test Development appear in Appendix 1.

Methods

The study applied a mixed-method approach. Students completed a simple tick box checklist (see Figure 2 below) each time they had completed between 10 to 15 minutes of silent reading. They were asked to indicate what type of text they had read and then to rate their enjoyment of the text on a three-point scale. The ratings were *I liked reading it, It was* OK and *I didn't enjoy it.* Teachers collected the forms each week and forwarded them to me as the primary researcher.



Figure 2:

A segment of the checklist completed by participants after their silent reading each day.

As indicated in Figure 2 above, distinctions were made between a number of types of text, each of which was easily distinguishable by students, and each of which included fiction and non-fiction texts. The category 'online story' refers mainly to texts that were made available to students through Moodle, an online learning platform that students accessed during class time and at home as part of their self-study.

Nine students (see Table 1) were randomly selected from an online random number generator and they (all) willingly agreed to be interviewed both before and after the intervention. The interviews focused on their reading habits and enjoyment of reading in English, and any difficulties encountered. Also discussed in the first interview was their prior experience of reading, particularly in their first language.

Table 1:Background of interviewees

	Nationality	Level	Age	L1 literacy	L2 literacy
S 1	Brazil	NZCEL Level 2	29	Yes	Yes
S2	China	NZCEL Level 2	50	Yes	Yes
S3	India	NZCEL Level 2	26	Yes	Yes
S4	India	NZCEL Level 1	25	Yes	Yes
S5	Thailand	NZCEL Level 1	19	Yes	Yes
S6	Congo	NZCEL Level 1	21	No	Starting to develop
S7	Iraq	NZCEL Level 1	40	Yes	Developing
S 8	Somalia	Level 1 Pt- time	35	No	Starting to develop
S 9	Somalia	Level 1 Pt-time	23	No	Developing

These interviews were captured by digital voice recorders and transcribed after both pre- and post-interviews had taken place using *Dragonware Naturally Speaking (http://www.nuance.com/dragon.html)*. In many cases when reporting the data below I have slightly adjusted the original syntax and lexis, without changing the essential meaning, in order to assist readability. I reviewed the first set of interviews and coded keywords and phrases according to the interview questions. After initially coding the items using highlighting to colour-code each student, the data was scrutinised by looking for commonalities. This process was repeated for the post-intervention interviews. The categories emerged and the data was examined again for similarities. The analysis took place following the research questions. After reading and rereading the data, key ideas were coded, underlined and grouped into categories, which were then compared with similar responses and then grouped together into themes which had emerged from the data.

Findings

Checklist results

The checklists were collected by teachers and samples have been selected from Week 3 and Week 10 from the part-time class, a Level 1 class and the Level 2 class to show how students responded to the material they were reading. In-class and out-of -class responses were combined, and these are presented in Figures 3, 4, and 5.



Figure 3:

'Like' responses to different reading texts in Week 3 and Week 10 in the part-time class.



Figure 4:

'Like' responses to different reading texts in Week 3 and Week 10 in a Level 1 class.



Figure 5:

'Like' responses to different reading texts in Week 3 and Week 10 in the Level 2 class.

The category 'story from a textbook' appeared the most popular, in each class receiving more 'like' responses than the other text types. This may be because the layouts and designs are colourful, and presumably therefore more appealing to new readers. Reading on-line (mainly through the Moodle platform) also received high ratings; this was a familiar activity as all students had two hours each week in the computer lab. Of the different text types, comics were the least read by students, perhaps due to low access.

When comparing the changes between Weeks 3 and 10, there was generally a more favourable response to reading in Week 10 than in Week 3. In particular, both graded readers and chapter books showed a small increase in 'like' responses in three of the four classes. These readers were actively promoted during the 10-week research project, so it is likely that more were read in response to the teachers' promptings. Chapter books too were read more in Week 10 than in Week 3. There was also a slight increase in favourable responses to reading magazines. The number of ER sessions remained the same, so it is likely that the participants' growing confidence in reading contributed to the increase in 'likes'. The interviews reveal a little more insight into this slight change in behavior.

The following figure presents, for each of the classes, the average weekly number of 'likes' for each of the text types.



Figure 6:

The text-types 'liked' by students when reading in and out of class.

