Surveying and Observing Students to Enhance a Prototype Smartphone and Tablet Software Application for Spoken English Language Learning

Maxim Woollerton
Gary V. Ireland
Joshua Rappeneker

Summary
The subject of this paper is a survey of students at several Japanese universities, designed to elicit feedback on the strengths and weaknesses of a prototype English language learning software programme running on smartphones and tablets. The paper explains the reasons for conducting the survey, the kind of information sought, the survey design, the method of implementation, the results obtained from the survey, and the observations and conclusions drawn by the authors, as part of an effort to make changes that improve the overall quality of the software programme.

Keywords
Spoken English, survey, smartphone, tablet, software

1. Introduction
During the past three years, the authors have been designing and creating
an English language learning application for use with Apple Inc.’s iOS devices: iPhone, iPad and iPod Touch. The core material for this application is being taken from a series of English language learning textbooks, *The English Course*, originally written by the authors between 2004 and 2010. The series is being dismantled and recreated in a modular format for use in this and future applications. The original materials for *The English Course* series comprised a textbook, a set of DVDs containing video and audio material, and an online student self-study web site. The concept of the series was to blend the use of the communicative approach to language learning in the classroom, with the best techniques and materials available from computer-assisted language learning (CALL) and thereby create a richer learning product.

Changes in the technological world, in the form of the large scale adoption of hand-held mobile computing devices, is propelling change in both the educational and the commercial world. Books, disks and desktop computers are being replaced by laptops computers, smartphones and tablet computers, as the devices that many, including students, desire to use for educational as well as recreational and employment purposes. These new devices require new forms of media to run on them, incorporating not just text, video and audio material, but which also allow for interactive use and conform to the new dimensions and interfaces that handheld devices impose on the user. Additionally, it is likely that learning itself is in a process of transformation to a more individualistic and less classroom and fixed schedule type of activity. These trends are the principal reason why the authors decided to transfer the educational materials within *The English Course* to these new technological and usage formats.

— 324 —
This project has proved to be a huge challenge and has presented an enormous learning curve for the authors. There are several reasons for this. Firstly, creating stand-alone modules suitable for use with smartphone and tablet devices from the core materials required a great deal of discussion and redesigning. Secondly, taking all of the ideas for use of materials and making them come to life as interactive activities has required enormous experimentation and engineering time. Many of the activities in the prototype module are unique and cannot be seen in other applications. Thirdly, due to the continuing development of new operating systems with better capabilities, and with continued research and understanding of the technology and the learning environment, the project has necessarily been through numerous changes and updates to date.

Having progressed with a prototype module to an advanced stage, as described in Ireland and Rappeneker (2013) and in Ireland, Woollerton, Rappeneker (2014), it was decided that for development to go further it was time to put the module into the hands of and elicit feedback from prospective users. To do this, an online survey was created and conducted during the period of December 2013 to January 2014.

2. Reasons for conducting the survey

Computer software for educational purposes can generally be categorised as being either tutorial in nature or as a tool, a distinction first suggested by Taylor (1980) and expanded upon by Levy (1997), who suggested that it be used as the basis of a framework by which software can be evaluated pedagogically. The prototype application largely but not exclusively fits the description of software that is in the tutor role. As such, the software
includes a virtual teacher within the machine, guiding and controlling the student through the use of the software (Levy: 182). The tutorial role also features evaluation of the performance of the student by the software.

The virtual teacher manifests itself in the form and language of instructions and guidance before an activity within the application, and in the evaluation and feedback following an activity. The teacher is also present within a dictionary or other form of reference used within the software, and the teacher can also be seen within the content of each activity, determining the language used or not used, setting the purpose of the task and the design and method used within the software to achieve that purpose. The teacher is also present within the overall design of the application in aspects such as the navigation system. As Levy as pointed out (ibid: 197), software operating in a tutor role ‘requires methodology governing presentation and interaction’.

In Ireland, Woollerton, Rappeneker (2014) the authors described a clear set of interaction design principles formulated by Kristof and Satran (1995) that deal with user orientation to software and usability issues. Users must be able to quickly get a feel for the software, to understand how it is organised and what it is for. In terms of usability, Kristof and Satran (ibid: 50–51) presented six ideas for maximising the usability of software. These are:

1. Remove obstacles [to users interacting with the software]
2. Minimise effort [for example, to navigation]
3. Give feedback [to let users know that something is happening]
4. Be explicit [so that users know what can be done with different objects on the screen]
5. Be flexible [to let users access or exit different features quickly]
6. Be forgiving [of user actions and do not limit the user to getting something right before moving on]

In Ireland, Woollerton, Rappeneker (2014) the authors examined each section of the prototype application, evaluating the issues of user orientation to the software and its usability. The authors felt that it was essential to get the views of potential users of the software, in terms of the application’s navigation system, how easily the purpose of activities could be understood, the quality of the instructions and the feedback, and whether this software was judged to be a useful for learning English.

There were also commercial and marketing considerations that the authors wanted to research. The points here were factors such as the extent of smartphone or tablet ownership, the operating system of the devices that potential users owned, attitudes towards purchasing software for language learning on smartphones and tablet devices, and what potential users felt the software was worth in monetary terms.

