

The Forgetting Curve and Learning Algorithms

Jeff Hanks* and Ping Zhan**

Abstract

With the globalization and impact of Internet, a new wave of English as public language is occurring in Japanese corporations. Learning English in an efficient ways is still a challenge. In this manuscript, we introduce the forgetting curve and spacing effect, algorithms in the software SuperMemo. We also propose an experiment project on how to test learning English words based on the spacing effect. A pre-experiment result is also given.

Index terms: forgetting curve, spacing effect, spaced repetition, SuperMemo, learning algorithm, experiment.

1. Introduction

There are now (1,750,000,000) one billion, seven hundred and fifty million people speaking English globally. In Europe and Asia, the non-English-speaking world, the number of companies where English is used as an official language is growing. Rakuten – the Japanese biggest online-shopping, announced this change recently and brought a lot of attention. It has been said that newly hired employees in Samsung Electronics, the South Korean multinational electronics company, are required to score over nine hundred on the TOIEC test. Other companies, such as Airbus, DaimlerChrysler, Nokia, SAP, Microsoft in Beijing and Tokyo, also use English the official language in the company [4].

The reason that a language strategy is needed now is that global business speaks English. Besides this, there are three more factors to be considered: The gradual downsizing of the Japanese market, the effect of the Internet and English media, and finally the rapid Englishnization in Asia. It is inefficient to communicate in multiple languages and English becomes a must in order to compete globally. To achieve Englishnization, there are still a lot of hurdles. The biggest one is the English level of individual employees [4].

One author is now researching on the topic in relation to supply chain management, where the phenomenon in that fluctuation in orders increase as one moves up the supply chain from retailers to wholesalers to manufacturers to suppliers is referred as the bullwhip effect. The bullwhip effect results in an increase in all costs in the supply chain and a decrease in customer service levels. Therefore, one objective of coordination is reducing the bullwhip effect. A lot of

2012年11月30日受付

* 江戸川大学 情報文化学科専任講師 語学

** 江戸川大学 情報文化学科教授 数理計画

managerial levels have been discussed. To make these managerial levels work, trust between partners is important and this trust can only be based on regular and relatively long-term communication, in the global case, a multicultural communication [6].

Now with the high-speed development, we have to learn more and learn more efficiently. Learning English based on the mechanism of the brain is the motivation for our work. A researcher of science and medicine, Professor Kuniyoshi Sakai researched the issue.

“English or Japanese, although they are two different languages, the same parts of brain are used. ...The most important thing is whether it is a first language, which is learned before 7 years old, or a second language, which is one learned after 7 years old. ...There are four separate parts of brain used in language processing, words, intonation or accent, grammar, meaning. If the field responsible for grammar has been damaged, it leads to language disability, which has been clarified recently. ...The process of learning language: When nothing about the knowledge of grammar changes, the activity of brain does not change. Then the activity of brain is in proportion to the knowledge of grammar. After six years or so, with high level of grammar, a lot of examples, conversely, show an energy-savings taking place in the brain. ...Although, like other fields, there are some disputes, that is, two assumptions: leaning—stimulated from their parents (古典な後天説), and acquisition—developed naturally with growth, an inborn ability (近年の先天説).” [5].

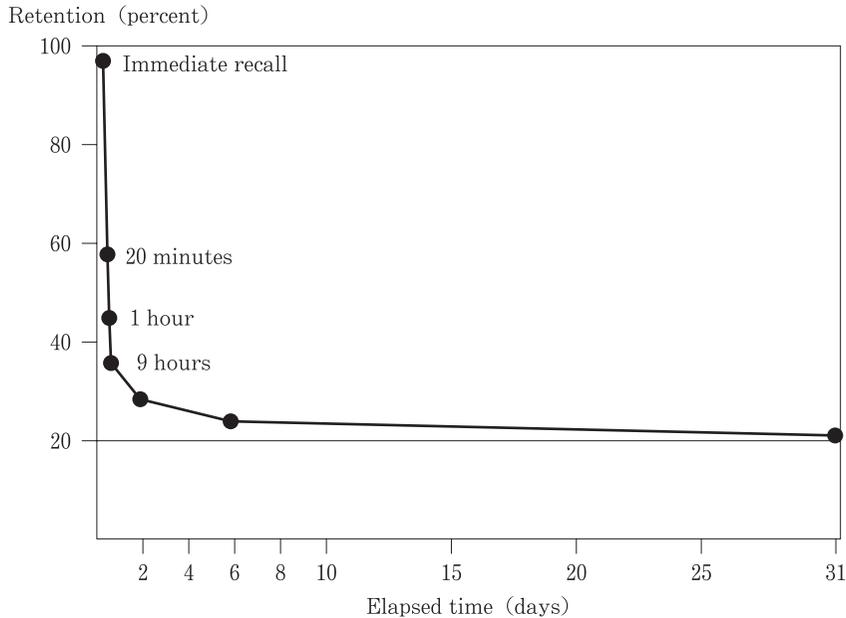
This is beyond our scope now. However, it is clear that there are some rules about learning language regarding the mechanisms of the brain. Knowing these rules and finding some efficient ways to learn English or other languages is the motivation of our current research.

