Acquisition of Kanji and German vocabulary using jMemorize

Markus Rude
Tsukuba University

Reference data:
Rude, Markus (2008). Acquisition of Kanji and German vocabulary using jMemorize.
In K. Bradford Watts, T. Muller, & M. Swanson (Eds.), *JALT2007 Conference Proceedings*. Tokyo: JALT.

There is no proof that vocabulary-training programs are better than traditional methods of vocabulary acquisition. However, instructors could use such programs for their own language studies and transfer their experience to their classes. This paper describes such an attempt using the SW (software) jMemorize: the instructor set his goal to study 1000 Kanji within 18 months, limiting his daily practice time. At the same time, he introduced the software to his students and encouraged them to input and share Japanese-German vocabulary. This paper presents findings from the individual study, including the marked negative effect of interrupting the regular study for just a couple of days, or the problem of TOT (tip of the tongue) phenomena, which can sometimes be overcome by a shift of attention. Though some of the results could be transferred to the classroom, the students' motivation to use jMemorize on their own was limited, possibly due to the requirement of regular usage or even daily usage of such a SW.

語彙トレイニングのソフトウェアが伝統的な語彙習得学習方法に勝っていることは証明されてはいない。しかし、教師がそういったソフトウェアを自分の言語習得のために用い、その経験を授業で活用することは可能である。本論ではドイツ語教師である筆者がjMemorizeというソフトウェアを用いた試みを紹介する。筆者は、毎日の学習時間を決め、18ヶ月かけて1000語の漢字を習得することを目標としている。同時にドイツ語の授業にそのソフトウェアを導入し、各学生がドイツ語の語彙を他の学生と協力して学習するように指導する。本論の試みが示唆することとして、以下のことが挙げられる。数日間でも規則的な学習を怠った場合、語彙の習得数の低下は著しい。思い出せそうで思い出せない語彙は一時的に他の語彙を想起することで思い出せることがある。この手法を授業の語彙チェックで用いると学生の緊張による失敗を回避することが可能である。これらの結果を実際の授業で用いたが、このようなソフトウェアを使用した継続的学習の習慣化を促すことは困難であり、学生が jMemorizeを自発的に活用するよう動機付けるには限界があった。

While there are many software programs for vocabulary acquisition, there is little scientific research about their efficiency. Should we – as *instructors* – apply software in our classroom as long as there is no proof of its usefulness? One possibility is to act as a test person, using the software for our own language acquisition. If we are sufficiently satisfied with the program, we can introduce it in our classes.

In the first section of this paper, one vocabulary trainer – jMemorize – is introduced. The main part of this paper is section 2, describing the acquisition of Japanese Kanji with this trainer (Application A). The learning goal of one autonomous learner (the author) and the learning schedule are presented. The section also discusses the amount of data to be recalled per item and time schedules for repetition. Finally, some techniques which support successful recalls are treated.

The following section describes the usage of jMemorize in class for German vocabulary acquisition (Application B). The underlying idea of cooperative vocabulary acquisition is being outlined and experience from the introduction of the software in class is reported, including some student feedback.

The last section contains an overall discussion and gives a tentative conclusion.

What is jMemorize?

jMemorize (Djemili 2007) was selected among many vocabulary training programs because it can handle various languages (Unicode), it is freeware and open-source, and it runs on various platforms (Windows, Mac-OS and Linux; Java required). Furthermore, it is easy to learn, the statistics are intelligible and the time schedules for repetition can be defined by the user (For other vocabulary training programs, see Rude 2008 and Barranco-Droege 2006).

jMemorize is a vocabulary training software, which uses a time schedule for *spaced repetitions* of language items (Spaced repetition means repeating items after increasing time intervals and is the opposite of *massed repetition* where repetitions happen within a short time, e.g. the day before a test). The users feed the program with vocabulary items they want to learn and repeat those like flashcards. They define a *time schedule* for repetition and then start a learning cycle. Upon seeing a stimulus (*frontside* of flashcard) they try to recall given entries (contained on the *flipside* of that flashcard) and decide manually, whether they knew the answer or not. Based on the time schedule and time-stamps of the current repetition (Fig. 1, column with header 'Last Test'), jMemorize calculates the expiration time of items, which means the due date of the next repetition (column with header 'Expires'). Cards which are past this expiration date (red icons), or which have been forgotten in the last test or not been learned at all yet (gray icons) can be repeated in a repetition cycle. Cards, which have not yet expired, count as 'learned' (yellow or green icons).