This shows that graded reading stories were valued more highly by students in the Level 1 classes than in the Level 2 classes. However, chapter book readers received a positive response in all except the Level 1A class, which seemed to prefer the easier graded readers or graded stories. This result suggests that graded readers with adult themes have an important role in lower level classes, especially at A1 level. As the vocabulary is graded and they have visual clues to support the text, students can read them easily and gain meaning and pleasure from a successful reading experience. This in turn encourages them to read more—the "reading book strap" theory (Day & Bamford, 1998).

Students in Level 2 awarded the highest positive rating for reading text books in class with their teacher. It seemed that students at A2 level were developing their reading skills so that they feel more secure and independent with material that was intensive.

Reading magazines and newspapers received similar responses to graded stories readers and chapter books. For the part-time classes, reading magazines and newspapers were read more frequently out of class than in class. "Other" text types also rated quite highly within all classes. In addition to graded readers, teachers had extra reading material in their classes which were available to students. These included travel and food magazines and back-copies of "Password," an out-of-print magazine with graded texts, for new residents in Aotearoa New Zealand. Students reported that they enjoyed these materials and wrote the names of the texts on their checklists. This result suggests that students were starting to engage with other text types as well as their class textbooks and graded readers.

Another result yielded by the quantitative data was that students seemed to be influenced by their teachers' input in their selections. Informal feedback from teachers about their ER activities was reflected in these results. Graded readers and magazine reading were promoted by the teacher in 1A and the result reflected this. Another teacher in level 2 used textbooks as a key source of reading material and this was also reflected in the result. The low response for comics may also support this result as comics/ graphic novels were not promoted specifically in any of the classes at that time, although they have been shown to be very valuable in other settings. Teacher input during the project seemed to be a factor that encouraged students in their reading habits. Teachers also modelled a reading habit by reading during each 10 minute ER session. Day and Bamford (1998) stated that teachers need to orient students to the goals of ER and be role models "demonstrating what it means to be a reader and the rewards of being a reader" (p. 8).

Interviews: Pre Interviews

Memories of reading in L1

When asked how they enjoyed reading in their first language as younger adults, those who had primary and secondary schooling in their home country were mostly positive about the experience and remembered it as rewarding. They remembered reading a range of material in their L1—stories from their own culture, history, magazines and newspapers. For example, S3 said:

I read all the books from a very famous author. He was a painter and an author and poet is very famous. Rabindra. He wrote dramas and novels and I read all his books. He also wrote songs and I like them very much too.

The one exception was S2 who said, "In my memory I don't think I was very good at reading. We lived in the country and my parents they were not able to read." Students

from refugee populations had varied experiences. They were all in neighbouring countries by the time they were adolescents and did not have the positive experiences of L1 reading that were experienced by many of the other students. Similarly S6 reported:

I was born in Congo but when I was six months my mum took me to Angola. So, I went to kindergarten then right through school and spoke Portuguese, and then when I was 17 and went back to Congo.

First encounter with English

When asked about when they had first learned English, five of the nine had started to learn at junior high school in their home nations, and the other four who had refugee backgrounds started learning on arrival in New Zealand. This information is significant in that those that first encountered English in junior high school or secondary school would already have had literacy in their L1 and a basis of syntax and lexis on which to build, and they could transfer that knowledge and strategies learned while acquiring L1 to the second language. Those with an interrupted schooling and/or who had limited literacy in their L1 had a different experience in that they were often learning to read and write in their second or other language, so they had rather "a mammoth task to face" (Field & Sellars, 2008, p. 137).

Current reading in L2

In the first interview, participants were asked what they read in English. Even in this first interview the students with interrupted education and low literacy spoke warmly about reading with their children. It was noticeable that they were active in seeking out books at the local and institutional library and then reading them with their children.

I get books at the local library every two or three weeks and I have my sister and we read them together at nighttime. (S5)

And I also take my children to the library about two times a week and so we get children's books. (S7)

The students with continuous education were reading more widely and seemed to enjoy reading books about other countries and travelling. One student spoke about romantic novels and romantic histories and another who liked cooking was very interested in recipe books.

Students were asked what they were reading online in English language, particularly through the class Moodle site. All but one student were using digital means both inside and outside the classroom. This process of using digital means seemed very natural and indicated that most of the interviewees were becoming normalised to the use of online media for reading as well as print (Bax, 2003).

Reading enjoyment and challenges

Prior to the intervention, four of the nine interviewees expressed an attitude of disinclination towards reading. These students indicated that although they were trying to read, they did not find it enjoyable, and in each case they linked this to difficulties in understanding vocabulary. By contrast, three of the five interviewees who gave a positive response were parents of young children and specifically mentioned that they liked reading with their children.

