An Analysis of the Effectiveness of an In-house CALL Software Package for the Learning and Teaching of kanji (Japanese Characters) and the Development of Autonomous Language Learning Skills

Debbie Corder (deborah.corder@aut.ac.nz)  
Grant Waller (grant.waller@aut.ac.nz)  
Auckland University of Technology, New Zealand

Abstract
Mastering kanji, Japanese characters, is one of the major hurdles for students in achieving proficiency in the Japanese language. With increasing class sizes, differences in prior knowledge, levels of ability, and approaches to learning, it became a priority to explore methods to increase the effectiveness of teaching and learning kanji for students on the Japanese programmes at the Auckland University of Technology. This article provides an initial report on the evaluation of a CALL software package developed in-house to provide an alternative way for students to learn kanji and to facilitate the development of autonomous learning. The evaluation combined computer tracking of 37 students in their first semester of study, with interviews, questionnaires, and student diaries. Initial findings suggest that the software is meeting different learner needs and that there is a correlation between student usage of the software and improved performance in assessments. Some students are developing autonomous learning characteristics, but it is clear that not all students know how to use the software effectively. It would seem that the software alone does not facilitate the development of autonomous learning and that there is a complex relationship between a number of factors, including differences in learning strategies and motivation.

Introduction
The difficulty of mastering character-based languages such as Japanese is well documented (McCarthy, 1996; Van Aacken, 1999). The contact hours for the BA (Japanese) over three years at the Auckland University of Technology (AUT), is around 1260 hours. However, mastery of such languages requires an average of 2600 contact hours compared to 960 for Western languages (Van Aachen, 1996). This is because to master the writing system - kanji (Japanese characters)? a student is required not only to know the meaning of each kanji, but also how to write it, read it (up to six or seven readings for each kanji), and to learn the related vocabulary of kanji compounds (when one or more kanji are combined to form related words). The time commitment required to gain a reasonable degree of mastery to read, say, a newspaper can be considerable. It is not surprising that the demands of kanji learning have been identified as one of the major contributing factors to high attrition rates in Japanese in the first year of study (Van Aachen, 1996, p.2). Students on the Japanese programme at AUT are expected to master 1000 kanji and even with four dedicated kanji papers amounting to 600 hours, it is challenging, especially for non-kanji background students. By 1999, addressing student needs, especially in relation to kanji, was becoming a significant learning and teaching
issue. The number of students studying Japanese had been gradually increasing since 1996, and along with this came changes in their cultural and ethnic mix. In particular, the number of Asian students increased, with a smaller increase in Maori and Pacific Island students. This brought an increased variability in student learning needs, styles and abilities, and prior knowledge. This CALL project grew out of the need to explore methods to increase the effectiveness of learning and teaching kanji. Because of the time commitment required for mastery, any method would also have to enable the students to work as independent learners in an autonomous learning environment.

There are various definitions and interpretations of the term ‘learner autonomy’ and it is often equated with learner independence whereby learners are independent to the extent that they can work on their own without help from a teacher (Blin, 1999, p. 134). The point of departure for this project was based on Holec's 1981 model: students become independent learners through autonomous learning, which is the process whereby they exercise control and assume responsibility for their learning by making decisions or choices, touching on all of its aspects from goal setting to self-assessment. In short, the common goal of all definitions is that students take more responsibility for their learning and move towards self-reliance (Boud, 1988, p. 23). The task, therefore, was to provide a framework to facilitate this development.

The decision to introduce CALL as a method to increase the effectiveness of learning and teaching kanji was based on a number of reasons: the desire to meet different learning styles and to provide students with a learning environment in which they could work autonomously; evidence from research that CALL software is effective for language learning (McCarthy, 1995, 1996), including kanji (Van Aacken, 1996); its potential to cater for individual learning needs, increase motivation (Levy, 1997; Soo, 1999) and promote independent learning (Levy, 1997; Van Aacken, 1996). It was not intended to use the software in isolation from the rest of the Japanese programme, but as one of the learning methods in the complete learning environment, along the lines suggested by Hoven (1999, p. 163).