Application A: Individual acquisition of Kanji with jMemorize and Excel

This section treats the learning goal, the learning schedule, the question of how many entries per item should be required from the flipside during recall and the time schedule. Concerning the actual retrieval process, the importance of previewing, delaying and skipping (related to priming), and of mnemonics (e.g. keyword method) are discussed. As for the contents of flashcards, it is argued to use also input related to the arts. It concludes with some hints for individual users of jMemorize.

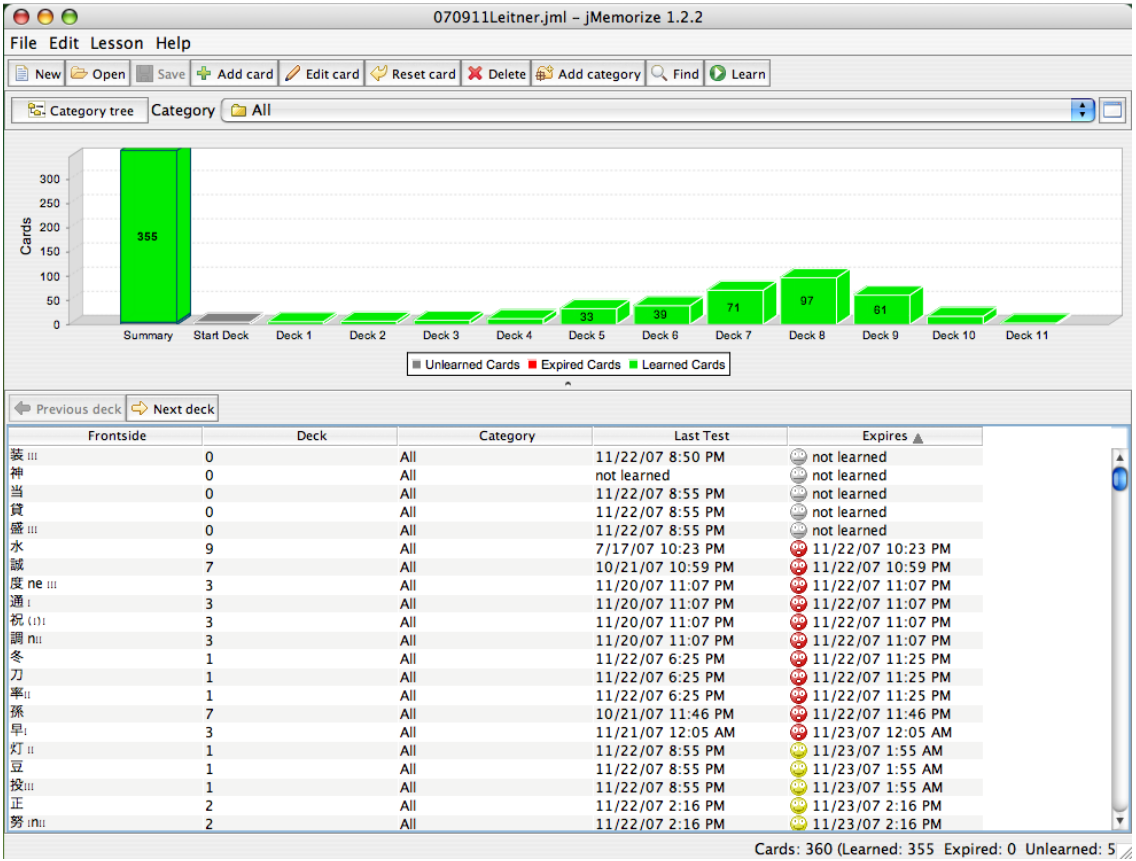

Figure 1. Main window of jMemorize for the application of Kanji acquisition.
Each line of the list represents one 'flashcard' to be learned. Frontside: Kanji. Flipside: data to be learned (invisible).

Learner and goal

The learning goal was to acquire the 1006 Kanji, which Japanese students master at primary school in a receptive way (only reading, eventually writing by computer) within about 18 months. I am a false beginner having already spent numerous hours with studying Japanese including Kanji. However, I always gave up at some point and never became a fluent reader.

Learning schedule: How much learning time is required?

