# Towards a Unified Theory of Chunking

Mahendra Indukuri

University College Dublin, Ireland
mahendra_indukuri@ucdconnect.ie

**Abstract.** "Psychologists seem to know a chunk when they see one" points out Herbert S. Terence in a chapter about chunking in avian visual cognition published in 2001(Terrace, 2001). The phenomena of chunking are extensively discussed in the literature in relation to human and animal learning mechanisms and behaviour. In the context of humans, which is the scope of this paper, the term chunking is used primarily in two different contexts in literature; memory chunking and motor action chunking. However, there is a lack of explanation in the literature about the term chunking itself and especially in relation to how these different types of chunking relate to each other. In this paper, different types of chunking mechanisms in humans will be discussed and a new framework of chunking that explains these different chunking mechanisms as a singular phenomenon will be developed. This paper will build on the recent work which points out this lack of proper definition and also the possibility of a unified framework that could define chunking in humans as a singular phenomenon (Gobet, 2016). Neurophysiological data relating to the observed function of basal ganglia in different types of chunking and potential similarities in underlying neurophysiology of these different types of chunking will be discussed as evidence in supporting a unified framework that poses chunking as a singular phenomenon (Graybiel, 1998).

**Keywords:** motor chunking, memory chunking, chunking in humans.

## 1 Introduction to Chunking

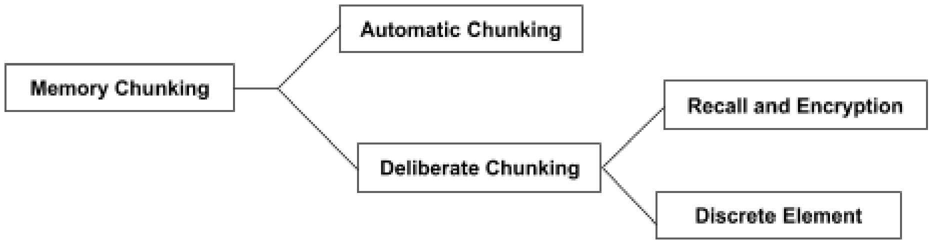
To understand more about what chunking means, let us look at some examples. Remembering a phone number can be given as an example of chunking (Li, 2013) where chunks of sequences form a larger sequence which would have otherwise been not possible so is typing on a QWERTY keyboard (Yamaguchi, 2014). At its core chunking can be defined as the ability to overcome objections to linear models of serially organized behaviour (Terrace, 2001) and this definition includes both memory as well as motor action chunking. By remembering chunks one can remember a long string of digits and by using hierarchical models of production for typing, people can type faster by typing more characters as part of a single chunk. The contextual nature of the effects of chunking can be best understood from the example of typing where people who are good at typing can type faster but are not necessarily better at recognising typos in a text which tells us their knowledge of the keyboard and typing is a hierarchical production system that activates only during typing. The chunks formed are very local to the typist's fingers and the typist need not necessarily be better at recognising incorrect spellings while glancing.

Just to understand the extent to which chunking is used to define various phenomena, we can look at the case of chunking in robotics where chunks of schemas(Schemas used for robot learning are sequences of actions that accomplish a goal-directed behaviour, or a task) formed to solve subproblems will be remembered as such in the robot's memory (Tang, 2008). This is a case where the idea of chunking is applied in a non-living entity but that doesn't make chunking a robotic or a computational process. But drawing from this idea, we can study motor chunking in humans as a similar mechanism and applying some of the frameworks currently used in robotics can be of use as well. We can see closely related ideas about chunking, problem-solving using chunking and expertise discussed in animals and even robots. This phenomena of chunking is thus an emergent pattern in learning behavioural tasks in order to overcome limitations in problem-solving. Whether the problem is regarding memory or motor action. But we will see later in the paper that even some memory chunking processes in humans can be restated as motor action-based learning which makes a better case for a unified theory.

The examples ranging from the purely cognitive task of remembering numbers to a motor production task of typing and many tasks that have a motor as well as a cognitive component can be studied in order to develop a unified theory of chunking. All though chunking phenomena have been studied as distinct aspects of human learning and behaviour similar terminologies have been used to define chunks and chunking and this tells us the similarities are to be explored and unification can be useful. The argument for unification is not that these phenomena cannot be studied independently, even though these phenomena can be studied independent to one another, the commonalities that can be clearly identified are important in developing a better understanding of the nature of these phenomena and this will help in avoiding some of the confusion with terminology to which researchers are not immune. In addition, developing an overarching ontology can help researchers in the domain borrow ideas and from one another and help towards unification in an interdisciplinary discipline like cognitive science.