It was decided to develop the software in-house because of the lack of relevant commercial software. Research has shown that for software to be effective, it must be relevant to the curriculum (Levy, 1997; Van Aacken, 1996). The sourcing of a HyperCard database of kanji ecards called ‘QTkanji’ from Saeko Komori (1996) of Chubu University, was the catalyst. The software was trialled in semester one 2000, and formal evaluation has been ongoing from semester two 2000. This paper focuses on the findings of the initial evaluation of the data from two cohorts of first-year students in 2001 (37 students).

The evaluation of the software sought to determine whether it could be effective in improving student learning of kanji, its relationship with the students’ learning process, whether it provides the environment for autonomous learning, and whether this, in turn, promotes the development of autonomous learning.

Software description

The QTkanji-based software is being replaced by a cross-platform version, kanjiMaru, so it will only be described briefly here. The new version will be described elsewhere. The software was for both classroom use and self-access with the focus on drilling with
repetitive exercises. We supported McCarthy's (1995) thinking that drilling is an essential part of language learning especially when the learner is outside the country of the target language.

The design had to meet the needs of non-kanji, Chinese, and Korean. Komori and Zimmerman (2001) point out the different word recognition strategies used by Chinese subjects (reliance on visual information), whilst English subjects rely more on phonological information. Non-kanji students have to make the transition from the alphabetic form of writing to a logographic system. However Chinese students, whilst familiar with the logographic system, have to unlearn the Chinese pronunciation of the kanji, and in some cases the number of strokes or even meaning and relearn the Japanese version. The amount of prior knowledge of Korean students depends on their age, and whilst they have an advantage in that the pronunciation of many Japanese words is the same in Korean, they have difficulty in pronouncing some sounds.

To cater for individual learners, activities have been incorporated to develop reading and listening skills, and tasks that require students to identify kanji or their readings either by typing or by pointing and clicking with the mouse. Komori’s original database came with a video clip of each kanji showing the stroke order, providing a visual dimension, and a facility to draw the kanji using the mouse.

The program consists of different levels of study, and within each level, there are four stacks:

1 Lessons
2 Practice writing (reading/typing)

3 Practice listening
4 Tests

The design is simple with the layout of menus, and the positioning of buttons the same throughout. Immediate feedback is in the form of a buzzer, flashing light, or correct answer. The point scoring system and summaries of errors against correct answers allow students to analyse and pinpoint gaps in knowledge and understanding.

Unlike commercial software, the program can be easily customized to match any kanji textbook, at any level of kanji learning, so it is always relevant.

Implementation

The software was formally introduced into the first-year papers in semester one 2001. There were no trial and control groups because of the ethical issue about whether one group of students would be gaining an advantage over the other. Taking McCarthy's (1995) approach of not adding to the students’ workload, students were timetabled to use the software for one session per week with a teacher available and had self-access opportunities. Teacher assisted sessions were not made compulsory for two reasons: the aim to cater to individual learning needs and to provide an autonomous learning environment for students to study independently. Apart from a brief introduction to the program, students did not have any more formal sessions on how to use the software. However, it soon became apparent that students need ongoing support and guidance, not in terms of technology, but in terms of using it to meet their needs at various stages in their learning.

From observation, discussions with students, and analysis of the tracking data, it was clear after six weeks that a number of students were not using the software effectively,
and that usage was trailing off. Workshops on strategy awareness were held in week seven after the first formative assessment, to help students become more aware of their learning process and more receptive to different ways of learning including the software. Attendance was voluntary, and all first-year students were invited during a timetabled student forum hour to avoid additional time demands. Disappointingly, fewer than half attended. Feedback from those who attended indicates that sharing each other's learning strategies was the most useful exercise. These students started using the software and cue cards more consistently, and their assessment results improved. For example, one student went from 38% for the first test, to 72% in the second and went on to pass the paper. It is not clear to what extent this can be attributed to increased use of the software alone, but it would suggest that it provided some benefits.