Because of individual differences, this question cannot be answered in a univocal way. It could be reformulated to 'How much learning time is available?' I decided to limit the pure *repetition time* including the imput of new Kanji to 1 hour per day (6 cycles x 10 min)

However, for increasing the success rate, a quick preview before and a quick review after a repetition cycle have been added. Additionally, being subject *and* experimenter, I need some minutes per session for the modification of cards (adding item related data and statistical data). Therefore, the complete *study time* is about 2 hours per day (6 sessions x 20 min).

Hence, I mostly study six sessions per day, three with jMemorize and three with Excel, a newer version of (Rude 2006). Fig. 2 shows the learning curve of Kanji – accumulated from jMemorize and Excel – over a period of roughly 12 months (December 2006 to November 2007). 'Learned' in jMemorize means, that all entries of an item could be recalled during the last repetition, and that the card has not expired yet. It does not mean 'learned' in a psychological sense, e.g. being stored in long-term memory. Each dot represents the average number of learned Kanji in a given calendar week, calculated from the seven values from the end of each day of that week. The progression is more or less constant at about 1.7 Kanji per day in the average. Though globally increasing, locally the curve decreases four times for a week or two, always due to lack of daily study time.

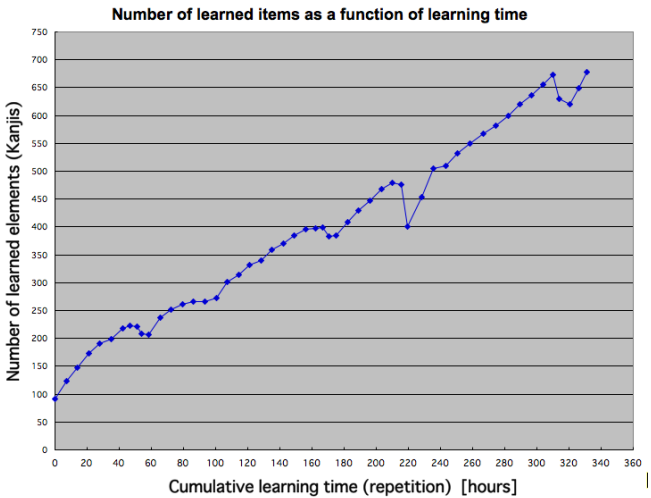

Figure 2. Learning curve

This figure shows how many Kanji counted as learned over one year (52 dots represent 52 weeks). About 590 Kanji were newly acquired during this period. The effort is shown in terms of pure repetition time (about 340 hours). The setbacks are due to study pauses.

Each of the last two setbacks of the learning curve shows a sharp drop: both were due to a study pause of 3 to 5 days. The effect of these rather short pauses is an eye-opener and I am not surprised, that former attempts to master the Kanji were not successful. Though jMemorize yields only a very rough model for forgetting, differentiating only between the discrete states 'learned', 'expired' and 'unlearned' (a probabilistic model would be more appropriate), the message is clear: we should study on a daily base, at least until we have sufficient knowledge to maintain our knowledge otherwise, e.g. through regular reading.

One problem after study pauses is the large number of meanwhile expired Kanji, which *accumulate*. A second problem comes from the decreasing *success rate* when trying to recall cards that expired some time ago. The third problem stems from *motivation*: It is demotivating, if it takes one or two weeks just to reach the old level again from before the study pause.

Number of entries per item: How much information should be recalled per item?

The frontside of a flashcard (stimulus) contains an individual Kanji; the flipside (entries to recall from memory) is a given Kanji's Chinese and Japanese readings (Onyomi and Kunyomi) and meaning. Fig. 3 is an example for an item, where three entries have to be recalled from memory.

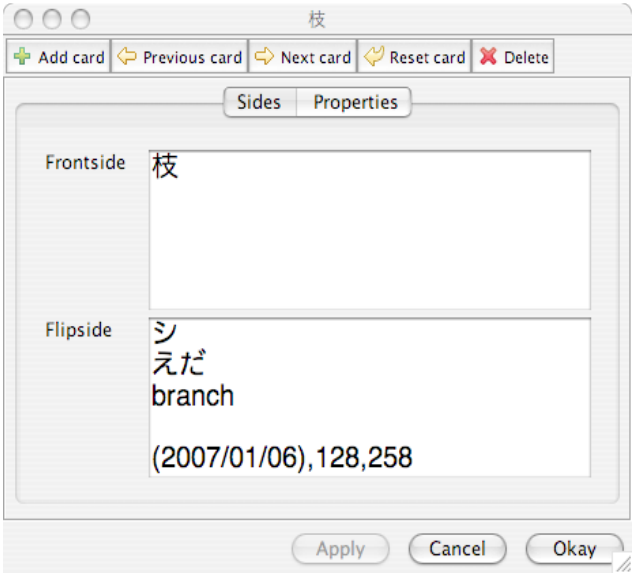

Figure 3. Simple item with three entries to be recalled. *Frontside: Kanji character '枝'. Flipside: Onyomi, Kunyomi and meaning 'branch'.*

