International Review of Social Sciences and Humanities
Vol. 7, No. 2 (2014), pp. 106-113
www.irssh.com
ISSN 2248-9010 (Online), ISSN 2250-0715 (Print)

Chunking — A Good Way of Memorizing and Teaching Vocabulary

Xu Fang
School of Foreign Languages
Qingdao University of Science and Technology
PO Box 502, 69 Song Ling Road, Qingdao 266061, China
E-mail: sailqd_2008@163.com

(Received: 25-4-14 / Accepted: 29-5-14)

Abstract
By elaborating the definition of chunking, explaining characteristics and evidence of chunking, analyzing chunk types, expounding mental lexicon, differentiating between working memory and short-term memory, this paper states how to apply chunking to vocabulary memorization and teaching.

Keywords: Chunking, Mental Lexicon, Working Memory, Vocabulary.

1. Definition of Chunking

A term alludes to the process of taking individual units of information (chunks) and separating them into larger units. Perhaps the most common example of chunking occurs in phone numbers. For example, a phone number sequence of 8-6-1-9-3-2-5 would be chunked into 861-9325. Chunking is often a useful way when putting large amounts of information into memory. By grouping disparate individual elements into larger blocks, information becomes easier to retain and memorize.

2. Characteristics and Evidence of Chunking

The first feature is that, when individuals mistakenly recall an item in a serial recall task, it is likely to come from a similar item. For example, it may be an item that they placed in the same grouping. Individuals that use this strategy for recall will commonly misplace items they are grouping. Since one must recall items in the exact order they were presented during the serial recall task, any item that is even one position out of place is considered incorrect. Therefore, according to how many groups an individual breaks the list into, misplacement of an item will be limited to within the confine of the size of the group.

Another characteristic of the "chunking" effect is that a modality effect is present. That is, the mechanism employed to convey the list of items to the individual has an impact on how much "chunking" happens. Experimentally, it has been found that auditory presentation causes in a larger amount of grouping in the responses of individuals, as compared to visual presentation. Studies, such as George Miller's *The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information* have demonstrated that the probability of recall is greater when the "chunking" strategy is used. As stated above, the grouping of the

responses occurs as individuals divide them into categories according to their inter-relatedness based on semantic and perceptual properties. Lindley (1966) indicated that the groups produced have meaning to the participant, therefore; this strategy makes it easier for an individual to recall and retain information in memory during studies and testing. As a result, when "chunking" is obvious in recall tasks, one can anticipate a higher proportion of correct recalls.

The most persuasive evidence for the existence of "chunking" is explained in the analysis of response times. When studying this aspect of the test/response phase of a recall task, one considers response time as a function of output position. Consequently, this analysis allows for the measurement of the process of recall in each task tested. The recall or forgetting curve explains that each item in a cluster characteristically demands about the same amount of time to recall. This can be noticed as strings of items where the response times are both similar, as well as very rapid. However, one can also observe in these responses time curves that the time between the "chunks" follows a different tendency wholly. Items or periods of output where the individual is not recalling items that are connected with a group need a significantly larger amount of time. As a result, before the beginning and end of recall of a group of items in a "chunk," there is a jump in response time in the curve.

3. Chunk Types

The chunk types are based on the syntactic category part (i.e. without function tag) of the bracket label in the Treebank (cf. Bies (1995) p.35). Approximately, a chunk includes everything to the left of and containing the syntactic head of the constituent of the same name. Some Treebank constituents do not have relevant chunks. The head of S (simple declarative clause) for example is normally considered to be the verb, but as the verb is already part of the VP chunk, no S chunk happens in our example sentence.

Besides the head, a chunk also contains premodifiers (like determiners and adjectives in NPs), but no postmodifiers or arguments. This is why the PP chunk only contains the preposition, and not the argument NP, and the SBAR chunk consists of only the complementizer.

There are some difficulties when altering trees into chunks. In the simplest case, a chunk is just a syntactic component without any further fixed components, like the NPs in our examples. In some cases, the chunk contains only what is left after other chunks have been taken away from the component, cf. "(VP loves (NP Mary))" above, or ADJPs and PPs below. We will discuss some special cases during the following description of the individual chunk types.

Here we summarize 12 types of chunks: ADJP, ADVP, CLP, DNP, DP, DVP, LCP, LST, NP, PP, QP, VP (Xue et al., 2000).