## 2 Classification(s) of chunking

Due to the vast variety of phenomena included, chunking mechanisms can be classified using multiple methods of classifications that are not mutually exclusive. One of the basic classifications that are crucial is the distinction between chunking in memory and chunking in motor action (Terrace, 2001). As we saw in the example with phone numbers, chunking in memory involves remembering chunks of digits together so that long lists can be remembered. For example, chunking helps in storing the list of digits (9010-83-4224), by just using three chunks the recall can be easier. However, this classification of memory chunking is not the end, if we take the example of (IBM-NASA-USA) we can still see the same phenomena of memory chunking but there is a clear difference in how the chunks are formed even though the resulting phenomena of chunking is similar. In the second example, recalling from pre-existing knowledge of the three distinct elements helps chunking. These types of memory chunking can be classified together into "deliberate chunking" which can be distinguished again from automatic memory chunking like first language learning (or early language learning in bilinguals and trilinguals).




In the above classification, both (IBM-NASA-USA) and (9010-83-4224) come under Deliberate Chunking, however, the former is an example of Recall and Encryption based chunking while remembering phone numbers is discrete element chunking. However, the common factor of using three-chunks instead of ten-chunks in remembering can be applied to all cognitive chunking mechanisms i.e., it is more likely someone will forget the whole USA or 4224 or a word/simple phrase in their first language than forgetting parts of USA or 4224 or a word/simple phrase in their first language. In addition to the above examples, strategy games like chess can be described to contain memory chunks where a Recall and Encryption kind of chunking occurs (Gobet, 1998).

Now let us look at chunking in motor action which can be clearly distinguished from memory chunking. A top-down classification of various types of motor chunking would be difficult to imagine. Motor chunking can be involved in dexterity, verbalisation, perception/gaze or in any other voluntary motor capability that can be improved by tackling serially ordered behaviour with chunking. Some examples of motor action chunking are, typing and playing a musical instrument (Piano, Guitar etc.). Even in this type of chunking, there can be classifications based on factors such as involvement of active cognitive problem-solving aspect for which Tower of Hanoi can be an example. Whereas typing requires lesser thinking and video games are probably somewhere in the middle. The classifications here are a bit more tricky and I believe an additional discussion is required.

While solving ToH quickly, some actions happen as if they are automatic just like in driving but the core task is cognitive in the case of ToH. The motor practice may help with speed but cannot be sufficient but the chunking involves a significant motor component as well. Phenomena, where both cognitive and motor control are required at the same time to improve on performance at the given task, can be classified as in a different class from purely motor-action chunking if such phenomena actually exist. In this hybrid chunking, a hierarchical control system that tries to help overcome the linear action includes both memory and motor-action and can be commonly observed in human-computer interface interactions which involve rapid action and learning of chunks. A better understanding of these phenomena and how we define expertise in these phenomena can be important in the domain of neural-interfaces and other interfaces of the future.

Now to look at the commonalities, among all chunking processes is the assumption that all of the processes that are studied under chunking involve bottom-up learning and by extension, we can say to characterise chunking means believing in bottom-up learning. In addition, there is the aspect of the formation of hierarchical structures whether in memory or in motor control.

## 3 Chunking and computational models of cognition

Chunking is used in parts of cognitive models and some models of cognition are based on chunking. In addition to studying these models to understand more about how the concept of chunking is currently used, we can also learn about chunking as a phenomenon in general learning of sequential behaviour and/or sequential thinking. We are going to look at three of the most influential or widely used models that incorporate chunking.

### 3.1 Chunking in ACT-R

In ACT-R (Adaptive Control of Thought—Rational) a chunk is a unit of declarative memory. Chunks can be retrieved into working memory when a production rule is invoked while a learning task is performed. These retrieved chunks can be modified or placed back in the LTM without any change. (Taatgen, 2002)

### 3.2 Chunking in Soar

In Soar, a chunk is a unit of working memory. Chunking is a learning mechanism that acquires rules from goal-based experience. (Laird, 1986).

