

Verbal working memory, long-term knowledge, and statistical learning

Satoru Saito, Graduate School of Education, Kyoto University, Japan

Masataka Nakayama, Kokoro Research Center, Kyoto University, Japan

Yuki Tanida, United Graduate School of Child Development, Osaka University, Japan

Correspondence to:

Satoru Saito

Graduate School of Education, Kyoto University

Yoshida-honmachi, Sakyo-ku, Kyoto

606-8501 Japan

Phone & fax: +75 753 3066

E-mail: saito.satoru.2z@kyoto-u.ac.jp

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article is available via its DOI shown below.

Saito, S., Nakayama, M., & Tanida, Y. (2020). Verbal working memory, long-term knowledge, and statistical learning. *Current Directions in Psychological Science*, 29(4), 340–345. <https://doi.org/10.1177/0963721420920383>

Abstract

Evidence supporting the idea that serial order verbal working memory is underpinned by long-term knowledge has accumulated over more than half a century. Recent studies using natural language statistics, artificial statistical learning techniques, and the Hebb repetition paradigm have revealed the presence of multiple long-term knowledge that underlies serial order verbal working memory performance. These include (1) element-to-element association knowledge, which slowly accumulates through extensive exposure to the exemplar; (2) position-element knowledge, which is acquired through several encounters with the exemplar; (3) whole sequence knowledge, which is captured by the Hebb repetition paradigm and acquired rapidly with a few repetitions. Arguably, the former two are a basis for fluent and efficient language usage, and the third for vocabulary learning. Thus, statistical and possibly episodic learning mechanisms may form the foundation of language acquisition and language processing, which characterize linguistic long-term knowledge for verbal working memory.

Keywords: Working memory, Statistical Learning, The Hebb repetition effect,

Long-term memory

Verbal working memory, long-term knowledge, and statistical learning

The transient retention of information in the service of mental processes is an essential cognitive function in our daily activities such as reading, reasoning, mental arithmetic, and car navigation. Because of its centrality in human cognition, this memory function - *working memory* - has been one of the most active research targets in psychological science (e.g., Baddeley, 2012; Barrouillet & Camos, 2015; Engle, 2018; Logie & Cowan, 2015). Accordingly, there have been a large number of research debates on the characteristics of working memory. Among which, its relation to long-term memory is always the front line topic in the field (e.g., one of the benchmarks of working memory; Oberauer et al., 2018).

On the one hand, this short-term retention mechanism provides a foundation for long-term learning. Working memory for verbal material, for example, forms the basis of vocabulary acquisition, as it is achievable by temporary retention of the phonological form of new vocabulary (Baddeley, Gathercole, & Papagno, 1998; Gathercole, 2006; Majerus, & Boukebza, 2013; but see, Melby-Lervåg et al., 2012). Besides, there is plenty of evidence for the contribution of long-term knowledge to short-term retention

of verbal sequences assessed by working memory tasks, which has been reported in seminal studies (Baddeley, Conrad, & Hull, 1965; Miller, 1956) and in other subsequent studies (see, Thorn & Page, 2009 for a comprehensive collection of related papers).

In this article, we focus on verbal working memory and its relationship to linguistic long-term knowledge with an emphasis on the serial order information. We first reviewed a set of studies that highlight the influences of two types of long-term knowledge on verbal working memory functioning: (1) element-to-element association and (2) position-element association. Thereafter, the Hebb repetition effect, which is assumed to reflect (3) whole sequence knowledge, is introduced as a phenomenon reported from one of the laboratory-based learning paradigms. According to recent findings from this and other related paradigms, we conclude that multifaced long-term knowledge is acquired from the language environment and underpins verbal working memory performance.

The Effect of Element-to-Element Association Knowledge

Learning from the Natural Language Environment. The transitional statistics of

linguistic elements in the natural language environment influence verbal working memory performance. It is known that bi-phone frequency, which is a co-occurrence frequency of two successive phonemes in the language environment, affects immediate serial recall in English. That is, people performed better with a list of nonwords that consist of high frequency bi-phones (e.g., rin = ri + in) than those that are low frequency (e.g., keb = ke + eb) (e.g., Gathercole, Frankish, Pickering, & Peaker, 1999). The generality of this phenomenon is confirmed by the effect observed at a level of subsyllabic unit called mora in Japanese (the bi-mora frequency effect, Tanida, Ueno, Lambon Ralph, & Saito, 2015; Tanida, Nakayama, & Saito, 2019).