Mostly the additional recall of up to three examples like compounds, phrases or sentences is required. Fig. 4 is an example for such a complex item, where seven entries have to be recalled from memory.

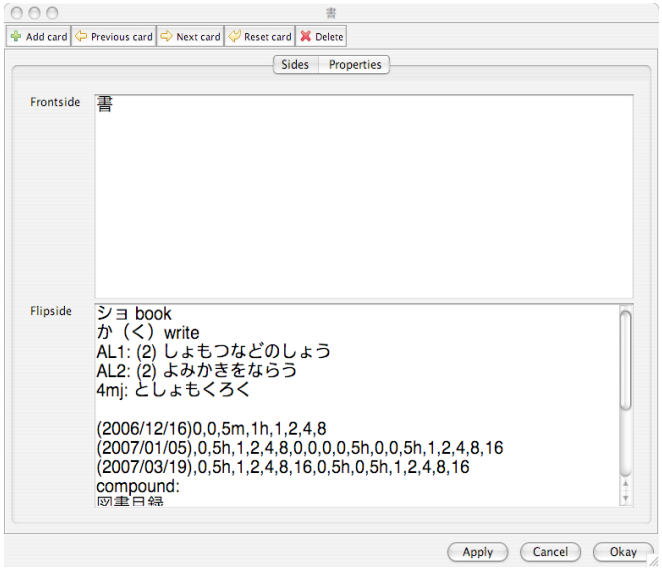

Figure 4. Complex item with seven entries to be recalled. *Frontside: Kanji character '書'. Flipside: Onyomi, its meaning 'book', Kunyomi, its meaning 'write' and three examples.*

On one hand, it is advisable to require additional examples in order to grasp the whole concept of a Kanji. Additionally, each example constitutes a mental link to other Kanji and thus helps to create a densely interconnected *mental lexicon*. Often, phrases and sentences are better retained then Onyomi or Kunyomi alone.

On the other hand, the retrieval time will increase considerably with each additional entry required for recall. Since jMemorize offers only two answers ('Yes' or 'No') for each recall trial, complex items create an additional problem: how should the user decide in cases, when all but one of the entries could be recalled? For such borderline cases, I introduced the special marker '!!!', to be added to the frontside (see Fig. 1). It allows a quick *review* of all borderline Kanji after the repetition session, in which only the missing entries are recalled or looked up. This procedure strengthens exactly the memory traces of the missing entries and thus increases the chances for complete recalls in the next repetition cycle.

Each time, a missing entry can be successfully recalled in successive trials, one exclamation mark can be deleted. The marker must disappear after a while, since the goal is to recognize individual Kanji without additional hint. However, if the missing entry can not be recalled in a successive trial, the answer should be 'no', giving the card the status 'unlearned' (icon turns gray).

Alternately, some programs offer more than just two answers 'yes' and 'no'; they calculate the due dates for the next repetition accordingly.

Time schedule: which spacing is best?

Should we repeat each item every day (constant schedule), or after 1, 2, 3, 4, 5, 6, 7, 8, ... days (linear schedule), or after 1, 4, 9, 16, 25, 36, 49, 64, ... days (quadratic schedule) or after 1, 2, 4, 8, 16, 32, 64, 128, ... days (exponential schedule)? The answer depends on the context of learning, on the

characteristics of the items, on individual differences, etc. Therefore the following comments should be understood as individual experience that cannot be generalized:

I tried the exponential schedule and the quadratic schedule and prefer the latter: The exponential progression is rather strong, such that I often had to repeat Kanji that were completely forgotten. On the other hand, with the quadratic, more modest progression, I occasionally thought to have forgotten a Kanji, but – just an instant later – its entries popped up from my memory. Or I could not recall some essential part but felt that I knew it.