**Research Methodology**

The methodology used both quantitative and qualitative approaches and looked at the learner and the learning environment, as well as evaluating the effectiveness of the software. By looking at student approaches to their language study as a whole, it sought to identify whether students were showing characteristics of autonomous learning, and how this related to their use of the software. The student approach to study was monitored by a longitudinal case study involving questionnaires, focus groups and interviews, and student diaries. The software was modified to automatically track student use of the various activities. The tracking data and coded questionnaire responses were analysed using SPSS.

The first questionnaire given in week five of the first semester sought background information including age, gender, first language, language study (prior knowledge, study preferences), computer experience, and initial impression of the software. A second questionnaire given at the end of their second semester of study focused on the use of the software. The tracking data and coded questionnaire responses were analysed using SPSS.

Participants

As shown in Table 1, there were 37 students in the initial evaluation.

<table>
<thead>
<tr>
<th>Non-kanji</th>
<th>Chinese</th>
<th>Korean</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 (F)</td>
<td>5 (F)</td>
<td>8 (F)</td>
<td>23</td>
</tr>
<tr>
<td>3 (M)</td>
<td>7 (M)</td>
<td>4 (M)</td>
<td>14</td>
</tr>
</tbody>
</table>

Of the 37 students, 23 returned questionnaires (see Table 2). Their ages ranged from under 20 to 50, with approximately 75% in the 25 and under age group. The Korean students tended to be in the older range and the non-kanji in the younger. Seven students had less than one year of Japanese study.
Table 2.
Composition of the students who returned questionnaires (n=23)

<table>
<thead>
<tr>
<th></th>
<th>Non-kanji</th>
<th>Chinese</th>
<th>Korean</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (F)</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>Male (M)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Keeping diaries was voluntary, and over half who kept them were non-kanji students because of the reluctance of the Asian students to write diaries (see Table 3).

Table 3.
Composition of students who agreed to be interviewed and to keep diaries (n=17)

<table>
<thead>
<tr>
<th></th>
<th>Non-kanji</th>
<th>Chinese</th>
<th>Korean</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (F)</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Male (M)</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

The students were asked to use diaries to record their reflections on their learning experiences and management of the study. They were also invited to take part in focus groups to get used to discuss their learning experiences and to listen to other students talking about their learning difficulties and strategies, which in turn could give them ideas to reflect upon their learning. They were asked to comment on their reasons for studying Japanese, how they went about learning it, what strategies/methods worked for them, including the software, and how they monitored their progress. A number of students also agreed to be interviewed one-on-one, and similar questions were put to them.

Findings and discussion

Five research questions provided the framework for the data collection. Two questions related to software design, and three to effectiveness for kanji study. The findings relating to design have been used to develop the new version and are discussed elsewhere. The latter three will be discussed here.

1) **Do students actually use the software, and how are they using it? Are there any trends based on individual differences (background, gender, prior learning)?**

The software is being used by students in the three main language groups? non-kanji, Chinese, and Korean, and on the whole the non-kanji students have made the most use of the software, closely followed by the Korean female students. The extent to which the female Korean students made use of the software was unexpected and possibly significant that they were high achievers in Japanese, and many were mature students. The trends show that there are differences in use by individuals, and by gender and language group and that the non-kanji male students have launched the software the most, followed by the female Korean students. The male Korean students have launched it the least. In terms of how they use the software, preferences for stacks are similar for all students, except
for Korean males. This might not be truly representative of Korean males as the sample size was small. The Chinese students tended to be the most resistant on the whole to use the software. However, subsequent findings indicate changing attitudes.

The period of greatest use was in the first six weeks, after which there was a gradual overall decline with small peaks leading up to test and examination weeks. This reflects the trends in other CALL research projects (Hatasa and Hatasa, 1997; McMeniman and Evans, 1998) and was anticipated in this project from the weekly tracking data and classroom observation. However, despite the decline, the software was used continuously by 90% of the students (33 out of 37) until the final few weeks of the semester: only four (10.8%) had stopped using it by week eight, nine (24%) showed reasonably consistent use and two (5.4%) showed high-frequency use. This trend is promising, bearing in mind that the use of the software was not compulsory. Whether this continuous use would have occurred without teacher intervention, is open to question.