Type	**Definition**
ADJP	Adjective Phrase
ADVP	Adverbial Phrase
CLP	Classifier Phrase
DNP	DEG Phrase
DP	Determiner Phrase
DVP	DEV phrase
LCP	Localizer Phrase
LST	List Marker
NP	Noun Phrase
PP	Prepositional Phrase

QP	Quantifier Phrase
VP	Verb Phrase

3.1 NP

Our NP chunks resemble the ones of Ramshaw and Marcus (1995). Especially, possessive NP constructions are divided in front of the possessive marker (e.g. [NP Eastern Airlines] [NP' creditors]) and the handling of coordinated NPs follows the Treebank annotators. Nevertheless, as Ramshaw and Marcus do not depict the details of their conversion algorithm, results may not be the same in difficult cases, e.g. including NAC and NX.

An ADJP constituent inside an NP constituent are part of the NP chunk: (NP The (ADJP most volatile) form)→ [NP the most volatile form]

3.2 VP

In the Treebank, verb phrases are highly fixed; see e.g. the following sentence which includes four VP constituents. Following Ramshaw and Marcus' V-type chunks, this sentence will only include one VP chunk:

((S (NP-SBJ-3 Mr. Icahn) (VP may not (VP want (S (NP-SBJ) (VP to (VP sell ...))))) .)) →[NP Mr. Icahn] [VP may not want to sell] ...

It is still plausible however to have one VP chunk directly follow another: [NP The impression] [NP I] [VP have got] [VP is] [NP they] [VP 'd love to do] [PRT away] [pp with] [NP it]. In this case the two VP constituents did not overlap in the Treebank.

Adverbs/adverbial phrases are part of the VP chunk (as long as they are in front of the main verb):

(VP could (ADVP very well) (VP show ...)) →[VP could very well show] ...

Contrary to Ramshaw and Marcus (1995), predicative adjectives of the verb do not become part of the VP chunk, e.g. in "[NP they] [VP are][ADJP unhappy]".

In inverted sentences, the auxiliary verb does not become part of any verb phrase in the Treebank. Therefore it does not become part of any VP chunk:

((S (SINV (CONJP Not only) does (NP-SBJ-1 your product) (VP have (S (NP-SBJ) (VP to (VP be (ADJPPRD excellent)))))), but ...

→[CONJP Not only] does [NP your product] [VP have to be] [ADJP excellent], but ...

3.3 ADVP and ADJP

ADVP chunks mostly yally with ADVP constituents in the Treebank. However, ADVPs inside ADJPs or inside VPs if in front of the main verb are absorbed into the ADJP respectively VP chunk. On the other hand, ADVPs that include an NP make two chunks:

(ADVP-TMP (NP a year) earlier) → [NP a year] [ADVP earlier]

ADJPs inside NPs are absorbed into the NP. And parallel to ADVPs, ADJPs that include an NP make two chunks:

(ADJP-PRD (NP 68 years) old) →[NP 68 years] [ADJP old]

It would be interesting to see how changing these decisions affect the chunking task.

3.4 PP and SBAR

Most PP chunks just are composed of one word (the preposition) with the part-of-speech tag IN. This does not demonstrate, though, that finding PP chunks is wholly unimportant. INs can also make up an SBAR chunk (see below) and some PP chunks include more than one word. This is the case with fixed multi-word prepositions such as *such as, in spite of, due to*, with prepositions preceded by a modifier: *well above, just after, even in, particularly among* or with coordinated prepositions: *inside and outside*. We think that PPs perform sufficiently differently from NPs in a sentence for not wanting to chunk them into one class (as Ramshaw and Marcus did in their N-type chunks), and that on the other hand tagging all NP chunks inside a PP as I-PP would only puzzle the chunker. As a result we chose not to deal with the recognition of true PPs (prep.+NP) during this first chunking step.

SBAR Chunks mostly are composed of one word (the complementizer) with the part-of-speech tag IN, but like multi-word prepositions, there are also multi-word complementizers: *even though, such that, just as, even if, as though, only if.*

3.5 CONJP, PRT, INTJ, LST, UCP

Conjunctions can be composed of more than one word as well: *as well as, instead of, rather than, not only, but also*. One-word conjunctions (like and, or) are not annotated as CONJP in the Treebank, and are therefore no CONJP chunks in our data.

The Treebank employs the PRT constituent to annotate verb particles, and our PRT chunk does the same. The sole multi-word particle is *on and off*. This chunk type should be easy to identify as it should coincide with the part-of- speech tag RP, but through tagging errors it is sometimes also assigned IN (preposition) or RB (adverb).