I bought the book about beauty at the Warehouse. If I'm interested in something I want to read it. (S6)

Yes it's still a little hard for me, that I asked my sister or use a dictionary it if I need to. (S5)

I find the grammar is very different and hard. Some words I didn't understand but other words have fairly the same meaning. (S6)

I find reading quite difficult, but my classmates and teacher help me a lot (S9).

Post-Interviews

The following themes appeared from the post interviews.

Increased confidence and enjoyment

In the pre-interview, only a very few responded to the question about enjoying reading, as if this was a goal that they had not realised at that stage. In the post interviews, most were still finding reading quite difficult because of the vocabulary load, but were able to see that reading was beneficial to their future studies. Most students answered positively to the query about enjoyment and seemed more engaged with a variety of texts than at the pre-interview.

I really love reading, so I am reading more and I think I will continue to read because reading is food for your brain. (S1)

I really like reading together in a class because you can share ideas with the other person. If you maybe just reading by yourself it's a little bit hard. (S9)

Students seemed to enjoy the autonomy of being able to read by themselves. S3 said "I like reading but I wait until I'm alone so I can read". S2 was discovering the enjoyment of reading with other classmates and her own children, and said she loved reading, mentioning magazines and cartoons. This aligns with previous findings that suggest enjoyment increases motivation to read (Day & Bamford, 1998).

Interesting material increased enjoyment

One factor that helped students to enjoy reading was finding books and material that interested them. Students enjoyed newspapers with current events and magazines with pictures. Others liked reading online in both L1 and L2. The print material read by most of the students were library books or books from school, magazines, real life stories or books about travel or other countries.

There was also evidence of students more specifically identifying genres of particular interest to them. For instance, one participant explained:

I really like to read adventure books so I'll look for those. I really like reading stories about people who have come to New Zealand to get more understanding about that. (S7)

This level of specificity was not present in his pre-intervention interview, and suggests that the additional reading experience enabled him to identify particular areas of reading interest, suggesting an improvement in motivation to read.

As mentioned in the quantitative data, students in Level 2 particularly valued reading with their teacher. This activity was 'liked' more than any other in the checklist. It was noticeable that Level 2 students were more confident than the other students which in turn increased their motivation to read more. They then sought more opportunities to read and select their own texts, with several mentioning going to a library outside of class for the first time. This suggests that they were taking further steps towards becoming independent learners.

The reading process in L2 was becoming somewhat easier

Students were reading similar texts to those they were reading in the pre-interviews and most commented that they found the task of reading a little easier.

I stop at the library and read the easy books. I think if I can read step-by-step I will get better. (S2)

And the environment encourages you. When you look around and see people reading and writing you want to do it to. You want to be able to show that you can do it as well. (S6)

Other students appreciated the intensive reading they were experiencing in class. A number of students commented that the dictionaries on their phones were assisting them to read. It seemed that students at Level 2 were starting to experience some fluency when reading and although they still encountered unknown vocabulary, they felt they had more strategies to manage this cognitive challenge.

Vocabulary. When I have a difficult word, I use my electronic dictionary to find the meaning of the word. Electronics very good because it's so quick. (S1)

If I just want a quick word or I don't know the word I'll have a look on the phone first. Then I will use my paper dictionary if I want to look up the meaning (S3).

Engagement with their children's reading material

One interesting finding that stood out was the intergenerational benefit gained by students reading alongside their own children. Those students who were parents reflected the sense of pleasure and enjoyment they experienced and mentioned that they thought this helped them to become better readers. There are social as well as linguistic gains for all learners in this informal environment.

I am reading a lot of my children's books. They bring them home and I am reading them down about the same level at the same time. I actually bought some books for them, but they bring home reading books and books from school so we can read these together. (S7)

Most students maintained L1 reading.

Students were reading more in both L1 and L2. They read novels and traditional stories in L1 and used L1 for chatting online and keeping in touch with friends. Those who were keen readers in their first language and had developed a habit of reading since they were children seemed to sense that developing a habit of reading in their L2 was beneficial. They were reading in both their first and second languages and choosing which language they would read in different contexts and for different purposes.