Initially, the general loss of the sustained effort to use the software was simply attributed to a lack of student motivation to keep using it. However, this was a rather simplistic explanation and failed to take into account the range of possible contributing factors including:

- The change in perception of the software as an effective method for learning, or the novelty of the software wearing off (Levy, 1997) especially if results were not immediately obvious, and the realization that drilling was necessary. This might have stemmed from students’ inability to adjust the use of the kanji software to meet changing needs, which might, in turn, be due to lack of language learning strategy awareness (Toyoda, 2001; Hoven, 1999; Blin, 1999; Van Aachen, 1996).
- The amount of teacher guidance received by a student on how to use the software to meet individual learning needs.
- The usual ebb and flow of motivation of students over some time in response to assessment results, fatigue, and the general effect of learning a language in an institutional setting (Ushioda, 1996).
- The demands of the non-language parts of the course (assignments and examinations).
- External factors such as personal and financial problems.

Some factors cannot be changed, such as the demands from other parts of the course, or as Van Aachen (1996, p.12) points out, there will always be a handful of students who will not accept using computers as a learning method. However other factors, such as motivation and learning style can change (Ellis, 1994, p. 479). This was evidenced by data from several students who repeated the kanji paper. Their patterns of use show very different trends from the first semester, indicating more consistent use throughout the repeat semester, with marked periods of reduced use that seem to coincide with assessment periods in other parts of the course. These changes in use appear to have coincided with changes in approach, and with what Ushioda (1996) identifies as essential elements in autonomy: being able to sustain self-motivation and engage intrinsic motivation. For example, one student who is repeating the paper writes... I failed kanji I in the first semester... I decided to make more of a constructive effort to learn and be more enthusiastic... The more I see, read, write, say, and pretty much breathe the kanji...
characters? the more it is becoming common knowledge . . . I love the feeling of knowing something, being able to read and write kanji that just baffled me last semester. As L2 learning is ‘sustained deep learning,’ it requires this sort of continued practice (persistency) to gain progress, and additional motivational factors to sustain attention and effort (Dornyei and Otto, 1998).

When students were first introduced to the software, the lesson stack was the most popular as students appeared to like the drawing function and the video clip. However, after the second week, the writing (reading/typing) stack proved the most favoured stack for most students. Many said they liked the demands on accuracy from having to type. The test stack (point and click) proved the least favoured but the Chinese students opened it the most frequently and the Koreans the least. Not surprisingly, the non-kanji students made the most use of the video clips showing the kanji stroke order, and the Chinese students the least. The Korean students made the most use of the sound files, possibly reflecting the fact that one of their main weaknesses is pronunciation. There appeared to be no real difference in preference of stacks in terms of gender except for the Korean males who preferred the lesson stack while all other groups preferred the writing (reading/typing) stack.

From the tracking data, it has not been possible to identify how students have been using the software in terms of their learning approach, for example identifying whether they have been using metacognitive strategies to plan how they are going to use the software, whether they have goals in mind, and whether they are evaluating themselves. The focus groups and interviews, and some diary entries especially from the second semester provided some insight, but the informal discussions with the students during class time proved the most useful. In hindsight, it might have been useful for students to keep logs of how they were using the software and why, or even more simply, to complete a checklist after each session.

**Effective use**

It was evident that some students knew clearly what they wanted to achieve in a session and would use different activities for revision, practice, and testing, whilst others would spend a great deal of time on less demanding interaction such as drawing kanji using the mouse, and/or using the sound files to listen at great length to the native speaker pronouncing the vocabulary in the lesson stack. Whilst the sound files provide excellent guidance on pronunciation, it might have been beneficial for these students to move more quickly onto the listening stack where they could also test their kanji recognition. The students struggling to draw kanji using the mouse might have been better served using a pencil and paper. Others went from one activity to another without reaching the error summary, or finished activity with 50% or less correct. They got neither feedback on their level of mastery nor spent time analyzing and correcting their errors. These tended to be students who used the software irregularly or ceased altogether. Some students responded to teacher guidance, and their patterns of use changed to some extent. A number of students only just began to use the software again in the last week before the final test and examination.