The manuscript is organized as follows. We begin by reviewing the forgetting curve and spacing effect in Section 2. In Section 3 we introduce the algorithm in the software SuperMemo. A pre-experiment outline is discussed in Section 4 and a related simple experiment result is shown in the final section.

2. Forgetting curve and spacing effect

The famous forgetting curve was hypothesized by Hermann Ebbinghaus in 1885. His curve proposes to show how memory of a piece of data declines over time when there is no attempt to reinforce it. As items are reinforced the curve flattens, until the data is retained long-term.

“Ebbinghaus...used himself as a subject and memorized lists of nonsense syllables until they could be perfectly recited. Later after varying delays of up to 31 days, he relearned those same lists and measured how much less time was needed to learn them again relative to the time required to learn them in the first place. If 10 minutes were needed to learn them again after a delay of 6 hours, but only 4 minutes were needed to learn them again after a delay of 6 hours, then his memory was such that 60% savings had been achieved. As the retention interval increased, savings decreased, which is to say that forgetting occurred with the



Source: Herman Ebbinghaus, *Memory, a Contribution to Experimental Psychology*, 1885.

Figure 2-1 Ebbinghaus' classic forgetting curve.

passage of time.” [8]

By charting this decay over a period of 31 days produced his now famous forgetting curve (Figure 2-1). Ideally, reinforcement should come just before an item is forgotten, but not too soon. This is termed the “spacing effect”. The spacing effect shows that reinforcement over a longer period is more effective than repetition over a short period of time (cramming). This is where “spaced repetition” comes in.

Spaced repetition is a method of learning designed to take advantage of the spacing effect. It was first proposed by C.A. Mace in *Psychology of Study* in 1932. Pimsleur language courses, developed in the 1960s, are the most famous example of second language teaching based on spaced repetition. The Leitner system, developed by Sebastian Leitner in the 1970s is a method of using flashcards based on spaced repetition.

Mathematical models of forgetting have also been researched in industrial production. Three types of models are classified, VRVF (Variable Regression to Variable Forgetting), VRIF (Variable Regression to Invariant Forgetting) and LFCM (Learn-Forget Curve Model) [3] These models were developed in an attempt to maximize worker production by determining the optimum length of breaks for factory employees in order to minimize the forgetting of tasks upon returning to their work. As these models focus on tasks as opposed to words or data, they are of little use to us in our proposed experiments.

3. Learning algorithm

Learning algorithms about some special patterns have been developed in machine learning, artificial intelligence and neural networks. Surprisingly however, there is little research into the optimal spacing of repetition for new ESL vocabulary. To our knowledge, referenced papers on a general learning algorithm have not been published in official academic journals. Here we introduce the learning algorithm in the software SuperMemo provided on its web site. SuperMemo is the learning software that we intend to use in our work. SuperMemo, developed in Poland by Piotr Wozniak in 1985, was the first example of PC software created to exploit spaced repetition. Since that time, a number of programs and web based applications have been developed, one of the most popular of which is “Anki”, an open source flashcard application which is available on multiple platforms, including most recently, smart phones. For ease of understanding, we begin with the early version of the algorithms, Version 2 [7].

First, we define some notations.

