In press, in D.C. Noelle & R. Dale (Eds.), Proceedings of the 37th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Individual Differences in Chunking Ability Predict On-line Sentence Processing

Stewart M. McCauley (smm424@cornell.edu)
Morten H. Christiansen (christiansen@cornell.edu)
Department of Psychology, Cornell University, Ithaca, NY 14853 USA

Abstract

There are considerable differences in language processing skill among the normal population. A key question for cognitive science is whether these differences can be ascribed to variations in domain-general cognitive abilities, hypothesized to play a role in language, such as working memory and statistical learning. In this paper, we present experimental evidence pointing to a fundamental memory skill—*chunking*—as an important predictor of cross-individual variation in complex language processing. Specifically, we demonstrate that chunking ability reflects experience with language, as measured by a standard serial recall task involving consonant combinations drawn from naturally occurring text. Our results reveal considerable individual differences in participants' ability to use chunk frequency information to facilitate sequence recall. Strikingly, these differences predict variations across participants in the on-line processing of complex sentences involving relative clauses. Our study thus presents the first evidence tying the fundamental ability for chunking to sentence processing skill, providing empirical support for construction-based approaches to language.

Keywords: Chunking; Sentence Processing; Language Learning; Usage-based Approach; Memory

Introduction

Language processing takes place in the here-and-now. This is uncontroversial, and yet the consequences of this constraint are rarely considered. At a normal rate of speech, English speakers produce between 10 and 15 phonemes—5 to 6 syllables—per second, for an average of 150 words per minute (Studdert-Kennedy, 1986). However, the ability of the auditory system to process discrete sounds is limited to about 10 auditory events per second, beyond which the input blends into a single buzz (Miller & Taylor, 1948). To make matters worse, the auditory trace itself is highly transient, with very little remaining after 100 milliseconds (e.g., Remez et al., 2010). Thus, even at normal rates, speech would seem to stretch the capacity for sensory information processing beyond its breaking point. Further exacerbating the problem, human memory for sequences of auditory events is severely limited to between 4-7 items (e.g., Cowan, 2001; Miller, 1956). Thus, both signal and memory are fleeting: current information will rapidly be obliterated by the onslaught of new, incoming material. We refer to this as the Now-or-Never bottleneck (Christiansen & Chater, in press).

How, then, is the language system able to function within the fundamental constraint imposed by the Now-or-Never bottleneck? We suggest that part of the answer lies in *chunking*: through exposure to language, we learn to rapidly recode incoming speech into "chunks" which are then passed to higher levels of representation (Christiansen & Chater, in press). As a straightforward example of chunking in action, imagine being tasked with recalling the following string of letters, presented individually, one after another: *b h c r l t i a p o a k c e a o p*. After a single exposure to the string, very few individuals would be able to accurately recall sequences consisting of even half of the 17 letters. However, if asked to recall a sequence consisting of exactly the same set of 17 letters as before, but re-arranged to form the string *c a t a p p l e c h a i r b o o k*, most individuals would be able to recall the entire sequence correctly. This striking feat stems from our ability to rapidly chunk the second sequence into the familiar words *cat, apple, chair*, and *book,* which can be retained in memory as just four chunks and broken back down into letters during sequence recall. Crucially, this ability relies on experience: a sequence comprised of low-frequency words is more difficult to chunk, despite being matched for word and sequence length (e.g., *e m u w o a l d i m b u e s i l t*, which can be chunked into the words *emu*, *woald*, *imbue*, and *silt*).

We suggest that language users must perform similar chunking operations on speech and text in order to process and learn from the input, given both the speed at which information is encountered and the fleeting nature of sensory memory. Importantly, this extends beyond low-level processing: in order to communicate in real-time, language users must rely on chunks at multiple levels of representation, ranging from phonemes and syllables to words and even multiword sequences: children and adults appear to store chunks consisting of multiple words and employ them in language comprehension and production (e.g., Arnon & Snider, 2010; Bannard & Matthews, 2008; Janssen & Barber, 2012). While the exact role played by chunking in abstracting beyond concrete linguistic units in language learning differs across the theoretical spectrum, both usage-based (e.g., Tomasello, 2003) and generative (e.g., Culicover & Jackendoff, 2005) accounts have underscored the importance of multiword units in sentence processing and grammatical development.