It became apparent that several factors were linked to consistent and effective use. These factors are not mutually exclusive, and include:
• Sustained intrinsic motivation and self-motivation for study as a whole, and not necessarily high achieving in terms of tests and examination.
• Enthusiasm for learning kanji.
• Some awareness of and willingness to use different learning modes (including the software) and strategies generally to meet learning needs.
• Perception of the value of spending time using the software, especially for drilling.
• Regular attendance to receive teacher guidance on how to use the software more effectively to suit individual needs.

2) **Is there a correlation between student usage of the software and improved performance in tests and examinations?**

The kanji paper has three tests and one final kanji exam. Pearson correlation coefficients, with a two-tailed test of significance, were calculated as an indication of the effectiveness of the software in improving student performance in assessments. Because student feedback from the quantitative data had indicated a strong preference for the writing (reading/typing) stack, correlations were also calculated between assessment results and the frequency of usage of the writing (reading/typing) stack.

There is no significant correlation between the overall use of the software (as indicated by the total number of times a student launches the program) and performance in the assessments. However, the figures for total writing (reading/typing) stack launches (the total number of times a student launches and uses the stack) display a moderate correlation (0.412) at the 0.05 confidence level for test 2, and a stronger correlation (0.521) at the 0.01 confidence level for test 3, and the final examination (0.471).

This relationship is indicated graphically in Table 4, which shows the average assessment results in each of the four assessments for the 30%(n=37) most frequent users of the writing stack (high users) contrasted against the 30% least frequent users (low users).

Table 4: Correlation between writing (reading/typing) stack use and performance in assessments (n=37)
Analysis of individual cases confirms that there appears to be a correlation between how the software was used and the subsequent performance in assessments. Students who mainly used the writing (reading/typing) stack as a part of their study tended to maintain or improve their assessment grades throughout the paper. There also appears to be a correlation between the number of practice scores accumulated by an individual within the writing (reading/typing) stack (which indicates the length of time spent practising within this stack on drilling until errors were eliminated) and performance in the assessment.

The significance of this finding is apparent when compared to trends from previous semesters. The 250 kanji studied in the first-semester range from easy pictographs to more complex kanji beyond the prior knowledge of the average high school student. Test scores have tended to fall as the kanji become progressively more difficult. Consequently, students who initially achieved a low first test mark have often ultimately failed the paper. For example, in the year before the introduction of the software, 82% (19 out of 23) of those who failed the paper, had failed or had borderline passes for the first test. However, when individual students in this study were analysed, a different pattern emerged. Although the failure rate was still high with 14 out of 21 who failed the first test ultimately failing the paper, those who had been using the software consistently and effectively passed the paper. The same applied to students who gained average marks in the high 60s or 70s. These students would have tended to achieve a bare pass, but many were able to maintain their average because they used the software effectively. Students in the 90s who would have tended to fall to 80s or high 70s by the end of the paper were maintaining their high averages. The repeating students who passed well on their second attempt showed much more effective use of the software than they had done in the previous semester.

3) Does the software provide an environment in which the learner can work autonomously? Allied to this, to what extent are students showing signs of independent learning, and how much does the software feature in the total learning environment?

From the tracking data, it is clear that students have been making their own decisions about which stacks to use and which kanji to study. In this respect, it can be said that the software provides an environment in which the learner can work autonomously. From the diary entries, interviews, focus groups, and teacher/student dialogue, some students are demonstrating characteristics of autonomous learning, such as reflection on their learning process and experiences, goal setting, and increased decision making in their learning, including consciously using the software as part of their overall kanji learning process (Corder and Waller, in press). This is especially apparent by the end of their second semester of study, as can be seen from responses to questions in the second questionnaire:

Do you think you have developed any different learning strategies from using the kanji software?

- It has showed me a different way of learning language (Chinese)
Based on the writing stack, I have developed my own study strategy applying the same techniques, hiragana to kanji, and kanji to hiragana. (non-kanji)
• Yes, kanji cards don’t work for me therefore what else can I do? I use the writing stack and do each lesson at least twice. (non-kanji)
• Yes, using a different way to review the kanji. (non-kanji)

Has the software made a difference to your mastery of kanji?