INTJ is an interjection phrase/chunk like *alas, oh, hello, alas, Goodness*! It is quite rare.

The list marker LST is even rarer. Examples are *1, 2, 3, first, second, a, b, c*. It might be composed of two words: the number and the period.

The UCP chunk is suggesting the UCP (unlike coordinated phrase) constituent in the Treebank. Arguably, the conjunction is the head of the UCP, so most UCP chunks consist of conjunctions like *and* and *or*. UCPs are the rarest chunks and are possibly not very useful for other NLP tasks.

3.6 Tokens Outside

Tokens outside any chunk are mainly punctuation signs and the conjunctions in common coordinated phrases. The word not may also be outside of any chunk. This occurs in two cases: Either not is not inside the VP constituent in the Treebank annotation e.g. in

... (VP have (VP told (NP-1 clients) (S (NP-SBJ) not (VP to (VP ship (NP anything))))))

or *not* is not followed by another verb (because the main verb is a form of *to be).* As the right chunk dividing line is explained by the chunk's head, i.e. the main verb in this case, *not* is then in fact a postmodifier and as such not contained in the chunk: ".. [SBAR that] [NP there] [VP were] n't [NP any major problems]"."

4. Mental Lexicon

Knowledge of a word may contain its individual pronunciation, spelling, meaning, and grammatical usage and collocation patterns. Speakers of a language usually have some knowledge of this type for much more items than they actively employ. Cognitive psychology claims that vocabulary knowledge is retained in a 'mental lexicon'.

Aitchison (2003: ix) is typical in emphasizing that the large vocabulary size and efficiency of retrieval "suggest that these words are carefully organized, not just stacked in random heaps, in the nature of a human word-store, or mental lexicon".

Retrieval of information of vocabulary from the mental lexicon for better and smoother understanding is required not only by first language learners but also in a second language (Kadota, 2003). The long-term stored mental lexicon, much like a paper dictionary, contains language information such as spelling, phonemes, meanings and syntax. Then vocabulary knowledge is extracted from the mental lexicon and final understanding of contents can be achieved (Kadota). For example, if some information is given, but is spoken in a completely unfamiliar language, the language itself can not be recognized, not to mention the contents of the information understood, because there is no information about the language in your mental lexicon (Abe, 1995).

5. Working Memory and Short-Term Memory

5.1 Working Memory

Baddeley (1984:34) explains working memory as "a system for the temporary holding and manipulating of information during the performance of a range of cognitive tasks such as comprehension, learning and reasoning". Researchers in cognitive psychology claim that working memory is employed to hold incoming information temporarily in daily cognitive acts such as perception, conversation, reading comprehension, mental calculation, judgment, reasoning and thinking (Osaka, 2000). Working memory, the so-called mental memorandum" or "blackboard" (Oishi, 2006:87) is thought to be composed of two slave systems. One is the phono]ogical loop and the other is the visuo-spatial sketchpad. A system called the central executive dominates the two slave systems. Phonological loop especially concerns the understanding of sounds. There is a phonological short-term store on the phonological loop. Miyake (1995) notes that after listening to incoming information, the phonetic memory is said to be crossed out in 1.5-2.0 seconds from the phonological store unless there is attention to and rehearsaI of incoming phonetic sounds. For instance, when completely unfamiliar information such as a new phone number or a new address is given as auditory input, the phonetic memory will vanish in a moment if there is no vocalizing or subvocalizing rehearsal. The process of mentally vocalizing forms is called "subvocalization". In case of visual input (looking at words), the words are vocalized and rehearsed on the phonological loop, with the effect of retaining incoming information. As long as the incoming information with phonetic sounds is rehearsed on the phonological loop, phonetic memory can be held indefinitely (Miyake, 1995). The central executive in working memory plays a crucial role in dominating the allocation of cognitive resources for various cognitive procedures. The central executive has the function of assigning the two slave systems to their respective tasks and it also plays a vital role in extracting information from long-term memory.

5.2 The Definition and Nature of STM

A very important event in the history of cognitive psychology was the development of a theory of STM (short-term memory) in the 1960s. Jack C. Richards, John Platt and Heidi Platt (2000: 283) define STM as: “Short-term memory refers to that part of the memory where information which is received is stored for short periods of time while it is being analyzed and interpreted. Once the message or information in an utterance is understood the data may become part of permanent memory (or long-term memory). The utterance itself is now no longer needed and may fade from short-term memory.”