### 3.3 Chunking in CHREST

In CHREST (Chunk Hierarchy and REtrieval STructures) a chunk is a node that is the initial element of a series of sequential elements (mostly refers to declarative). A chunk is formed in the long term memory and short term memory feeds into long term memory where chunks are formed and can be retrieved (Gobet, 2010).

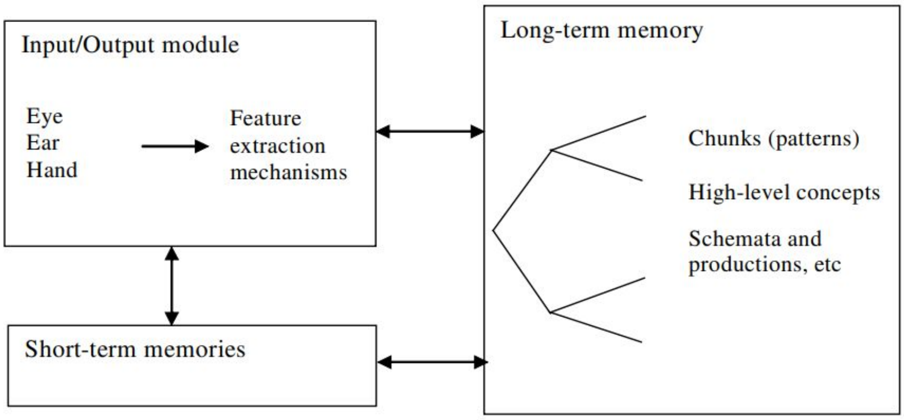



Image source (Gobet, 2010)

Although these models give us varied explanations of chunking as a phenomenon, the application of these models is mostly limited to cognitive chunking i.e., memory chunking. Hence, motor production tasks are not studied using these computational models. Regardless of this limitation, we can clearly understand important facts from further studying these models and their conceptualisation of chunking. One feature of all these explanations of chunking is that they follow a computational approach towards the understanding of chunking and not a biological one, as mentioned earlier a biological theory of chunking based in evolution is possible with a redefinition of some of the features of memory chunking as motor chunking mechanisms like verbalisation and redirection of attention makes it possible for a theory of chunking based on the biology of the organism.

## 4 Towards a Unified Theory of Chunking

In a unified theory of chunking, specific types of memory chunking phenomena are defined as special instances of motor action chunking. The motor action can be the act of verbalising a number while committing it to memory or simply talking to oneself that can be considered a pseudo verbalisation or a pseudo motor action that helps one store a long string of numbers in their memory. The importance of chunking as a central problem in linguistic memory has been studied(Cánovas, 2020). Defining verbal chunking as a motor mechanism is not particularly inventive, we know that being able to repeat a phrase to oneself a few times will commit it to memory and most often the recall is made in a similar tone using mostly the same muscles used while trying to remember it in the first place. Even in non-verbal memory, a case for chunking has been made. For example, chunking observed in mind maps involves a pseudo motor component. Directing perceptual attention and noticing a particular pattern is very much a motor task.

As we have seen, when referring to chunking, researchers working within one of memory chunking or motor action chunking domains ignore the implications of their research and theories on the other aspect of chunking as if they are quite unrelated phenomena while using similar ideas in defining chunks and chunking. Chunks define the newly formed elements whether they are a set of elements that are being committed to memory as a "chunk" or a set of motor actions that are being entangled with each other to form a "chunk". This process, while helping humans be a bit more efficient in remembering and helping speed up some motor actions, also places some very similar constraints in both cases. Just to illustrate this similarity in the new constraints placed, let us consider examples of typing a phone number vs enunciating it. And let us say all the subjects are asked to practice both these tasks before we test them, so that chunk formation happens. In the example of typing, let us say typing a phone number on a NUM pad which is a case of motor chunking, a set of few digits will be typed in one go another set in the next go and so on, i.e., there will be a longer pause between chunks than within chunks. Also, if one is asked to switch positions of digits within a chunk, it will be harder than switching the positions of two whole chunks. It is important to remember that this is when people are asked to type at their regular pace. Similarly, in verbalisation, people generally pause between two chunks. If asked to switch positions of digits within a chunk, they might struggle while switching positions of whole chunks is an easier task. In both these examples, a set of muscles are acting to produce either a verbal or a haptic way of interacting with the world and a set of verbalisations which are motor actions and a set of taps on a keyboard are being chunked together to overcome certain constraints. By restating these two examples as the same core phenomenon, we can approach the topic of chunking differently.