Although chunking has been accepted as a key component of learning and memory in mainstream psychology for over half a century (e.g., Miller, 1956) and has been applied to specific aspects of language acquisition (e.g., word segmentation; Perruchet & Vinter, 1998), very little is known about the ways in which chunking ability shapes the development of more complex linguistic skills, such as sentence processing. Moreover, work on individual differences in sentence processing in adults has not yet isolated specific learning mechanisms, such as chunking, focusing instead on more general constructs such as working memory or statistical learning (e.g., King & Just, 1991; Misyak, Christiansen, & Tomblin, 2010).

The present study seeks to address the question of whether individual differences in chunking ability—as assessed by a standard memory task—may affect complex sentence processing abilities. Here, we specifically isolate chunking as a mechanism for learning and memory by employing a novel twist on a classic psychological paradigm: the serial recall task. The serial recall task was selected due to its long history of use in studies of chunking, dating back to the some of the earliest relevant work (e.g., Miller, 1956) as well as its being a central tool in an extensive study of an individual subject's chunking abilities (e.g., Ericsson, Chase, & Faloon, 1980).

We show that chunking ability, as assessed by our serial recall task, predicts self-paced reading time data for two complex sentence types: those featuring subject-relative (SR) clauses and those featuring object-relative (OR) clauses. SR and OR sentences were chosen in part because they have been heavily used in the individual differences literature, but also because multiword chunk frequency has previously been shown to be a factor in their processing (Reali & Christiansen, 2007).

Experiment 1: Measuring Individual Differences in Chunking Ability

In the first experiment, we seek to gain a measure of individual participants' chunking abilities. Rather than using a specifically linguistic task, we sought to draw upon previously learned chunks using a non-linguistic serial recall task. Participants were tasked with recalling strings of letters, much like the above examples; letters were chosen as stimuli in part because reading is a heavily practiced skill among our participant population. However, the stimuli did not feature vowels, in order to prevent them from resembling words or syllables. Instead, the stimuli consisted of strings of sub-lexical chunks of consonants drawn from a large corpus. Because readers encounter such sequences during normal reading, we would expect them to be grouped together as chunks through repeated exposure (much like chunked groups of phonemes of the sort necessary to overcome the Now-or-Never bottleneck during speech processing, as described in the introduction to this paper).

In much the same way that natural language requires the use of linguistic chunks in novel contexts, this task requires that participants be able to generalize existing knowledge—sub-lexical chunks of letter consonants previously encountered only in the context of words during reading—to new, non-linguistic contexts. Importantly, in order to recall more than a few letters (as few as 4 in some accounts; e.g., Cowan, 2001), it is hypothesized that participants must chunk the input string (in this case, we expect them to draw upon pre-existing knowledge of chunks corresponding to the *n*-grams in experimental sequences).

Furthermore, the inclusion of matched control strings—consisting of the same letters as corresponding experimental items, but randomized to reduce *n*-gram frequency—affords a baseline performance measure. Comparing recall on experimental and control trials (see Exp. 2) should thus yield a measure of chunking ability that reflects reading experience while controlling for factors such as working memory, attention, and motivation.

Method

Participants 70 native English speakers from the Cornell undergraduate population (41 females; age: M=19.6, SD=1.2) participated for course credit.

Materials Experimental stimuli consisted of strings of visually-presented, evenly-spaced letter consonants. The stimuli were generated using frequency-ranked lists of letter *n*-grams (one for bigrams and one for trigrams) generated using the Corpus of Contemporary American English (COCA; Davies, 2008). Importantly, *n*-grams featuring vowels were excluded from the lists, in order to ensure that stimulus substrings did not resemble words or syllables.