However, for other learners, other languages or other entries on the flipside, different time schedules might be better. The software SuperMemo seems to use a quite sophisticated method to calculate the time schedule, taking the history of past successes and failures into account (Wozniak 2002).

Priming: previewing, delaying and skipping

Priming is a psychological term and describes the phenomenon, that a subconscious stimulus influences our behavior, e.g. by decreasing our reaction time to related conscious stimuli. *Previewing*, *delaying* and *skipping* are three techniques, which help to prepare the actual retrieval process and make it more likely to be successful. The three terms are no psychological terms, but related to priming because subconscious processes are involved.

In this context, *previewing* denotes the quick scanning of a whole group of Kanji that will be repeated next, recalling just the Chinese readings (Onyomi). Although brief, this activation of some neural structures is likely to help the recall of the remaining entries from the flipside during repetition thereafter. This preview can additionally be considered as a modest speed-reading practice. Occasionally, an entry cannot be recalled, though we *know* that we know it. This phenomenon it called TOT (tip of the tongue) phenomenon. The conscious attempt to recall such an entry is often not possible and typically ends in a wrong guess or just in frustration. *Delaying* the retrieval attempt by proceeding to the next entry of the same item or to the next item – thus shifting the attention temporarily – sometimes helps. Letting the problematic TOT entry slip into subconsciousness and returning to it some time later helped me often to recall it. *Skipping* is a similar technique and refers to the act of not trying to recall the flipside of a card at all in a given session, but to leave it to the next repetition cycle.

I could observe cases of successful recall of TOT items with all three techniques. The rationale in using such tricky techniques is an underlying assumption about learning processes: active recall is always better than just looking up the answer.

Mnemonics: keywords, key sentences, abbreviations and acronyms

Mnemonics are not an implicit part of vocabulary trainers, however, their usage can considerably improve their efficiency. I recommend using the keyword method, key sentences, abbreviations and acronyms – without trying to be complete – by just introducing two representative examples:

A *key sentence* helped to remember a certain sound which I often mistook: the Kanji 尊 means 'respect' and I always mistook its Onyomi as 'zon' instead of 'son'. The key sentence '村長を尊重する / son-chou wo son-chou suru / to respect the chief of the village' helped, since the critical sound appeared in both homophones 'son-chou', one of which I already knew. A beneficiary side effect of such sentences involving homophones is the ease of remembering them and their meaning and – indirectly – the two contained pieces of vocabulary.

In another case an *acronym* helped to remember the meanings of a Kanji: the Kanji 急 means 'urgent, sudden, abrupt' and I often forgot one of the three. Adding the abbreviation 'USA" and remembering 'USA > urgent, sudden, abrupt' just adds a piece of *redundancy* to the information structure, which is so helpful for successful recall, and which is also an inherent characteristic of language processing in the brain. See for example Booji (2005) for redundancy rules with respect to the mental lexicon (p. 18).

Introducing the arts: music, proverbs and nonsense sentences

Repetitive patterns, slightly varied, are structures with which the human brain can deal with ease and with fun, and which can be found in music, poetry or in arts in general (Sousa 2006). Furthermore, these elements or – more general – elements which stimulate our emotions, our limbic system, stick better to our memory than linguistic structures without emotional content: we like to think them and we like to remember them (p. 213ff). Again just some examples:

Linking *songs* to individual Kanji and – if you like singing – just sing a line of it upon retrieval. E.g. '泣 / naku' means 'to cry, to weep'. Children song: '昔 泣き虫 神様 が、嬉しくても 泣いて、悲しくても 泣いて / mukashi na-ki-mushi kami-sama ga, ure-shi-ku-te mo na-i-te, kana-shi-ku-te mo na-i-te / Long time ago, there was a god who was a cry-baby. He cried when he was happy, he cried when he was sad.'

Japanese *proverbs* in the form of 4-Kanji-compounds, linked to an individual Kanji. E.g. '業 / gyou, gou' means 'occupation, business, industry, achievement'. 4-Kanji-Compound: '自業自得 / ji-gou-ji-toku / As a man sows, so shall he reap.'

Nonsense sentences linked to Kanji, best when they stimulate our imagination. E.g. '灯 / tou' means 'light, lamp'. Nonsense sentence: 'ホテル の 点灯係 は、蛍 か てんとう 虫 だ？ / hoteru no tentou-gakari ha, hotaru ka tentou-mushi da? / Who is in charge of lighting in the hotel, fireflies or ladybugs?'