Learning in experimental settings at a laboratory. A study by Majerus, Martinez Perez, and Oberauer (2012) demonstrated the influences of element-to-element association knowledge that is acquired in the laboratory to verbal working memory. They combined the standard statistical learning paradigm (Saffran, Aslin, & Newport, 1996) and an immediate serial recall task, where individuals were exposed to a series of stimuli (e.g., a syllable sequence) with immanent statistical structures, and were tested with the immediate serial recall of sequences either compatible or incompatible with the

to-be-learned structure. In some of their experiments, participants were familiarized with an unsegmented auditory sequence of 3,795 syllables for about 37 minutes. There are syllable transition rules in the sequence: e.g., "ko" followed "zin," which is a regular transition, but not "chu," is an irregular transition. Crucially, each of the regular combinations was presented more than 200 times - massive exposure to the target transitions. As was the case with the natural language statistics, these artificial rules restrict later recall performance with the higher scores for regular sequences than irregular ones.

The Effect of Position-Element Association Knowledge

Learning from the Natural Language Environment. There is another type of statistics in the natural language environment - distributional statistics of linguistic elements. For example, /ŋ/ appears at the end, but not the beginning, of a syllable in English. In contrast, it can appear at the head of a syllable in Vietnamese (Warker, Dell, Whalen, & Gereg, 2008). Thus, the occurrence frequency of each linguistic element differs among within-syllable positions or within-word positions, and the patterns of

these distributional statistics vary among languages. This type of statistics affects verbal working memory performance. Nakayama, Tanida, and Saito (2015) reported that the position-element frequency in the natural language affected immediate serial recall performance of Japanese nonwords lists - a high frequency position-mora combination was recalled better than a low-frequency one. This was observed even after controlling for other types of frequencies, e.g., bi-mora frequency (this means the position-element frequency uniquely predicts recall performance).

Learning in experimental settings at a laboratory. Thus far, only two studies have experimentally confirmed the effect of position-element association knowledge on verbal working memory (Nakayama & Saito, 2017; Majerus & Oberauer, 2019). In a series of experiments in Nakayama and Saito, participants performed trials (e.g., 704 trials in Experiment 1) of immediate serial recall of Japanese nonwords. Some nonwords appeared at a certain serial position more frequently than at the others. The repeated position-item associations improved immediate serial recall performance for those particular items at particular positions even after removing the influences of co-occurrence frequencies. Majerus and Oberauer have incidentally found

position-word association learning in an immediate serial recall - words that consistently appeared in only one serial position were recalled better over trials than word that appeared at random serial positions.

The Hebb Repetition Effect and Whole Sequence Knowledge

The Hebb repetition paradigm, which was invented by Hebb (1961), has been developed to specifically examine the contributions of long-term knowledge to working memory. The typical procedure of this paradigm requires participants to perform many trials of immediate serial recall of verbal items (e.g., digits, words, or nonwords).

Unbeknownst to participants, the same sequence of the items is presented repeatedly in, for example, every third trial. This repeated sequence shows a gradual improvement of memory scores over the repetitions. The learning that occurred here is called Hebb repetition learning and the increased performance is named the Hebb repetition effect.

It is natural to expect that the Hebb repetition learning is based on both (1) element-to-element association and (2) position-element association as the whole sequence is repeated. However, the experimental tests on the Hebb repetition learning

ostensibly disproved this prediction.

If element-to-element (item-to-item) associations were acquired through the list repetitions and they contributed to the Hebb repetition effect, the effect should be transferred to lists that contain the same element-to-element transition but in different serial positions on the list. This was not the case; such a transfer effect was not observed (Fastame, Flude, & Hitch, 2005). Similarly, if position-element (position-item) associations were acquired through the list repetitions and they contributed to the Hebb repetition effect, the effect should be transferred to "partial repetition" lists that contain the same position-element associations but only at odd or even positions on the list. Once again this was not the case; such a transfer effect was absent in the study by Cumming, Page, and Norris (2003).