Yes and now I am reading more romantic novel stories about people, and also reading cosmic type stuff. (S1)

I like to read our Bengali novels. (S3)

Constraints to reading for enjoyment

Vocabulary load

Every interviewee mentioned vocabulary load and also seemed to link vocabulary load to enjoyment or the lack of it. Most of them used phones and print dictionaries and support from peers to assist them with new vocabulary. They expressed a preference for phone dictionaries as they were faster than using print dictionaries, and so minimised the interruption to their thought processes.

Material that is difficult or not interesting

The last constraint was reading material that was not interesting or was too difficult. Many students mentioned that they stopped reading when they were not interested in the material, or that there were too many difficult vocabulary items. About half the students mentioned that they were selecting easy material which was helping them to read:

For me it's about if I'm interested in the reading. If I'm interested in the topic I will even read it with the hard words. If the book is not interesting, and you asked me about an hour later I won't even remember, so I really just want to read books that are interesting. (S6)

I read books that are about things that are true. When the book is interesting it encourages you to read more. When it's not interesting it becomes like a disappointment. (S7)

These comments and others made in the post-interviews support the claim by Nation (2004) that extensive reading material should contain the vast majority words known to the second language learner. This ensures success and encourages the learner to read more.

Discussion and Conclusion

This study reports on the findings from a teaching intervention in which an Extensive Reading programme was introduced to learners at early stages of language acquisition. Overall, through the intervention, students appeared to be gaining in confidence, were enjoying reading more, and were starting to develop some positive habits for reading outside the classroom. Many remembered the enjoyment they gained from reading in their L1, and this seemed to motivate them to keep reading. Although overall confidence in themselves as readers still seemed to be fairly low, this was perhaps to be expected after just ten weeks. Nevertheless, we may tentatively conclude that the ER intervention played a positive role in promoting reading, demonstrating that extensive reading need not be confined to higher proficiency levels.

The issue of unfamiliar vocabulary further suggests that prior to initiating ER programmes, especially for lower level learners, a working knowledge of high frequency words will be very useful and could be further consolidated with an ER programme. Convenient access to dictionaries, especially those that are accessible by smartphone will support students. Students acknowledged that phone technology greatly assisted them to mitigate the interruptions caused by unfamiliarity they were experiencing when reading.

During the interviews students at Level 2 displayed moves towards independence, through selection of material that interested them. Such learner autonomy is strongly encouraged in New Zealand educational institutions. One of the key principles of adult learning is that "adults are self-directed and are capable of independent learning and

that motivation factors for adult learner are deep seated and internally derived" (Tertiary Education Commission, 2008, p. 10).

The findings suggest some tentative principles for designing an ER intervention. Firstly, it was notable that all of the students who initially reported not enjoying reading cited the difficulties they experienced, particularly in relation to encountering new vocabulary. When students continue to read material that is too challenging they may start to disengage from the process, but when they read material that they are personally interested in and within their vocabulary range, they may start to develop positive attitudes towards reading. This suggests that an important principle for an ER programme would be to ensure that the texts provided in class are based on appropriate vocabulary selections. Although this is an established principle of extensive reading (e.g. Day & Bamford, 1998), it is worth reiterating because it could be that an ER programme could be undermined if the first selections made by a student are too difficult.

A further finding from the pre-intervention interviews was that the students with young children had an overwhelmingly positive attitude to reading. This was attributed to the time spent with their children, helping them to read. It could be that it was the social aspect of such encounters that made reading a more enjoyable experience, with parents and children engaged in a joint, collaborative activity with a positive educational goal. This is, of course, not the typical experience of reading for many people, which is instead associated with solo activity. It may therefore be worth exploring in future studies whether a social element for extension reading could be created in the classroom, particularly for those without (young) children. It might, for example, be worth exploring the possibility of shared reading tasks, where two students both have a copy of the same text, and one reads a page (or sentence) aloud while the other follows, before swapping roles, and with naturally occurring conversation occurring as they discuss or comment on the text. Another possibility that may be worth investigating could be the role of book clubs, where students pre-read a text at home, and then come together to re-read and discuss it.

A further finding that may be helpful when instigating an ER programme was that texts within the student textbook were the most popular source for reading materials. This may seem somewhat surprising as ER seems most often promoted through graded readers. Comments from the students suggested that this was related to the visual appearance of modern textbooks, which are often glossy and visually appealing images and layout design. It may also be that the textbooks appealed due to their familiarity and because they often involved pre-teaching of vocabulary and grammar. One way that teachers could make use of extensive reading from the textbook would be to encourage learners to scan ahead to readings in later chapters that may be of interest, or to go back to texts that had been read in previous chapters.