• Yes, it makes it easier to learn kanji. I am sure if I did not have the software I would have found kanji learning much more boring and difficult. (Korean)
• Yes, when I use it properly, I get much better marks in my tests. (non-kanji)

Because teacher intervention (increased guidance on software use and strategy awareness-raising workshops) was necessary, it is clear that the software alone was not sufficient to facilitate autonomous learning. This supports Hoven's (1999) argument that whilst software can give learners more control over their learning, the corollary is that they need to understand their learning processes and be able to use and develop effective learning strategies to use the software effectively. Initial analysis of diaries, interviews, and focus groups, as well as discussions with students, indicate that there is a close relationship between a student's overall approach to study, including the use of learning strategies, motivation and their ability to work independently, and their use of the software. This would support Blin's (1999) view that students need to be autonomous to some extent to benefit from the full potential of CALL software.

The extent to which the intervention, or indeed the research, influenced the development of characteristics of autonomous learning is not clear. Student feedback indicates that taking part in the research has made them reflect more on their approach to learning. However, it could have been the methods used in the research, namely those which lead to greater student/teacher dialogue that was a contributing factor. It would also seem that students may need time to adjust not just to CALL software but to the study demands of tertiary education and that some adjust faster than others. Most students need varying levels of support and guidance to develop the tools required for independent study, and this is why scaffolding is important.

Conclusion

Realistically CALL as a method of study will not appeal to all students. However, the findings from the initial evaluation indicate that the kanji software is effective as an alternative learning method for students to master a very difficult aspect of Japanese. The initial evaluation has limitations, primarily the small size of the group of students. However, findings do appear to be similar to a number of findings from other CALL research projects. These include declining use of the software, the link with motivation, the need for careful integration of the software into the learning programme, and the need for learner training in language learning strategies so that they can use the software effectively to meet their learning needs.
In addition, the initial evaluation of the software, which has also looked at autonomous learning, has highlighted the importance not just of motivation, but of the affective aspect of motivation, and the role of the teacher in helping students to engage intrinsic motivation, maintain self-motivation and develop the ability to reflect on learning. This latter point is significant when the main attraction of CALL software is the argument that students can work at their own pace with the minimum input from teachers.

The findings strongly indicate that taking steps to ensure the software is effective for language learning requires more than careful integration in terms of content. This has implications for introducing CALL software into a teaching programme. Providing students with software, no matter how relevant to the curriculum, will not in itself guarantee that students will use it, let alone use it effectively, or that it will facilitate autonomous learning. The findings point to a complex relationship between, on the one hand, student awareness of the language learning process, their needs, and ability to manage and maintain motivation, and on the other the language learning environment including assessment, expectations (including learning and teaching), the curriculum and the influence of teachers.

In terms of ensuring software is effective for learning, is meeting individual learner needs, and facilitates autonomous learning, the findings indicate the need for:

- Learner training on how to use computer software effectively.
- A support structure on the lines of the ‘scaffolding’ concept. It is not sufficient just to build support into the software; there needs to be interaction and dialogue with the teacher to ensure feedback, guidance, and ongoing development of skills such as analysis and reflection.
- Learner training in terms of language learning strategy awareness and awareness of individual learning needs.
- Integration of the software into the teaching programme not just in terms of content but also as a teaching and learning method.
- Increased awareness of all the teaching team, of the role of the software, and how it can be used to meet learner needs.
- Learner feedback on language learning experience to encourage them to reflect on their learning behaviour, and to capitalise on both negative and positive experiences.
- Increased staff awareness of the importance of motivation in language teaching, and its relationship with autonomous learning.
- Software design focusing on activities that actively engage students rather than more passive use. For example, typing as opposed to video clips and point and click, and inclusion of effective feedback to allow for error analysis.

The initial findings have confirmed the potential of CALL for enhancing student learning and encouraged further research and development. Further research will include analysis of data on student use of the software over two years of study and the role of the teacher and teaching team as a whole. Further software development will include grammar and reading comprehension practice, and software to help students manage their
learning, including activities to develop cognitive and metacognitive knowledge and strategies for learning kanji to reduce memory load.

References