Two hypotheses have been proposed to explain this unique effect of whole sequence repetition. One is a combined learning hypothesis and the other is an episodic learning hypothesis. The combined learning hypothesis postulates that element-to-element association learning and position-element association learning counteract each other (see, Majerus & Oberauer, 2019). Any experimental

manipulations that intend to separate the two learning mechanisms potentially create a situation in which one type of learning works toward facilitation and another toward interference. When element-to-element association (e.g., a combination of items X and Y) is tested at serial positions 3 and 4 after the Hebb repetition learning with a list "JKMCXY" where items X and Y had been presented at serial positions 5 and 6, element-to-element association works to facilitate memory performance whereas the position-element association interferes with the current serial position-element binding, thus works toward reducing memory performance. Likewise, in a situation that items are repeated at only odd or even serial positions, the position-element association at the repeated positions works to facilitate recall whereas element-to-element association learning interferes with the current element-to-element binding, thus works toward reducing recall performance. When a whole sequence is repeated, then, the two learning mechanisms are able to work together toward the same direction, showing a stronger effect than either one of them.

The episodic learning hypothesis postulates that the Hebb repetition effect is derived from episodic learning where a specific context is learned with just a few

encounters. For example, an item X in a sequence ABCDX could be episodically associated with a specific context ABCD without (effective) statistical learning operations (Burgess & Hitch, 2006; Page & Norris, 2009). Once such an episodic association is formed, it contributes to future recall performance when the same specific sequence is presented. It is important to note that this repetition learning is achieved with a small number of repetitions (typically four or eight times) in the standard Hebb repetition paradigm. Such rapid learning, at least initial learning, cannot be supported by any statistical learning mechanism that, by definition, requires a sample of multiple trials to extract the statistics. This episodic learning hypothesis is also broadly consistent with an integrative framework for serial order working memory and episodic long-term memory (Farrell, 2012).

Multifaceted Long-Term Knowledge for Verbal Working Memory

The findings from the studies using the natural language statistics, artificial statistical learning techniques, and the Hebb(-like) repetition paradigms have revealed the presence of multiple long-term knowledge that underlies serial order verbal working

memory performance. These include (1) element-to-element association knowledge, which is slowly accumulated through extensive exposure to the exemplar; and (2) position-element knowledge, which is acquired through many encounters with the exemplar; (3) whole sequence knowledge, which is captured by the Hebb repetition paradigm and is acquired rapidly with a few repetitions.

Functional roles of the three types of leaning. The three types of knowledge might have their respective characteristics and different functions. Slow learning is suitable for leaning and utilizing statistical structure in a language without overinfluence of each specific example. Rapid learning is suitable for learning a specific example in a language (e.g., a word-form) without overinfluence of prior knowledge. These competing functional goals are fulfilled by dissociated but interacting systems (Baddeley et al., 1998; McClelland, McNaughton & O'Reilly, 1995). Both element-to-element association and position-element association are a basis for fluent and efficient language usage (e.g., see Warker & Dell, 2015 for speech production). The whole sequence knowledge is assumed to be a basis for vocabulary learning, supporting the acquisition of new word-forms in children and adults. We can learn a new

word-form from a few encounters with the word, and this learning ability is particularly important for young children who need to acquire a large amount of new vocabulary within a few years. This is why the rapid Hebb repetition learning is regarded as a "laboratory analog" of natural language word learning (see Norris, Page, & Hall, 2018).

Influences of prior knowledge to laboratory learning. Learning from natural language environment and artificial statistical learning are not independent but former influences the latter. Majerus et al. (2012) showed the effect of element-to-element association on immediate serial recall in an artificial statistical learning paradigm with syllable sequences but not with digit sequences. Arguably, we are exposed to a dramatically large number of random digit sequences in our daily life and such exposures boost working memory performance with digit materials (Jones & Macken, 2015). Due to such saturated learning, a few hundred repetitions of digit pairs in the laboratory could not improve the immediate serial recall for lists that contain the repeated digit pairs. It is assumed that for adult participants, any types of verbal sequences (even of syllables or letters) might contain previously experienced pairs of items, which potentially interfere or diminish learning of experimentally provided item

sequences. This assumption is consistent with the finding that young children, who have less exposure to a language environment than adults, were sensitive to statistical structure as much as whole sequence repetition (Yanaoka et al., 2019).