Letter strings consisted of either 8 letters (4 bigrams) or 9 letters (3 trigrams). These sequences were divided into low-, medium-, and high-frequency bins (separately for bigram- and trigram-based strings): the high-frequency bins consisted of 7 sequences generated from the most frequent *n*-grams (28 bigrams for the bigram-based strings, 21 trigrams for the trigram-based strings). The low-frequency bins consisted of equal numbers of the least frequent *n*-grams, while the medium-frequency bins consisted of equal numbers of items drawn from the center of each frequency-ranked list.

The order of the *n*-grams making up each experimental stimulus was randomized. For each string, a control sequence was generated, consisting of the same letters in an order that was automatically pseudo-randomized to achieve the lowest possible bigram and trigram frequencies for the component substrings. All stimuli were generated with the constraint that none featured contiguous identical letters or substrings resembling commonly used acronyms or abbreviations.

An example of a high-frequency string based on trigrams would be *x p l n c r n g l*, with the corresponding control sequence *l g l c n p x n r*, while an example of a low-frequency string based on bigrams would be *v s k f n r s d,* with the corresponding control sequence *s v r f d k s n*.

Procedure Each trial consisted of an exposure phase followed by a memory recall phase. During exposure, participants viewed a full letter string as a static, centered image on a computer monitor for 2500ms. Letter characters were then masked using hash marks for 2000ms, to prevent reliance on a visual afterimage during recall. Then, on a new screen, participants were immediately prompted to type the sequence of letters to the best of their ability. There was no time limit on this recall phase and participants viewed their response in the text field as they typed it. After pressing the ENTER key their response was logged and the next trial began. The experiment took approximately 15 minutes.

Results and Discussion

A standard measure of recall is the number of correctly remembered items. In this study, recall for letters from the target string (irrespective of the sequential order of the response) was sensitive to frequency (High: 78.5%; Medium: 75.8%; Low: 72.9%). According to this measure, subjects were also sensitive to *n*-gram type (Bigram: 78.5%; Trigram: 72.9%) as well as condition (Experimental: 76.9%; Control: 74.6%). Logit-transformed proportions for this simple measure were submitted to a repeated-measures ANOVA with *Frequency* (high vs. medium vs. low), *Type* (bigram vs. trigram), and *Condition* (experimental vs. random control) as factors, with *Subject* as a random factor. This yielded highly significant main effects of Frequency ($F(2,138)=34.57$, $p<0.0001$), Type ($F(1,69)=138.4$, $p<0.0001$), and Condition ($F(1,69)=29.65$, $p<0.0001$).

To gain a more direct measure of chunking, we analyzed the responses for recall of the *n*-grams used to generate the stimuli (for the randomly-ordered control stimuli, items in the identical positions were used to provide a baseline). Participants' recall for chunks was sensitive to frequency (High: 58.5%, Med: 53.2%, Low: 48.8%), *n*-gram type (Bigrams: 59.6%; Trigrams: 47.5%), and condition (Experimental: 55.4%; Control: 51.7%). These proportions were logit-transformed and submitted to a repeated-measures ANOVA with the same factors described above. This yielded highly significant main effects of Frequency ($F(2,138)=71.83$, $p<0.0001$), Type ($F(1,69)=246.1$, $p<0.0001$), and Condition ($F(1,69)=30.52$, $p<0.0001$).

Thus, our findings demonstrate not only that readers are sensitive to sub-lexical chunks—which, consisting solely of letter consonants, do not correspond to syllables—but also that they can generalize to the use of these chunks in a novel context. Participants were sensitive to the frequency of letter bigrams and trigrams even when they appear in novel nonsense strings consisting of 8 or 9 letters. Moreover, participants showed considerable individual differences in their sensitivity to *n*-gram information.

As discussed above, the ability of many subjects to recall more than half of the items in the experimental strings is taken to involve chunking as a specific mechanism, given previously demonstrated memory limitations (e.g., Cowan, 2001). From this perspective, the notion that chunking is involved in the present *n*-gram frequency effect is consistent with over half a century of learning and memory research involving similar paradigms (e.g., Miller, 1956).