3. Survey design

The survey was divided into three sections. Part one of the survey was intended to find out basic information about each individual being surveyed. This included questions on each student’s age, gender, native language, English reading ability and whether he/she suffered from any condition that would be relevant as a challenge to using the application. The second part of the survey focused on the student’s relationship with technology used for educational purposes. This included questions about the student’s use of technology, his/her purpose for studying English, what kind of devices the student owned and which operating system that equipment was running.
There were also questions about how many language learning applications a student had used, the educational purpose for using those applications up until now and for what purpose he/she wished to use such applications in the future. The third section of the survey focussed on the student’s opinions about using the prototype application. The questions in this section asked for student opinions about the navigation of the application, the quality of the instructions, which sections of the application the student most enjoyed and least enjoyed, the student’s understanding of the purpose and use of each section of the application, the quality of the feedback provided after completing each section and the student’s impression of the overall usefulness of the application in learning English. Additional questions in this section asked the student for his/her opinions on buying applications for language learning and what he/she thought the application was worth in monetary terms.

The survey was designed to provide sufficiently detailed information in a format and length that would avoid fatigue or boredom when answering the questions. In all, there were 26 questions. The survey was written in English and then translated into Japanese. The English language version of the survey can be found in Appendix 1. It was decided to use closed rather than open questions throughout the survey. Although a preference for open questions is often seen as providing more useful information (Nunan: 143), the advantage of using closed questions was not only that the answers would be easier to collate and analyse (ibid) but also that if students had to respond in English, this would be more difficult and time-consuming for them, while if students responded in Japanese, those answers would have to be translated into English. Given the number of
questions that were to be asked and the size of the sample, closed questions were clearly more practical. Most of these were presented in a single answer, multiple-choice format, using Likert scale answer options. An effort was made to design questions that avoided the use of a median or non-decisive answer options, in order to prevent students from choosing the middle, non-committal or ‘easy’ answers. The Likert scale questions mostly have four possible answers. There has been some debate over whether such questions should have an odd or even number of possible responses with some researchers preferring to avoid an odd number of answers for fear that some responding may choose the central or ‘neutral’ answer to avoid giving an opinion. Dörnyei (2001: 207) dismisses this as relatively unimportant, but in this case it was felt that, in a consensus-based society such as Japan’s, this might be an issue and so an effort was made to avoid neutrally worded answers as the middle choice. All of the question types fall into the categories described by Youngman and presented in Nunan (144). Most questions are either scale questions used for eliciting student opinions with answer options such as as ‘very good’, ‘good’, ‘bad’ and ‘very bad’) or quantity/frequency questions used for determining student behaviour. In addition to these two types, ranking questions were also used. Further to this, where appropriate and possible, an effort was made to follow Dörnyei’s (2001: 203–204) guidelines for writing survey questions, which are based on the work of Anderson (1985) Oppenheim (1992). These can be summarised as follows:

1. Devise ‘characteristic’ statements; avoid neutral or extreme formulations;
2. Avoid statements that may be interpreted in more than one way;
3. Avoid statements that are likely to be endorsed by almost everyone or almost no one;
4. Select items that are believed to cover the whole range of the affective scale of interest in a balanced manner;
5. Include both positively and negatively worded items;
6. Keep the language of items simple, clear, short and direct;
7. Use simple words and avoid acronyms, abbreviations, colloquialisms, proverbs, jargon and technical terms;
8. Beware of loaded words such as ‘democratic’, ‘modern’, ‘natural’, or ‘free’;
9. Avoid items containing universals which may introduce ambiguity such as ‘all’, ‘always’, ‘none’ and ‘never’;
10. Avoid the use of double negatives.

4. How the survey was conducted

The survey was conducted during the period from the beginning of December 2013 to the end of January 2014 by giving iPhones to either individuals or small groups of up to four students studying English at university level in Japan. The students were initially asked to examine the prototype application by themselves without any explanation as to what the application was for or how it was to be used. Students were asked to spend approximately 20–30 minutes doing this. When the students felt that they had had enough time to examine the application they were then directed to answer the survey questions (in Japanese), which were located on a web site and which consisted of three pages of questions. In total, 167 students were surveyed at six different institutions. The overwhelming majority of the
students were Japanese native speakers (161 students, or 96 per cent), while there were four Chinese native speakers, one Korean native speaker and one English native speaker. Of the 167 students surveyed, 110 (66 per cent) were female and 54 (32 per cent) were male. Almost all of the students who gave their age (153 out of 154) were between 18 and 22 years of age. This equates to the survey being conducted mostly among first and second year Japanese university students.