$I(n)$: inter-repetition interval in days,

EF : easiness factor reflecting the ease of memorizing and retaining a given item in memory,

q : quality of the response, the 0–5 grade scale defined as follows:

5 – correct immediate response,

4 – correct response with hesitation,

3 – correct response recalled with serious difficulty,

2 – incorrect response; where the correct one seemed easy to recall,

1 – incorrect response; the correct one remembered,

0 – complete forgetting.

Initializing:

$EF = 2.5$,

$I(1) = 1$,

$I(2) = 6$.

The main procedure, the recursive formula:

$I(n) = I(n-1) \cdot EF$,

$EF' = f(EF, q)$

for all $n > 2$.

Where EF is restricted between 1.1 and 2.5, performs better if EF is larger than 1.3 (if EF is less than 1.3, set it to 1.3) and $f(EF, q)$ is given as $EF + (0.1 - (5 - q) \cdot (0.08 + (5 - q) \cdot 0.02))$.

Termination:

If $q=4$ or $q=5$.

Re-initializing:

If $q \leq 2$, set $I(1) = 1$, $I(2) = 6$, and without changing EF .

Denote $f'(q) = (0.1 - (5-q) \cdot (0.08 + (5-q) \cdot 0.02))$, note the values of $f'(q)$ are $-0.8, -0.54, -0.32, -0.14, 0, 0.1$ for $q=0, 1, \dots, 5$. And if q takes a lower value, the interval for repetitions is shorter. Since recursive occurs when $q < 4$, EF is decreasing. Also keeping q larger than 1.3 means the interval is increasing.

In version 4 of the algorithm, there are two steps in the recursive formula

Step 1: $temp = I(n-1) + I(n-1) \cdot (1 - 1/EF)/2 \cdot (0.25 \cdot q - 1)$

Step 2: $(1 - fraction) \cdot I(n-1) + fraction \cdot temp$.

Where $temp$ is an auxiliary value used in calculations, $fraction$ is any number between 0 and 1. Note for $q < 4$, $temp < I(n-1, EF)$ and also the result of Step 2, which is inconsistent with the concept of spacing effect. And it is revised in version 5.

In version 5, a new factor, optimal interval factor OF is introduced. And the recursive steps are as follows:

$$I(n, EF) = OF(n, EF) \cdot I(n-1),$$

where OF is calculated in two steps.

Step 1: $OF \cdot (0.72 + q \cdot 0.07)$

Step 2: $(1 - fraction) \cdot OF + fraction \cdot OF'$

Note again that $0.72 + q \cdot 0.07 < 1$ if $q < 4$.

Although EF as a function of q is renewed with the same formula f , no details about the relation of EF and OF have been shown. This is also the case in later versions of algorithms. In Version 11, Not OF but dOF , the decrement of OF is recursively calculated, and EF , easy factor, is replaced by AF , Absolute difficulty Factor, the description is also not complete.

4. Pseudo Optimal interval proposal

The ultimate objective is to find the pseudo-optimal inter-repetition interval. One opinion holds that the optimum inter-repetition interval is likely to be the longest interval that avoids retrieval failures. These are the maximum intervals that the human memory is capable of

between forgetting and remembering. This is consistent with the spacing effect and also coincides with the following in some sense. In chance theory or credibilistic theory (可信性理論) or uncertainty theory, 0.5 (the probability value, or the fuzzy characteristic value) is the most noticed value, in this case (farthest distance between the success (remembering) or failure (forgetting)). At this point brains work most actively.

Our pre-test procedures are just for an intuition into understanding the direction of the experiments. What is possible and what are the problems we need to solve? With so many variations of students' backgrounds and variations in words, there is a question as to whether we can obtain results that verify our assumptions.

In Section 3, we see that with the fact of easiness, the intervals are vary from 1.1 to 2.5, that is, vary from the unit interval and double roughly the former interval. For the [double], it is not only the medium, also with the accumulation of the words in previous days and some repetition, we HAVE TO and CAN decrease the time necessary.

The simplified version for pre-testing experiments is as follows.