Hints for individual Kanji acquisition with jMemorize

1. Decide your learning goal: learning for a test and learning for life require completely different approaches, the former allowing cramming or massed repetition, the latter requiring spaced repetition.
2. Decide how much time you can devote for learning on a regular base.
3. If a study pause is inevitable, try to catch up on the following days.

4. Choose appropriately sized portions of information to be recalled.
5. Use previewing, reviewing, delaying, skipping and mnemonics.
6. Use markers for borderline items with just one missing, forgotten entry.
7. Introduce the arts or other topics with emotional and imaginable content.
8. Modify your cards constantly, e.g. if you find structural regularities in language input (Ellis 2002): "Language learners have to figure language out" (p. 144).
9. Ask a native speaker to correct your input every once in a while.
10. Keep it simple. Do not try to follow all hints.

Application B: collective acquisition of German vocabulary in language classes

This section describes, how jMemorize has been introduced in our German classes as an optional method for vocabulary acquisition. It further reports on students' reactions, on an incentive to work harder on vocabulary, and on some transfer. It concludes with some hints for collective vocabulary acquisition.

Background and procedure

The vocabulary required in our standardized test after one year comprises about 500 words. Our textbook (Vögel 2002) covers a lot of this vocabulary. It also offers in its vocabulary section useful chunks and sentences and thus provides a solid base for a language class which emphasizes vocabulary acquisition.

Lüders (2005) describes how to cooperatively acquire vocabulary in a class: the vocabulary is divided among all students, the instructor needs to correct the input only once and provides all the students with the correct data by email. The goal was to establish a similar system in my German language classes.

One of the lessons was therefore held in the computer room of Tsukuba University and jMemorize was introduced through a detailed, bilingual worksheet. The vocabulary of a previous lesson was provided to the students. They just had to download the software jMemorize from the Internet, to install the program and to start it. Then the vocabulary file could be opened and a learning session could be started, using a time schedule that let cards expire within minutes ('Cram' time schedule). Optionally, students could enter the vocabulary of another lesson.

Some comments from the feedback sheets

Informal feedback sheets have been used and a majority of the students stated, jMemorize would be useful for vocabulary acquisition. Also a clear majority expressed the intention to use jMemorize in the future. However, we have to be careful with such feedback in the Japanese context, since some students

tend to write friendly feedback instead of critical feedback. Eleven students from almost 100 volunteered to input vocabulary and sent it to me by email. After correction, I sent all files to about 15 students, who wished to receive the files, whether they had supplied vocabulary or not.

Creating an incentive for vocabulary acquisition and spaced repetitions

Whereas the way of vocabulary acquisition is up to the students, it is up to the instructor to hold the students accountable for their effort. This is being done in two ways: on one hand, 20% of each of the eight weekly oral tests consists of a brief vocabulary check concerning the vocab homework. This homework realizes spaced repetitions, as it requires the students to review vocabulary of previous lessons less and less frequently (e.g. "Repeat the vocab of last week's lesson 15 plus the vocab of lessons 13, 9 and 1"), realizing an *exponential time schedule*. There is also one larger vocabulary test per trimester (20% of final grade).

Transfer from individual Kanji study to collective vocabulary acquisition

The application of the *delaying technique* (see above) for vocabulary checks in the classroom is an example for transfer: if a student cannot recall a certain word but seems to know it, I tend to ask for a second word, coming back to the first word later. In some cases, the recall was successful after such a shift of attention; the TOT problem could be resolved. Even if the final trial should be without success, the technique can contribute to reduce stress.

Hints for collective vocabulary acquisition with jMemorize

1. Prepare a detailed worksheet stating the goals of the class, including some hints from section 2.8.
2. Provide the vocabulary to be repeated, including some keyword examples.
3. Suggest the usage of the 'cram' time schedule, such that cards expire and turn red within the lesson time.
4. Divide the rest of the vocabulary among the students, or just ask for some volunteers who would like to enter some vocabulary and to share it with others.
5. Add an incentive, which encourages the students to work on their vocabulary, like vocab tests, or short, graded vocab checks in regular classes.