Whilst it is easy to be clear about the sample’s size and other features. It is more difficult to identify the exact nature of the population that the sample group should represent. In the past, much of the content material within the software would have been presented in textbooks, on DVDs or CDs and on web sites intended for students of English in Japanese universities. With the prototype application, the intended users are from a much broader range of people. Potentially, the software users could be of any age from early teen-age years and upwards, and from almost any developed nation and many developing ones. As smartphone or tablet owners, it is to be expected that the potential users would be relatively wealthy. Arguably, the potential users are more likely to be learning individually, to be self-motivated and usually able to determine independently whether or not to purchase and use the software. A larger sample size of around 380 students would have been preferable to achieve a confidence level of 95 per cent, but this was not possible given the time and resources available to the authors. It is probably true to say that a confidence level of 80 per cent can be achieved with the sample size (167 students) that was used for a margin of error of five per cent. With the smaller sample size (167 students), with a confidence level of 95 per cent, and a total population in excess of 20,000, the margin of error...
can be calculated to be 7.55 per cent.

5. Survey results

Within section one of the survey, there were two questions that were intended to differentiate between students based on their English level or other challenges that might have an effect on the students’ use of both the software and hardware. The first question asked students to rate their own English reading level. This might be particularly relevant to student responses to the questions about the software’s instructions and feedback, which at this stage, is exclusively presented in English. A majority (109 students or 65 per cent) categorise themselves as basic level non-native readers of English, while 52 individuals (31 per cent) categorise themselves as at an intermediate level. Only two students identify themselves as being at an advanced level. Four individuals claimed to be native English readers, which was somewhat surprising since only one student had responded that his/her native language was English. It might be expected that this discrepancy might be from the non-Japanese students reading the questions in Japanese, but in fact, the three native English readers had all previously identified themselves as Japanese native speakers. This indicates that there are clearly errors within the results of the survey. The survey also asked students to indicate if they suffered from any physical or mental challenges, as this could also be of relevance to ease or difficulty in using the software and hardware. Almost all students (166 out of 167 individuals) responded that they had no known challenges. Only one student indicated that he/she was challenged by dyslexia.

The second section of the survey focussed on student ownership of
Surveying and Observing Students to Enhance a Prototype Smartphone …

technology and if, how, why and how much it is used by students for educational purposes. The survey found that among these students, smartphones and tablets are currently the most popular form of technology used for educational purposes, with 34 per cent of the students indicating use of these items compared with 31 per cent who indicated the use of personal computers. In a more distant third place with 16 per cent came television and digital versatile disk (DVD) users. Fourteen per cent made no response to this question, which exceeded the ten per cent who use radio or compact disc (CD).

It is also interesting to note that 96 per cent of the students indicated that they owned a smartphone. When comparing this figure with the 34 per cent of students who choose to use a smartphone for educational purposes, it would seem that while the smartphone is a popular device to own and it is the first choice as the kind of technology to use for learning, there is a large number of smartphone owners who choose not to use it for learning. That said, other ubiquitous forms of technology (such as television and radio), are even less popular. It is possible to speculate as to why this might be, but unfortunately there is no clear reason within the data provided by this survey. It may be safe to say that the present time is a period of change for technology used for learning and that older forms of technology are being replaced by newer forms, but that the period of change is by no means over yet.

In answer to the question about the students’ current main focus when using technology to study English, the most popular reason given was for listening practice (40 per cent). Next was vocabulary practice (24 per cent). In third place was speaking practice (12 per cent). Grammar practice was the
main focus of only eight per cent of students, while reading practice and writing practice both received five per cent each. These clear priorities of the students for listening practice, vocabulary practice and speaking practice have provided a reassurance that the software’s focus on these same areas will match students’ needs.

Apple devices have a clear lead with the students surveyed. Fifty-nine per cent of students are using an Apple smartphone (iPhone) and almost four out of five of those are using the latest major version of the operating system (iOS7). Thirty-two per cent of students are using a smartphone running a version of Google’s Android operating system. Students with smartphones running a version of Microsoft’s Windows Phone operating system amounted to just two per cent of the total of 167 total students. Four students did not specify an operating system for the smartphone they claimed to own. Ownership of tablet devices was far lower than smartphone ownership. A large majority of the students (79 per cent) do not currently own a tablet device. Of those students who do own a tablet, Apple’s tablet (iPad) is the most popular, with 21 out of 35 tablet–owning students claiming ownership of an iPad. Use of the most recent operating system on these iPads exceeds use of the preceding operating system by three to one. These findings vindicate the choice to develop for Apple’s devices first. The reasons for doing so were not just to reach the greatest number of users (because ownership statistics especially in developing nations are not be the same as in countries like Japan), but also to develop for a platform that was the strongest in terms of commercial development, the least fragmented technologically and the least susceptible to viruses and other malware.

The next questions in the survey asked students about their use of
software for language learning on smartphones or tablets in terms of the number of applications they had used, the frequency and educational focus of that use. The number of applications used was low. Twenty-six per cent of students had only used one application, while 44 per cent had used between two and four applications. Use of five or more applications was limited to 13 per cent of the students. Coupled with the low number of applications employed by the students was their infrequent use. The most common response (38 per cent) was that applications were only used occasionally and 27 per cent of students responded that they never used applications on a smartphone or tablets. Frequent use of applications was limited to just 11 per cent of the students. The main reasons students gave for their current use of applications on these kind of devices were as follows: vocabulary practice (41 per cent); listening practice (15 per cent); grammar practice (14 per cent); speaking practice and writing practice (four per cent each); and other (one per cent). Seventeen per cent of students gave no response. In terms of the future use of applications, the main reasons were somewhat different, with substantially more interest in using the software to practise speaking (26 per cent) and to practise listening (23 per cent). This may not just reflect students developing a different set of priorities in the future, a lack of speaking and/or listening–focussed applications at present and/or also a hope that there will be more or better applications with that kind of focus in the future.