In this unified theory of chunking, a strong case can be made for redefining two specific types of memory chunking namely automatic & discrete element chunking explained earlier as motor chunking mechanisms where the motor action chunking that is happening in the act of saying the numbers for a phone number out loud or in some cases to oneself. If this were true we can use a common framework to refer to these two types of chunking that are seen to be fundamentally different from one another. This hypothesis can be tested in various ways including looking for different rates at which different people are able to remember discrete numbers in tests and comparing the results between people who remember numbers as words and other who remember a pictorial form / some abstract concept rather than a sound.

In addition to the improvements and unification in the fragmented topic, this new unified theory of chunking has implications on language itself, for example by considering chunking in verbal memory as a special case of motor learning brings sign language much closer to any of the verbal languages at least in any context where chunking is the core phenomenon. in understanding memory and recall from memory through chunks as more than just a purely mental process. Even in cases where there is no explicit verbalisation, chunking in memory can be considered as a pseudo motor phenomenon of sorts, i.e., people generally are talking to themselves if not explicitly verbalising the sounds. This helps one commit chunks to memory. This pseudo verbalisation in order to learn is a learnt behaviour, we know that verbal repetition is how children learn a language and construct linguistic memory at an early age. Oral recitals of poems etc. are common in almost all the world languages and it can be observed that children even develop specific head motions or motor patterns etc. in order to learn a linguistic memory chunk which with practice can be done easily. To define these performative motor actions that produce a sound a memory chunk takes us farther away from reality and closer to a computational understanding of language and linguistic memory while we can clearly observe constrained motor action and remembering constrained motor action is very much what verbal memory starts at.

## 5 Similarities in Neurophysiology of Chunking

Finding a neurophysiological activity that can be attributed to most or all kinds of chunking can potentially mean that the step towards developing a unifying ontology is correct. In regards to chunking, research points towards the basal ganglia although most research is in relation to motor sequence learning and chunking within motor sequence learning. Damage to basal ganglia and its correlation to impairments to motor sequence learning and chunking has been studied (Boyd, 2004 & Boyd 2009). Interestingly basal ganglia function is closely related to obsessive-compulsive disorder in addition to impairments in motor chunking processes (Rapoport, 1990; Giedd 1996; Carmin 2002 & Modell 1989). The characteristic lack of awareness of the algorithm that is being learned and the slow rate of learning these motor chunking sequences is well understood and the role of basal ganglia can be observed in multiple phenomena where S-R learning leads to chunking or learning of chunks (Graybiel, 1998). In an attempt to generalise cognitive and motor chunking specifically with regards to the role of basal ganglia’s function in learning production sequences, it has also been pointed out that the learning and memory functions of the basal ganglia can be seen as core features of the basal ganglia’s influence on the motor and cognitive pattern generators.

In addition, basal ganglia activity is related to auditory language perception as we have seen in the example of first language acquisition as a component of automatic memory-chunking and the fact that basal ganglia might be specifically involved in syntax processing where hierarchical sequencing might occur tells us there is a more central role to basal ganglia when it comes to chunking (Kotz, 2009). Research also suggests a more general role of basal ganglia in all learning that is incrementally acquired (Packard, 2002). Neurophysiological evidence pointing towards similarity in the physiology of how both motor and cognitive chunking learning happens in the basal ganglia helps the case for a unified theory.

## 6 Chunking and Expertise

Differentiating chunking from the general idea of conditional learning is important to set the boundaries of the domain. For instance, a pigeon learning a task A or B or C can happen without chunking playing a role but when a relationship develops between A, B, C, where activation of one would lead to the other means chunking is involved as there is a single pointer activation involved in initiating three different tasks (Terrace, 2001). A better definition of tasks with examples can better demonstrate this difference.

Coming to expertise, While chunking is a natural phenomenon, deciding expertise can involve cultural connotations. Chunking is often studied in expert behaviour including playing chess (Gobet, 1998 ), solving logical problems (Egan, 1979 & Lane, 2001) etc. Although the evolution of chunking phenomena in attempting to solve a problem need not necessarily be referred to as expertise, developing expertise almost always involves chunking. Cultural connotation attached to the word expertise (i.e., we don't talk about expert handwriting professionals even though handwriting involves a great deal of chunking) downplay the role for chunking mechanisms in many forms of expertise. To state simple, expertise and chunking are very closely linked. This becomes much more obvious when we look at the converse, it is hard to think of expertise without some form of chunking phenomena involved. Talking, painting, walking, writing, playing, and solving mathematical problems all involve chunking. The implications of this can help education and developing an academic curriculum etc. directly (Gobet, 2005). Just to be clear, we need to distinguish expertise from experts in this regard, one is an expert because of their expertise in one or more activities. I understand this claim is bold and I wish to develop it further to better present my case in this regard.