STM is transient memory. It has a limited capacity—it can only hold about seven or so unrelated chunks (a chunk is a meaningfully coded unit) once. LTM (long-term memory), on the contrary, is the place where more permanent information is stored. It is unlimited in capacity and retains information over a much longer interval, but it often takes a fair amount of effort to get information into it. LTM serves as a data base into which information is inserted through STM, and from which information is extracted to be employed in STM.

STM plays the role of a gateway into LTM. In Call's (1985) view, memory is composed of three parts: sensory store, short-term memory and long-term memory. Information comes in from the environment through a series of sensory memory systems (iconic and auditory memory) from which it is lost unless attended. The transitory sensory store retains information for a few hundred milliseconds; its characteristics are for the storage of visual information. Then the information goes into an intermediate STM where it has to be rehearsed before it can go into a relatively permanent LTM. Information is lost within 20-30 seconds if it is not rehearsed in STM. If the item leaves STM before a permanent LTM representation is developed, it will be lost forever. One can't hold information in STM forever since new information would always be coming in and would push out old information from the limited STM. STM can combine information from both the environment and LTM whenever a person tries to learn new information, make decisions, or solve problems. Once in STM, an item can be held there by rehearsal. As an item is rehearsed, information about it is transferred to LTM. As soon as rehearsal of an item is brought to an end, the item soon will be displaced by a new incoming item and thus lost from STM.

6. Applying Chunking to Vocabulary Memorization

According to several experiments conducted by Miller (1956), the capacity of short-term memory is restricted to about seven units, plus or minus two. This definition of each unit is equivalent to Miller's (1956) "chunking" account of memory capacity. Miller explains the capacity of short-term memory as a rather fixed or limited storage place, which makes room for seven (plus or minus two) "chunks" of information. A "chunk" alludes to a group of data arranged together based on the previously learned patterns which change the sound stream of the language into meaningful units. For example, the words, "the graceful teacher is talking loudly in the classroom" will be chunked into a few units like this: (the graceful teacher) (is talking loudly) (in the classroom).

That is to say, each unit can usually be explained syntactically as a word, a phrase, or a clause. The syntactic units, then, pass through the process of semantic interpretation before they are removed from short-term memory in order to accommodate new incoming data. Although the semantic information may go into long-term memory, the exact wordings in which the original messages were expressed are often lost or forgotten once the meaning has been retrieved (Jarvella, 1971). Therefore, short-term memory plays a crucial role not only in providing a space for language processing but also making a decision over whether the extracted meaning can go into long-term memory storage for further retention (Call, 1985).
There is less demand on cognitive capacity，because the lexical chunks are "ready—to—go"，and demand little or no additional processing. Although it is reasonable to a certain extent that grammatical knowledge allows for the creative recombination of lexis in novel and imaginative ways，it cannot function in its role until learners have accumulated a sufficiently large mental lexicon (Lewis 2000). On the other hand, Nattinger and DeCarrico (1992) claim that fluency is based exactly on a lot of "chunks" or "lexical phrases". They propose that prefabricated speech can be efficiently retrieved from speakers' mental lexicon and thus allows them to concentrate their attention on the larger structure of the discourse rather than keep it narrowly concentrated on individual words as they are produced. Knowing a word in isolation only may not aid to produce natural flow of speech. On the contrary, "chunking" could explore a word knowledge to the full and in Sinclair's (1991) words, it recognizes a words colligation (i.e. grammatical patterning), collocation (i.e. lexical patterning), semantic preference and (i.e. semantic association) semantic prosody (i.e. semantic connotation). I determinedly persist that employing all the opportunities to teach chunks rather than isolated words is an effective way that will be working well in college classes. In fact, it was early in 1990s that Lewis (1993), because of the accessibility of large a mounts of native speakers' natural language data, proposed that to chunk language was the most economical and feasible way of learning vocabulary.