The third section of the survey focussed on the prototype application. The most straightforward information to examine relates to the questions about which sections of the prototype application the students enjoyed most and least. The sections that were the most popular were Audio game (40 per

— 335 —
cent), *Pronunciation practice* (27 per cent), *Video* (24 per cent) and *Video task* (17 per cent). All of the other sections achieved a score of 11 per cent or less. The sections that were chosen as their least popular by the students were *Test* (26 per cent), *Audio game* (19 per cent), *Language Focus I & II* (16 per cent). It is interesting to see that *Audio game* was the most popular section, but also one of the least popular. The reasons for why this might be will be discussed in the next section of this paper. In general terms it is fair to say that the parts of the module that were the most popular were the ones that contained the most interactive content, while many of the less popular sections contained fairly static content.

Students were clearly positive about the navigation system within the application. Ninety-one per cent expressed a favourable impression of the navigation overall, with 74 per cent expressing the opinion that the software was easy to navigate, meaning that most students found it easy to move through the software from start to end and backwards and forwards between different sections of the application.

With the questions that asked the students for their opinions on the instructions, feedback, purpose and how to use the software, the results were again largely positive. Given that instructions and feedback in the prototype application are presented exclusively in English, the English reading level should be considered as a possible influencing factor on the students’ evaluation of these aspects of the software. Table 1 (below) shows the results broken down by each student’s claimed reading level.

There was no meaningful difference between the approval rates for the quality of the instructions contained within the software for students of basic and intermediate English reading levels. This was not the case for approval
Surveying and Observing Students to Enhance a Prototype Smartphone …

Table 1: Student Responses to Questions on Instructions, Feedback, Purpose and Use

<table>
<thead>
<tr>
<th>Area</th>
<th>Reading level</th>
<th>Very positive</th>
<th>Positive</th>
<th>Negative</th>
<th>Very negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructions</td>
<td>Native</td>
<td>2 / 50%</td>
<td>1 / 25%</td>
<td>0 / 0%</td>
<td>1 / 25%</td>
</tr>
<tr>
<td></td>
<td>Advanced</td>
<td>1 / 50%</td>
<td>1 / 50%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>40 / 77%</td>
<td>10 / 19%</td>
<td>0 / 0%</td>
<td>2 / 4%</td>
</tr>
<tr>
<td></td>
<td>Basic</td>
<td>80 / 73%</td>
<td>17 / 16%</td>
<td>5 / 5%</td>
<td>7 / 6%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>123 / 74%</td>
<td>29 / 17%</td>
<td>5 / 3%</td>
<td>10 / 6%</td>
</tr>
<tr>
<td>Feedback</td>
<td>Native</td>
<td>2 / 50%</td>
<td>2 / 50%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Advanced</td>
<td>1 / 50%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
<td>1 / 50%</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>37 / 71%</td>
<td>13 / 25%</td>
<td>1 / 2%</td>
<td>1 / 2%</td>
</tr>
<tr>
<td></td>
<td>Basic</td>
<td>54 / 50%</td>
<td>48 / 44%</td>
<td>6 / 6%</td>
<td>1 / 1%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>94 / 56%</td>
<td>63 / 38%</td>
<td>7 / 4%</td>
<td>3 / 2%</td>
</tr>
<tr>
<td>Purpose</td>
<td>Native</td>
<td>1 / 25%</td>
<td>0 / 0%</td>
<td>2 / 50%</td>
<td>1 / 25%</td>
</tr>
<tr>
<td></td>
<td>Advanced</td>
<td>2 / 100%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>31 / 60%</td>
<td>9 / 17%</td>
<td>12 / 23%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Basic</td>
<td>49 / 45%</td>
<td>22 / 20%</td>
<td>36 / 33%</td>
<td>2 / 2%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>83 / 50%</td>
<td>31 / 19%</td>
<td>50 / 30%</td>
<td>3 / 2%</td>
</tr>
<tr>
<td>How to Use</td>
<td>Native</td>
<td>1 / 25%</td>
<td>0 / 0%</td>
<td>2 / 50%</td>
<td>1 / 25%</td>
</tr>
<tr>
<td></td>
<td>Advanced</td>
<td>0 / 0%</td>
<td>1 / 50%</td>
<td>1 / 50%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>34 / 65%</td>
<td>7 / 14%</td>
<td>11 / 21%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td></td>
<td>Basic</td>
<td>45 / 41%</td>
<td>20 / 18%</td>
<td>42 / 39%</td>
<td>2 / 2%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>80 / 48%</td>
<td>28 / 17%</td>
<td>56 / 34%</td>
<td>3 / 2%</td>
</tr>
</tbody>
</table>

rates for the quality of the feedback, the purpose of the activities or how easy it was to use the software. With these aspects of the software, a clear distinction can be seen between students of basic and intermediate English reading levels.