## 7 Conclusion

This fact about expertise and the fact that animals including birds show chunking mechanisms as part of their behavioural tool kit in trying to learn and behave tells us that chunking, including memory chunking, can be seen as an evolutionary property in higher-order living organisms. This puts chunking in a different light than the traditional definitions given to define memory chunking which take a computational approach rather than a biological one. How we approach the topic can also be based on various schools of thought within cognitive science and considering chunking, including memory chunking, as a motor learning phenomenon at its core can be better developed in an embodied approach to the phenomenon of chunking. Although a specific framework which can be used to replace the wonderfully useful memory chunking theories discussed in the computational models section is not developed in this paper, developing one based on motor action is possible and can help us take a unified approach.

## References

1. Boyd, L. A., Edwards, J. D., Siengsukon, C. S., Vidoni, E. D., Wessel, B. D., & Linsdell, M. A. (2009). Motor sequence chunking is impaired by basal ganglia stroke. Neurobiology of learning and memory, 92(1), 35-44.
2. Boyd, L. A., & Winstein, C. J. (2004). Providing explicit information disrupts implicit motor learning after basal ganglia stroke. Learning & memory, 11(4), 388-396.
3. Cánovas, C. P. (2020). Learning formulaic creativity: Chunking in verbal art and speech. Cognitive Semiotics, 1(ahead-of-print).
4. Carmin, C. N., Wiegartz, P. S., Yunus, U., & Gillock, K. L. (2002). Treatment of late‐onset OCD following basal ganglia infarct. Depression and anxiety, 15(2), 87-90.
5. Giedd, J. N., Rapoport, J. L., Leonard, H. L., Richter, D., & Swedo, S. E. (1996). Case study: acute basal ganglia enlargement and obsessive-compulsive symptoms in an adolescent boy. Journal of the American Academy of Child & Adolescent Psychiatry, 35(7), 913-915.
6. Gobet, F., & Lane, P. (2010). The CHREST architecture of cognition: The role of perception in general intelligence.
7. Gobet, F., Lloyd-Kelly, M., & Lane, P. C. (2016). What's in a name? The multiple meanings of "Chunk" and "Chunking". Frontiers in psychology, 7, 102.
8. Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. Memory, 6(3), 225-255.
9. Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. Neurobiology of learning and memory, 70(1-2), 119-136
10. Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. Cortex, 45(8), 982-990.
11. Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. Machine learning, 1(1), 11-46.
12. Li, G., Ning, N., Ramanathan, K., He, W., Pan, L. I., & Shi, L. (2013). Behind the magical numbers: hierarchical chunking and the human working memory capacity. International journal of neural systems, 23(04), 1350019.
13. Modell, J. G., Mountz, J. M., Curtis, G. C., & Greden, J. F. (1989). Neurophysiologic dysfunction in basal ganglia/limbic striatal and thalamocortical circuits as a pathogenetic mechanism of obsessive-compulsive disorder. The Journal of neuropsychiatry and clinical neurosciences.
14. Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. Annual review of neuroscience, 25(1), 563-593.
15. Rapoport, J. L. (1990). Obsessive compulsive disorder and basal ganglia dysfunction. Psychological medicine, 20(3), 465-469.
16. Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say "broke"? A model of learning the past tense without feedback. Cognition, 86(2), 123-155.
17. Tang, Y., & Parker,L. E. (2008, May). Towards schema-based, constructivist robot learning: Validating an evolutionary search algorithm for schema chunking. In 2008 IEEE International Conference on Robotics and Automation (pp. 2837-2844). IEEE.

18. Terrace, H. (2001). Chunking and serially organized behavior in pigeons, monkeys and humans. In R. G. Cook (Ed.), Avian visual cognition [On-line]

19. Yamaguchi, M., & Logan, G. D. (2014). Pushing typists back on the learning curve: Revealing chunking in skilled typewriting. Journal of Experimental Psychology: Human Perception and Performance, 40(2), 592.