Discussion and tentative conclusion

Discussion

Long-term retention needs a considerable study effort (Application A). The goal of acquiring 1000 Kanji in 18 months seems to be reachable, but still some of the knowledge is not really in long-term memory, as the curve for one individual (the author) in Fig. 2 demonstrates. One positive result is the confidence in one's knowledge of Kanji: the systematic repetitions help to keep the learned items in a mental state that is easily accessible, also when seeing Kanji in daily life.

It might be interesting to investigate, whether students using jMemorize know vocabulary better compared to other students (Application B). However, learning for an announced test and learning for long-term retention are quite different, such that students cramming for a test the night before with conventional methods (massed repetition) might have an advantage compared to students, who use jMemorize on a daily basis (spaced repetition).

Tentative conclusion

As for Application A, the acquisition of Kanji by an autonomous learner, jMemorize can support a steady progress in the acquisition process through systematic repetition of learned Kanji in conjunction with continuously entering new Kanji (Fig. 2). jMemorize mainly calculates due dates for the next repetition for each item and presents them after that due date. However, experience has shown, that a learning pause of just five days has a marked negative effect onto the otherwise progressive learning curve. Experience has also shown the positive effects of techniques such as previewing, delaying and skipping, as well as using mnemonics like keywords, key sentences or acronyms. The data reveals, that the initial goal of mastering about 1000 Kanji within about 18 months seems reachable, assuming a learning time of about 1 hour/day (pure repetition of Kanji and basic input of new Kanji) to 2 hours per day (overall study time including previewing, reviewing, modification of cards, e.g. by adding mnemonics, key sentences, etc.). The progression corresponds to about 2 new Kanji per day.

As for Application B, the acquisition of German vocabulary in a classroom setting it can be said, that jMemorize can easily be introduced, explained and used in one lesson of 75 min, assuming that the students have basic knowledge in using a computer and an Internet browser and that a well-prepared worksheet and the vocabulary is available. The short-term response of the students was very positive, however, since the future usage was voluntarily, only few students used jMemorize thereafter. This finding reflects a strong and a weak point of jMemorize: it is very transparent and easy to learn, though, it does not encourage the usage of the techniques described above (this is up to the user who enters the data); also audio is not yet available (the next version will have audio). If the usage of a SW trainer shall be made mandatory in a class, a program which supports audio should be chosen.

Though both applications seem different, the experience the author gained through using jMemorize for his own study could partially be transferred to the classroom use, e.g. spaced repetition, the delaying technique for overcoming the TOT phenomenon or the keyword method.

Markus Rude is an Associate Professor at Tsukuba University, where he teaches German. He formerly taught German and English at Dokkyo and Rissho University, respectively. His interests are teaching pronunciation or prosody and the usage of computers for language learning.

References

Barranco-Droege, R. (2006). Memorization Software Reviewed. www.quingle.com/softarea/flash-su.htm [retrieved on December 15, 2007].

Booji, G. (2005). *The Grammar of Words*. Oxford: Oxford University Press.

Djemili, R. (2007). jMemorize Version (Version 1.2.2) [Software for Windows and Mac OS. Open-Source. Freeware]. jmemorize.org [retrieved on December 15, 2007].

Ellis, N. C. (2002). Frequency Effects in Language Processing. *Studies in Second Language Acquisition, 24*. 143-188.

Lüders, J. (2005). Wortschatzlernen mit der Lernkartei? Anspruch und Wirklichkeit [Vocabulary acquisition with the word card learning box? Claim and reality]. *PRAXIS Fremdsprachenunterricht. Die Zeitschrift für das Lehren und Lernen fremder Sprachen, 2*(1), 28-32 [München: Oldenbourg].

Rude, M. (2006). How to create a vocabulary learning file with Excel. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.), *JALT2005Conference Proceedings*. Tokyo: JALT.

Rude, M. (2008). Does vocabulary-training software support neuro-compatible vocabulary acquisition? In K. Bradford-Watts (Ed.), *JALT2007 Conference Proceedings*. Tokyo: JALT.

Sousa, D. A. (2006). *How the brain learns*. (3rd ed.). London: Sage Publications Ltd.

Vögel, B. & Azra, J.-L. (2002). *Gespräche im Unterricht* [Conversations in lessons]. ISBN: 4-9901072-0-9.

Vögel, B. & Azra J. (2002). *Gespräche im Unterricht* [conversations in class]. Kyoto: ALMA Verlag.

Wozniak, A. P. (2002). SuperMemo [Software for Windows]. www.supermemo.com/ [retrieved on December 29, 2007].