On the question of the quality of the feedback (meaning the information that is displayed to the student after he/she has completed an activity, such as a score, comments or suggestions for actions to take to improve his/her
future performance), intermediate level students were more likely to rate this 'very good' than basic level students (71 per cent compared with 50 per cent), while basic level students were more likely than intermediate level students to rate the feedback as 'quite good' (44 per cent compared with 25 per cent). There was no meaningful difference between the rates of disapproval for the quality of the feedback between basic and intermediate level students.

On the question of how quickly students were able to understand the purpose of different sections within the application, a similar pattern in the data was observed to the data for feedback, but unlike the case of feedback, there was also a detectable difference between the disapproval rates for the purpose of the sections in the software. Intermediate level students were more inclined to choose 'very quickly' than basic level students (60 per cent compared with 45 per cent) and basic level students were slightly more likely to choose 'quickly' than intermediate level students (but at three per cent, this difference was not statistically meaningful). More significantly, basic level students were more inclined to choose 'quite slowly' than intermediate level students (33 per cent compared with 23 per cent).

With regards the question of how quickly students were able to understand how to use the application, a very similar pattern was observed to that with the question about how quickly it was possible to understand the purpose of the different sections within the application. Intermediate level students were more likely to use 'very quickly' than basic level students (65 per cent compared to 41 per cent) and again basic level students were slightly more likely to choose 'quickly' than intermediate level students (but at four per cent, this difference was not statistically meaningful). Basic level
students were more inclined to choose 'quite slowly' than intermediate level students (39 per cent compared with 21 per cent).

Clearly there is a link between the students’ self-assessed English reading level and their feelings about the quality or effectiveness of the instructions, feedback, purpose and ease of use of the software. This strongly indicates that, since these aspects are arguably more important for students with a lower level of English, it is important to make their comprehension easier. The obvious way to do this is to localise the software. That is to put the instructions and feedback in the native language of the students to the extent that this is practical.

Overall, the response to the effectiveness of the software as a language-learning tool was overwhelmingly positive. The table in Table 2 (below) shows that, taken as a whole, 96 per cent of students had a positive or very positive attitude towards the software. There would appear to be a slightly higher tendency for intermediate level students to be very positive about the software compared to basic level students (50 per cent compared with 38 per cent), while basic level students were slightly more likely to be positive rather than very positive about the software compared with intermediate

<table>
<thead>
<tr>
<th>English reading level</th>
<th>Very positive</th>
<th>Positive</th>
<th>Negative</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>1 / 25%</td>
<td>3 / 75%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td>Advanced</td>
<td>1 / 50%</td>
<td>1 / 50%</td>
<td>0 / 0%</td>
<td>0 / 0%</td>
</tr>
<tr>
<td>Intermediate</td>
<td>26 / 50%</td>
<td>24 / 46%</td>
<td>1 / 2%</td>
<td>1 / 2%</td>
</tr>
<tr>
<td>Basic</td>
<td>41 / 38%</td>
<td>63 / 58%</td>
<td>3 / 3%</td>
<td>2 / 2%</td>
</tr>
<tr>
<td>All</td>
<td>69 / 41%</td>
<td>91 / 55%</td>
<td>4 / 2%</td>
<td>3 / 2%</td>
</tr>
</tbody>
</table>
level students (58 per cent compared with 46 per cent).

The final two questions in the survey asked students under what circumstances they were willing to purchase software for learning English on a smartphone or tablet and what they thought the value of the prototype application was worth in monetary terms. The answers here were interesting, although disappointing to anyone hoping to produce software commercially. Only 20 per cent of students indicated a willingness to pay, despite so many students assessing the prototype application useful for learning. Few students appear willing to pay for software to learn English on a smartphone or tablet device. This unwillingness to pay may be equally strong with educational material in older forms (such as books). Moreover, there was no meaningful difference between students according to their English reading level, suggesting that there is no link between paying for software previously and having achieved a certain level of educational ability. It would also appear that purchasing products with an educational value or education itself is either of low priority to these students or that students believe that they can learn effectively without paying or that some other factors have a more controlling effect on discretionary spending by these students.

When asked to evaluate what they believed the prototype application was worth in monetary terms, almost a quarter replied that the product’s value was zero. Thirty-seven per cent of students felt that the value of three modules of the product was worth between one and 249 Japanese Yen. Twenty-two per cent felt that three modules of the software were worth between 250 and 499 Yen. On the one hand, such results may seem disappointing to those involved with commercial educational software
development. The process of creating the software and its audio-visual content is extremely complex and laborious. The cost is not only in time and expertise; there is also a substantial financial cost. There would be little incentive in making such an investment of time and money when it would seem that there is such a limited scope for reward. On the other hand, it has to be remembered that, as pointed out earlier, the Japanese university students in the sample group are just one sub-section of the population that the software is targeted at. University students are likely to have different economic circumstances and spending priorities from members of other sub-sections in the target population, such as those in regular paid employment.