7. Cultivation Measures

7.1 Employing Chunking Strategy and Increasing Short-Term Memory Capacity

Chunking theory was first suggested by G. A. Miller in 1956. Since then, it has been widely employed in second language acquisition to a large extent. Gui Shichun (1991) claimed that chunking is an efficient method to overcome limitations of short-term memory. Chunking strategy is employed to enlarge information load of individual units and to increase the capacity of the short-term memory. As for this, there are chiefly two reasons: 1. Chunking can make each information unit rich and lessen the burden of short-term memory by diminishing amounts of information units. 2. When all the information is grouped into several large chunks, and information in each chunk is closely related to each other, representations in each large corresponding chunk of long-term memory will be excited. It is widely admitted that short-term memory can retain about 7+2 chunks. And information capacity of each chunk counts on the size of each chunk. Meanwhile, there must be three preconditions for chunking. 1. Enough time, 2. The nature of information in each chunk, 3. Corresponding chunks in the long-term memory before chunking (Liu, 2005).

7.2 Raising Students' Chunking Awareness and Improving their Ability to Chunk Language

As Lewis (1993) claims, language is composed of multi－word prefabricated chunks and an important element of language teaching is to raise students' awareness of chunks, and to develop their ability to chunk language successfully. We can imply from his argument that why most college students with a large vocabulary size still have problems with fluency. Their weak collocation competence may very well result in the possible reason. As a result, it is crucial to make students aware of chunks, giving the opportunities to recognize, organize and record chunks. Recognizing chunks is not always easy, and at least in the beginning, students demand a lot of help and guidance. Teachers then should play the important role in accelerating "chunk noticing" to occur. Noticing characteristics of the input, particularly the nature of the component chunks of the text has a facilitative value. Noticing language as chunks helps storage as chunks and therefore helps acquisition (Lewis 1997). As for language chunking two channels have been generally regarded as appropriate —chunking from known parts and chunking from the memorization of unanalyzed chunks (Nation 2001).

Educating students to recognize chunks is just the beginning of helping them change input to intake. The next step is needed for us teachers to activate these items in the classroom and offer students with chances to use these items.

2012 Qingdao University of Science and Technology's Humanity and Society Science Research Project (Grant No.11XC18).

References

[1] M.E. Call, Auditory short-term memory, listening comprehension and the input hypothesis, *TESOL Quarterly*, 19(4) (1985), 765-781.

[2] S. Kadota, Mentarurekishikon to wa namka (What is mental lexicon?), In Syohei Kadota (ed.), *Eigo no Mentarurekishikon (The English Mental Lexicon)*, (2003), 2-11, Tokyo: Shotosha.

[3] L. Ramshaw and M. Marcus, Text chunking using transformation-based learning, *In Proceedings of the Third ACL Workshop on Very Large Corpora*, (1995), 82-94.

[4] M. Lewis, *The Lexical Approach: The State of ELT and a Way Forward*, (1993), Hove, England: Language Teaching Publications.

[5] M. Lewis, *Implementing the Lexical Approach: Putting Theory into Practice*, (1997), Hove, England: Language Teaching Publications.

[6] M. Lewis, Learning in the lexical approach, In M. Lewis (ed.), *Teaching Collocation: Further Developments in the Lexical Approach*, (2000), Hove, England: Language Teaching Publications.

[7] K.P. Liu, The role of memory in second language cognitive process, *Foreign Language Teaching*, 4(2005), 54-57.

[8] G.A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, *The Psychological Review*, 63(2) (1956), 81-97.

[9] A. Miyake, Tankikioku to sadokioku (Short-term memory and working memory), *Ninchi Shinrigaku (2: Kioku) (Cognitive Psychology (Vol. 2: Memory)*, (1995), 71-99, Tokyo: Tokyo University Press.

[10] I.S.P. Nation, *Learning Vocabulary in another Language*, (2001), Cambridge: Cambridge University Press.

[11] H. Oishi, Gengo to wakingu memori (Language and working memory), *Nokagaku Kara no Dainigengo Shutokuron (Second Language Acquisition from Neuroscience)*, (2006), 77-90, Kyoto: Showado.

[12] N. Osaka, Wakingu memori to daini gengo shori (Working memory and second language acquisition), In N. Osaka (ed.), *No to Wakingu Memori (Brain and Working Memory)*, (2002), 243-254, Kyoto: Kyoto University Press.

[13] J.C. Richards, J. Platt and H. Platt, *Longman Dictionary of Language Teaching & Applied Linguistics*, (2000), Foreign Language Teaching and Research Press.

[14] J.M. Sinclair, *Corpus, Concordance, Collocation*, (1991), Oxford: Oxford University Press.

[15] F. Xu, Short-term memory, foreign language listening comprehension and the input hypothesis, *Foreign Language Teaching Abroad*, 1(2005), 28-36.