Future Directions

This review provides directions for future research as well as a summary of current findings. Further endeavor to elucidate as to how prior associative knowledge (if any) affects the effects of element-to-element association, position-element association, and the Hebb repetition, and as to whether multiple statistical learning mechanisms work counteractively and synergetically is required for theoretical development in verbal working memory research. These questions might drive an expansion of the current framework within and beyond verbal domains where different types of domain knowledge could be learned possibly with different learning mechanisms. Scattered examinations have demonstrated learning of differential characteristics of knowledge in different paradigms. Japanese speakers can exploit knowledge on regularity of pitch accent patterns in Japanese for immediate serial order recall of verbal materials (Ueno et

al., 2014; Tanida, et al., 2015). Within non-verbal domains (e.g., visuospatial domain), the presence of several types of statistical learning (Fiser, Berkes, Orbán, & Lengyel, 2010), and the Hebb repetition effect (Couture & Tremblay, 2006; Sukegawa, Ueda, & Saito, 2019) have been confirmed. A systematic examination of differential types of knowledge in different domains and paradigms will provide insight into the domain general and specific nature of working memory and its relation to long-term knowledge.

Conclusions

This review identified three different types of learning that characterize linguistic long-term knowledge for verbal working memory. Statistical (and possibly episodic) learning mechanisms of linguistic knowledge may form the foundation of language acquisition and language processing. This view coincides with the position that verbal working memory functioning is embedded within a large language network system (e.g., Majerus, 2013) and is an emergent property of the interaction among processes within the language system (Acheson & MacDonald, 2009; Saito & Baddeley, 2004).

Acknowledgements and end note

ORCID iD

Satoru Saito <https://orcid.org/0000-0002-0403-3606>

Masataka Nakayama <https://orcid.org/0000-0003-3904-4986>

Acknowledgements

This research was financially supported by JSPS KAKENHI Grant Number 16K04424.

Declaration of Conflicting Interests

The authors declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Notes:

1. Address correspondence to Satoru Saito, Department of Cognitive Psychology in Education, Graduate School of Education, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, 606-8501 Japan

References

- Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, 135(1), 50–68. <https://doi.org/10.1037/a0014411>
- Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63(September), 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Baddeley, A. D., Conrad, R., & Hull, A. (1965). Predictability and immediate memory for consonant sequences. *Quarterly Journal of Experimental Psychology*, 17(2), 175–177.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173.
- Barrouillet, P. & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove: Psychology Press.
- Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, 55,

627–652.

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition*, 34(8), 1720–1729.

<https://doi.org/10.3758/BF03195933>

Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory*, 11(1), 43–63. <https://doi.org/10.1080/741938175>

Engle, R. W. (2018). Working memory and executive attention: A Revisit. *Perspectives on Psychological Science*, 13(2), 190–193.

<https://doi.org/10.1177/1745691617720478>

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, 119(2), 223–271.

<https://doi.org/10.1037/a0027371>

Fastame, M. C., Flude, B., & Hitch, G. J. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, 13(3–4), 247–258.

<https://doi.org/10.1080/09658210344000314>

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception

- and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3), 119–130. <https://doi.org/10.1016/j.tics.2010.01.003>
- Gathercole, S. E. (2006). Nonword repetition and word learning : The nature of the relationship. *Applied Psycholinguistics*, 27, 513–543.
- Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 84–95. <https://doi.org/10.1037/0278-7393.25.1.84>
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). New York: Oxford University Press.
- Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*, 144, 1–13. <https://doi.org/10.1016/j.cognition.2015.07.009>
- Logie, R. H., & Cowan, N. (2015). Perspectives on working memory: introduction to the special issue. *Memory & Cognition*, 43(3), 315–324. <https://doi.org/10.3758/s13421-015-0510-x>

Majerus, S. (2013). Language repetition and short-term memory: An integrative

framework. *Frontiers in Human Neuroscience*, 7(Article 357), 1–16.

<https://doi.org/10.3389/fnhum.2013.00357>

Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports

vocabulary development: New evidence from a novel word learning paradigm.

Journal of Experimental Child Psychology, 116(4), 811–828.

<https://doi.org/10.1016/j.jecp.2013.07.014>

Majerus, S., Martinez Perez, T., & Oberauer, K. (2012). Two distinct origins of

long-term learning effects in verbal short-term memory. *Journal of Memory and*

Language, 66(1), 38–51. <https://doi.org/10.1016/j.jml.2011.07.006>

Majerus, S., & Oberauer, K. (2019). Working memory and serial order: Evidence

against associations. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are

complementary learning systems in the hippocampus and neocortex: Insights from

the successes and failures of connectionist models of learning and memory.