6. Analysis and comments based on results and observations of the survey

This section will detail some issues the users experienced when attempting to use the application, and when completing the survey. These challenges and difficulties were observed and recorded by the authors whilst monitoring the students.
The only practical way to allow participants to access the application was to give smartphones (Apple iPhones) on which the prototype was installed to each participant. This meant that each of the authors had to lend his own personal iPhone to participants to allow them to investigate the module. This limited the number and kind of participants to those that the authors came into personal contact with. As the authors are all university lecturers in Japan, this meant that almost all of the participants were university students. Unfortunately, this system of surveying limited the amount of time that students were able to experiment with the module.

One of the biggest difficulties encountered in the process was simply the amount of time needed for students to effectively use and evaluate the module. Students were put into small groups and told to look at and try each section (activity) of the module, but not to try to complete the sections or the whole module. Generally speaking, students were given 20–30 minutes to experiment. This is certainly not long enough to complete all of the question/answer sections or the final test. To actually complete the whole module takes approximately an hour for a basic level English user. As a result of this time constraint, it was only possible to survey up to ten students per hour.

Because time was at a premium in each case, it was difficult for students to find a balance between using each section of the module for an appropriate amount of time. Some students spent a lot of time on a few of the activities and neglected or skipped others completely. This appeared to sometimes have been because they did not fully understand some of the instructions. In other cases, it seemed to have been because students were enjoying an activity and wanted to finish it to see their feedback. In many
cases, students tried to complete the whole end of module test.

Another problem was that in some cases, students did not make enough effort to properly evaluate the module. The survey, and indeed proper use of the module, relies upon taking time to read and fully understand the instructions for each section of the module. It was noticed almost immediately upon commencing the survey process that some students were not spending enough time to understand the instructions. These students then often skipped through sections or simply did not really understand what they were doing. To remedy this problem, it was decided to give brief explanations to each group of students before they were given the iPhones. Both to save time and to ensure proper evaluation, students were briefly shown each section of the module and told what the aim of the section was. This improved the quality of the user experience for those surveyed and also improved the quality of the survey feedback data.

Attention in this paper will now focus on specific aspects of the software and the survey to narrow down areas for consideration and improvement in future development of the software and any further testing.

6.1 Module Selection

The initial screen with which the user is presented upon opening the application instructs the user to ‘Choose a module’ from the list. Beneath this instruction, there is currently a single option, ‘Module One: A job interview’. Tapping this option takes the user to the activity list of the first module. Whilst this user interface flow seemed simple to the authors, it turns out that the concept of a ‘module’ is not clear without external explanation – a person opening the app for the first time will not necessarily
understand what they are learning beyond the basic title ‘a job interview’, or why. Some changes will need to be made, either to the interface by adding extra descriptive text and explanations, or to the name ‘module’.

6.2 Activity List

The module screen presents a list of the activities the user can perform. The list has a single button on each line, centred horizontally in the screen. It was the programme author’s intent that the user start with the topmost activity, ‘Video Task’, and work down the screen through each activity sequentially. However, it soon became clear that the users either were unaware of our design or unwilling to follow such a linear path through the programme. Users during the testing phase would attempt activities such as the ‘Video Task’ that requires information provided in other, earlier tasks. Some students appeared frustrated when they were unable to answer questions without realising that they had not yet been exposed to the answers. The simplest remedy for this issue is to ‘lock’ each task until the prerequisite tasks are completed. This has the added advantage of possibly providing the users with a greater sense of achievement and direction as they work their way through the application, unlocking each task as they proceed.

6.3 Instructions

All of the more complicated activities in the application have instruction screens that appear before a user first attempts a task. The instructions have been written in as simple English as possible, however it was observed that the users almost universally dismissed the instruction pages without reading
them. When asked why they did not read the guidance, several students replied that it was ‘bothersome’ or ‘too difficult’. Whilst a more motivated user might spend more time attempting to understand the task and instructions, it is clear that the instructions should be made more accessible to users of all levels.

6.4 The Audio Game

The Audio Game activity was unusual in that according to the survey it was the most enjoyable task, but also the second least enjoyable task – it certainly polarised the users. The game screen presents the users with a randomised pile of diamonds laying on the ‘ground’ and a grid of possible positions into which the diamonds can be placed. (See the screen image in Appendix 2)

Tapping on a diamond will play a line from a conversation. It is the user’s task to drag each diamond into the position representing the correct part of the conversation. The conversation runs from the top–left position and continues left–to–right down the screen. As a diamond is dragged over a possible position, that position darkens. If the user releases the diamond in the correct position, a ‘correct’ sound plays and the diamond locks into place. Otherwise, a sound effect representing ‘incorrect’ plays and the diamond drops back into the pile.

In practice, there were several problems with the audio game. Firstly, without reading the instructions it is a fairly difficult activity to grasp initially. Multiple users were observed randomly dragging and dropping the diamonds without realising they were trying to reconstruct a conversation. Part of the problem here is that there is no obvious order into which the
diamonds should be placed. Modifying the position images, or adding number labels may make this clearer.