The study also reinforced the notion that at the early stages of reading for language learners, the teacher has an important role in promoting, modelling and guiding the reading process. It seems to suggest that when there is an affirmative context for reading in the class, and access to further material available outside the class such as institutional and civic libraries, students feel motivated to read both at home and in class.

One of the limitations of this study was that the number of students interviewed was too small to be able to generalise, but it seems that many of the students were starting to enjoy the experience of reading and finding it pleasurable. This study concurs with the findings of Rodrigo et.al. (2014) that beginner language students can start to enjoy reading and respond to a stimulating reading environment when there are quantities of printed material available inside and outside the class that are interesting and age and level appropriate. The digital environment also seems to be fostering students' interest in reading, as they interact with the social and business world through digital means. In short, this study suggests that beginner students can start to manifest a reading habit.

References

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language

problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1–27). London: Longman.

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

Institute for Test Research and Test Development (n.d.). CEFR Reading Assessment Descriptors Retrieved September 29, 2017, from <u>https://www.languagetesting.com/pub/media/wysiwyg/cefr/CEFR-Descriptors-</u> Reading-Assessment.pdf

- Day, R. & Bamford, J. (1998) *Extensive reading in the second language classroom*. Cambridge, England: Cambridge Language Education.
- Field, J. & Sellars, A. (2008). From remembering to reading: The mammoth task of beginning to function in a literate world. CLESOL Proceedings 2008: TESOLANZ.
- Freeman, D. & Freeman, Y. (2003) Teaching English language learners to read: Learning or Acquisition. In Garcia, G. (Eds.), *English learners: Reaching the highest level of English literacy*. (pp. 34-54). Newark, NJ: International Reading Association.
- Krashen, S. (2004, April). Free voluntary reading: New research, applications, and controversies. Paper presented at the RELC conference, Singapore. Retrieved September 29, 2017, from

http://www.sdkrashen.com/content/articles/singapore.pdf

Maley, A. (2009, December). *Extensive reading: Why it is good for our students...and for us*. Retrieved September 29, 2017, from

https://www.teachingenglish.org.uk/article/extensive-reading-why-it-good-ourstudents%E2%80%A6-us

- Nation, P. (2004) Vocabulary learning and intensive reading. *EA Journal, 21*(2), 20-29.
- Nation, P. (2015) Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136-145.
- Rodrigo, V., Greenberg, D., & Segal, D. (2014). Changes in reading habits by low literate adults through extensive reading. *Reading in a Foreign Language*, *26*(1), 73-91.
- Scales, A.M., & Rhee, O. (2001) Reading habits and patterns. *Reading psychology*, 22(3), 175-203.
- Tertiary Education Commission (2008). *Starting Points. Supporting the learning progressions for Adult Literacy.* Wellington: Tertiary Education Commission
- The Extensive Reading Foundation (2011). *The Extensive Reading Foundation's guide to extensive reading*. Retrieved November 15, 2017, from <u>http://erfoundation.org/ERF_Guide.pdf</u>
- Witten, I., Wu, S., & Yu, X. (2011). *The book of FLAX: A new approach to computer assisted language learning*. Retrieved November 15, 2017, from <u>http://flax.nzdl.org/greenstone3/flax</u>

Appendix 1:

The A1 descriptor for Reading is as follows:

Readers at the A1 level can understand very sort and simple texts a single phrase at a time picking up familiar names, words and basic phrases often with a great amount of rereading. They can understand short, simple greetings and messages, for example in emails and text messages or post cards. They can find specific, predictable information in very simple everyday material such as advertisements, webpages, timetables and catalogues. They can understand information about people and very short and simple descriptions especially if there is visual support. They can follow short simple instructions and directions relate to the most common everyday situations although frequent misunderstandings may occur.

The A2 descriptor for Reading is as follows:

Readers at the level A2 can understand short, simple texts containing frequently use words and phrases, names, cognates, and shared international vocabulary. They can understand short simple emails and letters from friends and colleagues. They can find specific, predictable information in simple everyday material such as advertisements, brochures, web pages, timetables and catalogues. They can understand the main points in short, simple descriptions of persons, places and things with which they are familiar. They can understand simple instructions and directions.