Psychological Review, 102(3), 419–457.

<https://doi.org/10.1037/0033-295X.102.3.419>

Melby-Lervåg, M., Lervåg, A., Lyster, S.-A. H., Klem, M., Hagtvet, B., & Hulme, C.

(2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, 23(10), 1092–1098.

<https://doi.org/10.1177/0956797612443833>

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

Nakayama, M., & Saito, S. (2017). Position-element frequency learning is dissociable from Hebb repetition learning. *Journal of Memory and Language*, 94, 235–253.

Nakayama, M., Tanida, Y., & Saito, S. (2015). Long-term phonological knowledge supports serial ordering in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1570–1578.

Norris, D., Page, M. P. A., & Hall, J. (2018). Learning nonwords: The Hebb repetition effect as a model of word learning. *Memory*, 26(6), 852–857.

<https://doi.org/10.1080/09658211.2017.1416639>

- Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A. R. A., Cowan, N., ... Morey, C. C. (2018). Benchmarks for models of short term and working memory. *Psychological Bulletin*, 144, 885-958.
- Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1536), 3737–3753. <https://doi.org/10.1098/rstb.2009.0173>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by eight-month-old infants. *Science*, 274(5294), 926–928.
- Saito, S., & Baddeley, A. D. (2004). Irrelevant sound disrupts speech production: exploring the relationship between short-term memory and experimentally induced slips of the tongue. *Quarterly Journal of Experimental Psychology*, A57(7), 1309–1340.
- Sukegawa, M., Ueda, Y., & Saito, S. (2019). The effects of Hebb repetition learning and temporal grouping in immediate serial recall of spatial location. *Memory & Cognition*, 47, 643-657.

- Tanida, Y., Nakayama, M., & Saito, S. (2019). The interaction between temporal grouping and phonotactic chunking in short-term serial order memory for novel verbal sequences. *Memory*, 27 (4), 507-518.
- Tanida, Y., Ueno, T., Lambon Ralph, M. A., & Saito, S. (2015). The roles of long-term phonotactic and lexical prosodic knowledge in phonological short-term memory. *Memory & Cognition*, 43, 500-519.
- Thorn, A. S. C., & Page, M. P. A. (Eds.) *Interactions between short-term and long-term memory in the verbal domain*. Hove, England: Psychology Press.
- Ueno, T., Saito, S., Saito, A., Tanida, Y., Patterson, K., & Lambon Ralph, M. A. (2014). Not lost in translation: Generalization of the Primary Systems Hypothesis to Japanese-specific language processes. *Journal of Cognitive Neuroscience*, 26, 433–446.
- Warker, J. A., & Dell, G. S. (2015). New Phonotactic Constraints Learned Implicitly by Producing Syllable Strings Generalize to the Production of New Syllables. *Journal of Experimental Psychology. Learning, Memory, and Cognition*.
<https://doi.org/10.1037/xlm0000143>

Warker, J. A., Dell, G. S., Whalen, C. A., & Gereg, S. (2008). Limits on learning phonotactic constraints from recent production experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1289–1295.

<https://doi.org/10.1037/a0013033>

Yanaoaka, K., Nakayama, M., Jarrold C., & Saito, S. (2019). Determining the developmental requirements for Hebb repetition learning in young children: Grouping, short-term memory, and their interaction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(4), 573–590.

Recommended Readings

Burgess, N., & Hitch, G. J. (2005). Computational models of working memory: putting long-term memory into context. *Trends in Cognitive Sciences*, 9(11), 535–541. This provides a detail model of verbal working memory and also a clearly written, relatively comprehensive review for readers who wish to expand their knowledge on Hebb repetition learning and the relationship between working memory and long-term memory.

Logie, R. H., & Cowan, N. (2015). (See References). This is an introduction to the special issue of Memory and Cognition on working memory, which provides not only simple introduction of the included papers, but also discusses future perspectives on working memory research.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). (See References). This is one of the first papers to raise attention about statistical learning in infants.

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160059. A comprehensive,

highly accessible overview of what is known about statistical learning.