Another problem is that the user is not given clear enough feedback when a diamond has been dropped on an incorrect position. Although an ‘incorrect’ sound is played, there are no visual cues that the user has made a mistake. This lack of negative feedback would seem to be a source of frustration for some users. Adding a visual effect, such as shaking the screen, would give the user a more immediate and clear indication of a mistake. In the same vein, the lack of any visual feedback at the successful completion of the task seemed to leave some users feeling somewhat disappointed. Overall, the activity would benefit from clearer visual feedback.

Finally, the Audio Game suffered from a poor design decision that made it sometimes quite difficult to use. As explained above, the position images darken when a diamond is dragged over them. The user then has a visual indication of which position cell will be selected if he or she releases the diamond. Whether or not to darken a position cell is calculated in the application by testing if the polygons that represent the diamond and the position cell intersect. This is the most immediately obvious implementation, however it poses several problems. The first, and most serious problem, is that it is possible for the diamond to intersect with two position cells – this means that both position cells are darkened, and the user is now unable to determine which position the diamond will drop into. The second problem is that a diamond spinning with the momentum of the user’s touch may move in and out of an intersecting state – again making it difficult for the user to determine if the diamond is correctly placed. The final problem is that determining whether two or more polygons intersect whilst being animated
and interacted with has a non-trivial CPU cost. Put simply, it uses a lot of computational power, and could slow down the device and reduce the user's enjoyment as well as use more battery life. These problems are all fixed by a small and simple change to the program implementation. Instead of detecting whether the diamond and position polygons intersect, we create a circle that circumscribes the position cell and then simply detect whether the user's finger sits inside the circle (a test which is much less computationally expensive). This solves all the problems outlined above and should make for a more intuitive user experience.

6.5 Survey Translation

As the majority of students taking the survey were Japanese native speakers, the authors had the survey translated and gave each user the option of taking the survey in Japanese or English. It is believed that, all students elected to take the Japanese survey. There were, however, a few problems with the translation. Firstly, there were a few direct transliterations, for example モジュール (or mojuuru), for 'module' which many students did not understand. Secondly, and more problematic was the translation of the terms 'quite' and 'very'. The English survey listed options in the following order: very good, quite good, quite bad, very bad. This was translated as: とても良い (totemo yoii), かなり良い (kanari yoii), かなり悪い (kanari warui), とても悪い (totemo warui), which seems like a reasonable translation. The problem is, however, that while とても (totemo) is often translated as 'very' and かなり (kanari) as 'quite', they don't vary in degree nearly so much in Japanese. It was necessary to explain in each survey session that the first option was the most positive and the second less

— 347 —
positive, and so on. The same problem existed in other parts of the survey and one author was asked several times whether かなり分かりやすい (kanari wakariyasui) was better than 分かりやすい (wakariyasui). After explaining which option was meant to be more emphatic, the students seemed able to complete the survey without difficulty, however it is possible that some responses were affected by these issues.

7. Conclusion

After three years in development, the prototype application is nearing completion, but it has not yet reached an acceptable final form. The software is largely tutorial in nature and as Levy stated ‘requires methodology governing presentation and interaction’ (Levy: 197). Attempting to perfect this methodology is a large part of the reason why the development of the application has taken so long. There is only so much that software authors can do in the absence of users, however. By the end of 2013, the development of the application had reached the stage where it had become essential to elicit the opinions of potential users and to observe novice users trying to navigate and cope with the tasks contained within the software.

The survey instrument did not attempt to extract detailed responses about the prototype application and also covered some issues that were not necessarily pedagogic in nature. In addition, it is recognized that the survey sample was smaller than it should have been and the short time allowed for examining the software may have been problematic for those surveyed. There were also problems with the translation of survey items and options from English into Japanese. Nevertheless, the data obtained can still be seen
as valuable in terms of giving broad feedback in many general areas that were of concern to the authors, such as navigation, instructions, feedback, purpose, value and ease of use of the software.

In addition to the survey data obtained, of equal or greater value was the opportunity to observe people other than the software developers attempting to use the prototype application for the first time. No author of software can definitively pre-judge how users will cope with the design or content of the material that he or she produces. In many cases, what seems obvious and sensible to an author can be foreign and incomprehensible to a novice user of the software. That is what was observed during the survey process for this application.

As a result of this survey process, the authors now know that it in order to improve upon the design of the software, some pre-existing ideas were confirmed and more work needs to be done in several key areas:

1. Instructions do need to be presented in the native language of the user where possible;
2. Users need to be guided in terms of the direction and sequence they follow when progressing through the application;
3. The visual representation used in some activities (for example and in particular, Audio game), need to be redesigned to make it easier for users to conceptualize tasks and the method for completing those tasks.