- 1) Based on Ebbinghaus' famous forgetting curve, that is, one hour, the next day, the week-end, and the end of month.
- 2) Double the former interval (medium, or average procedures), Note that the intervals here also form a geometric sequence with the factor 2.
- 3) If the experiments are done in the class lecture, a geometric sequence may be difficult because it would last more than one semester, in this case, we approximate it with a sequence, $a-1, a-2, \dots, a-n, \dots$, with $a-(n+1)-a-n=n$.
- 4) Equal intervals (a common or simple one for comparing)

An approximate example of a geometric sequence appears at <http://www.supermemo.com/help/faq/algorithm.htm>.

- 1) The 1 st repetition should be scheduled 30 minutes after reading words for the first time
- 2) after an hour
- 3) after 9 hours
- 4) after 24 hours
- 5) after 3 days
- 6) after 6 days
- 7) after 12 days

Our experiments will concentrate on words because a vital component of language learning is vocabulary. However in teaching a second language two questions arise. First is what vocabulary to teach. The second is how to deliver that vocabulary to students in an effective way. There are no other methods except for simple accumulation. Without a rich and varied vocabulary, you can only transmit simple messages. It is hard to communicate your personality and

make an impact on others, which is one main shortcoming of Japanese (or Japanese leaders) now. The learning process is relatively simple, that is, not dependent too much on personal factors.

As for what to teach, there are a variety of word lists used in ESL. The most famous of which is the British National Corpus (BNC). In Japan there is the JACET 8000. Another alternative is the General Service List (GSL) developed by Michael West, in 1953. There are pros and cons to each of these lists, but for the purpose of our proposed research, we have selected the Academic Word List (AWL), developed by Averil Coxhead at Victoria University in Wellington. It was developed as a companion to the GSL, but focuses on words occurring frequently in academic texts. The reason for selecting this list is that a majority of the words will be unfamiliar to our experiment's subjects (Edogawa University students), but will be potentially useful to them in their careers.

5. Experiments

We are very fortunate that before we begin this work, one student of the Department of Communication and Business, Tetsuya Inaba, who is interested in language and the mechanisms of the brain, did some experiments on retention of English and German words according to Ebbinghaus' forgetting curve. We summarize some of them.

The words are from textbooks, 1,550 English words and 1,320 German words were selected. The experiment period lasted from August to November. Inaba thought that German is much more difficult to remember than English. He grouped 26 English words and 20 German words into one unit. It was also assumed that the forgetting rate of German words would be much higher than for English words. Table 5-1 is the result.

Where, the first column indicates interval time, the second and third columns indicate the average words remembered versus words forgotten and the remembering rate.

Because the number of German words is only 20, less than the 26 of English words, German had an advantage over English here. The average remembering rate, 78% of English, is unexpectedly near to 73%, the rating of German. Mr. Inaba also said that he was a little astonished by the results.

Table 5-1 The remembering rate

Intervals	English	German
20 min	21 : 5 26	18 : 2 20
1 hour	23 : 3 26	15 : 5 20
1 day	19 : 7 26	14 : 6 20
a week	19 : 7 26	14 : 6 20
a month	19 : 7 26	12 : 8 20

Acknowledgments: The authors are very grateful to Tetsuya Inaba for providing us the pre-experiment results.

References

- [1] John Henderson, *Memory and Forgetting*, Routledge, 1999.
- [2] Michael Jacob Kahana, *Foundations of Human Memory*, Oxford University, 2012.
- [3] David A. Nembhard, Napassavong Osothsilp, An Empirical Comparison of Forgetting Models, *IEEE*

Transactions on engineering management, Vol. 48, No. 3, 2001.

- [4] 特集, グローバル英語力, *Harvard Business Review*, p. 108-115, 21012/10.
- [5] 酒井邦嘉, Language Acquisition and Brain Development, (脳科学が明かす言語習得のメカニズム), *Harvard Business Review*, p. 108-115, 21012/10.
- [6] Sunil Chopra, Peter Meindl, *Supply Chain Management* 3rd, Prentice Hall, 2006.
- [7] www.supermemo.com/english/ol/sm4.htm, <http://www.supermemo.com/english/ol/sm2.htm>
[http://www.supermemo.com/english/algsml1.htm#Advanced repetitions](http://www.supermemo.com/english/algsml1.htm#Advanced_repetitions)
- [8] Sergio Della Sala, *Forgetting*, Psychology Press, 2010.