In addition to the three points above, the authors have identified additional areas, such as feedback, where attention needs to be paid to improving the design of the software.
The final question is, after three years work in development: Will anyone buy this ‘app’, even if it has been perfected? Roughly seventy-five per cent of the students in Japanese universities who took part in this survey indicated that they did not want to pay for English language learning software. Who is willing, able and competent to meet their requirement for a free solution to their educational needs? The answer to this question remains unknown. On the other hand, to the relief of the authors, the market for smartphone applications is automatically and virtually worldwide, and the interest in paid ‘apps’ in other territories may be substantially different from that of the students in universities in Japan.

References
Appendix 1

Survey for The English Course iApp 2013

All information that you provide for this survey will remain strictly anonymous. All answers received for this survey will be compiled and analysed statistically.

Please choose the letter of any answer that is correct for each question and/or write/type in the blank space.

Section 1 – About yourself

1. Are you a student, teacher or neither?
A) Student  B) Teacher  C) Neither a student or teacher

2. What is your age?
A) ____ years of age  B) Prefer not to say

3. What is your gender?
A) Male  B) Female  C) Prefer not to say

4. What is your native language?
A) Japanese  B) Chinese  C) Korean  D) Other Asian language  E) English  F) Other European language  G) Other

5. How would you rate your English language reading ability?
A) Native reader  B) Advanced level non–native reader  C) Intermediate level non–native reader  D) Basic level non–native reader

6. Do you have any of the following challenges?
A) Differentiating between colours  B) Co-ordination of hand–eye movement  C) Dyslexia  D) Other  E) None that I know of  F) Prefer not to say

Section 2 – About you and technology for educational purposes

7. What kind of technology do you regularly use to help you study English?
A) Television or DVDs  B) Radio or CDs  C) Computer  D) Smartphone or Tablet  E) Other technology (               )  F) Nothing

8. What area is your main focus when using technology to study English?
9. Do you have a smartphone?
A) Yes   B) No

10. If your answer to question 9 is 'yes', what is the name of the operating system software (OS) on the smartphone that you have?
A) iOS7  B) iOS6  C) Android Jellybean  D) Android Ice Cream Sandwich  E) Android Gingerbread  F) Android (system unknown)  G) Windows Phone 8  H) Windows Phone 7  I) BlackBerry  J) Other

11. Do you have a tablet computer?
A) Yes   B) No

12. If your answer to question 11 is 'yes', what is the name of the operating system software (OS) on the tablet?
A) iOS7  B) iOS6  C) Android Jellybean  D) Android Ice Cream Sandwich  E) Android Gingerbread  F) Android (system unknown)  G) Windows Phone 8  H) Windows Phone 7  I) BlackBerry  J) Other

13. Have you used any language learning software (apps) on a smartphone or tablet device?
A) Yes, everyday  B) Yes, often  C) Yes, sometimes  D) Yes, occasionally [Go to question 14]  E) No, never [Go to question 17]

14. How many different apps have you used?
A) 1  B) 2–4  C) 5–7  D) more than 7

15. What are the main reasons for you to use the app(s)?
A) grammar practice  B) vocabulary practice  C) listening practice  D) speaking practice  E) reading practice  F) writing practice  G) Other (    )

16. What would you like to do more with apps in the future?
A) grammar practice  B) vocabulary practice  C) listening practice  D) speaking practice  E) reading practice  F) writing practice  G) Other (    )

Section 3 – About the TEC App

17. How was the navigation system on the TEC module? (Was it easy to move around the app?)
A) easy to navigate  B) quite easy to navigate  C) quite difficult to navigate  D) difficult to navigate
18. Taken as a whole, how were the instructions for the different sections on the TEC module?
A) easy to understand  B) quite easy to understand  C) quite difficult to understand  D) difficult to understand
19. Which section of the module did you enjoy using the most? Choose 1 or 2
A) Video  B) Video task  C) Language and vocabulary  D) Audio game  E) Pronunciation practice  F) Questions and answers  G) Language Focus 1 & 2  H) Test
20. Which section of the module did you not enjoy using the most? Choose 1 or 2
A) Video  B) Video task  C) Language and vocabulary  D) Audio game  E) Pronunciation practice  F) Questions and answers  G) Language Focus 1 & 2  H) Test
21. Taken as a whole, how quickly were you able to understand the purpose of each part of the module?
A) very quickly  B) quite quickly  C) quite slowly  D) very slowly
22. Taken as a whole, how quickly were you able to understand how to use each part of the module?
A) very quickly  B) quite quickly  C) quite slowly  D) very slowly
23. Taken as a whole, how do you feel about the feedback you received in the module?
A) very good  B) quite good  C) quite bad  D) very bad
24. Do you feel that using this module helped you learn any English?
A) Yes, a lot  B) Yes, a little  C) No, it didn’t  D) Not sure
25. How do you feel about paying for software to learn English on a smartphone or tablet?
A) I will not pay  B) I will only pay if I have to (for example, if a teacher tells me to buy it)  C) I will pay sometimes  D) I often pay
26. If you were going to use this app in the future, how much would you be willing to pay for a package of three modules?
A) ¥1,000 or more  B) ¥750–¥999  C) ¥500–¥749  D) ¥250–¥499  E) ¥1–¥249  F) ¥0
Appendix 2

Audio game screen image.