The question remains, though, whether differences in chunking ability might relate to language processing skills, as we hypothesized above. Under the view put forth in the introduction, chunking ability as assessed in the present task is predicted to be strongly intertwined with both language abilities and experience. Next, we therefore test whether individual differences in our chunking task predict variation in on-line sentence processing.

Experiment 2: Individual Differences in Language Processing and Chunking

To test whether chunking ability may play a role in language processing, we asked the same participants from Exp. 1 to take part in a self-paced reading task involving sentences featuring SR and OR clauses. We chose these sentence types because they have been the focus of much previous work on individual differences in sentence processing (e.g., King & Just, 1991). Examples of SR and OR sentences are presented in (1) and (2), respectively:

1. *The reporter <u>that attacked the senator</u> admitted the error.*
2. *The reporter <u>that the senator attacked</u> admitted the error.*

In both sentences, *the reporter* is the subject of the main clause (*the reporter ... admitted the error*). The two sentences differ in the role that *the reporter* plays in the underlined relative clause. In SR clauses as in (1), *the reporter* is also the subject of the relative clause (*the reporter attacked the senator*). This contrasts with the OR clause in (2), where *the reporter* is the object of the relative clause (corresponding to *the senator attacked the reporter*).

We suggest that chunking may reduce the computational burden imposed by long-distance dependencies during sentence processing, consistent with previous work showing that word-pair frequency decreases reading times for pronominal relative clauses (Reali & Christiansen, 2007). In line with the finding that ORs, which involve a complex backwards dependency with the head noun, create more processing difficulty than SRs (e.g., Wells, Christiansen, Race, Acheson, & MacDonald, 2009), we hypothesized that the impact of chunking skill may be more visibly observed for OR processing.

Method

Participants The same 70 subjects from Exp. 1 participated directly afterwards in this experiment for course credit.

Materials There were two sentence lists, each consisting of 9 practice items, 20 experimental items, and 30 filler items. The experimental items were taken from a previous study of relative clause processing (Wells et al., 2009) and consisted of 10 SR and 10 OR sentences. A yes/no comprehension question followed each sentence. The condition within experimental sentence sets was counterbalanced across the two lists.

Procedure Materials were presented on a computer monitor using a self-paced, word-by-word moving window display (Just, Carpenter, & Woolley, 1982). At the beginning of each trial, a series of dashes appeared, one corresponding to each nonspace character in the sentence. The first press of a marked button caused the first word to appear, while subsequent button presses caused the next word to appear and the previous word to return once more to dashes. Reaction times were recorded for each button press. Following each sentence, subjects answered a comprehension question using buttons marked "Y" and "N."

The experiment took approximately 10 minutes.

Results and Discussion

Comprehension accuracy on experimental items across participants was reasonably high (M=78.1%, SE=1.7%), with slightly higher accuracy scores for SR sentences (M=80.1%, SE=2.0%) than for OR sentences (M=76.1%, SE=2.2%), though this difference did not reach significance. Only RTs from trials with correct responses were analyzed.

We focused on the same sentence regions used in previous individual differences studies on relative clause processing (e.g., King & Just, 1991; Misyak et al., 2010; Wells et al., 2009); the mean RTs for each of the four regions is shown in Figure 1.

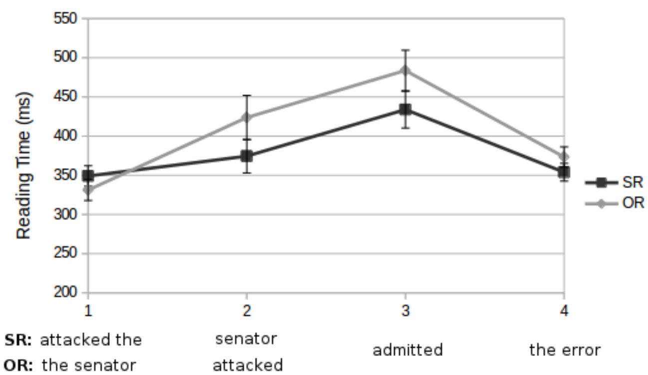

Fig 1: Mean reading times for each region of interest for SR and OR sentences, using sentences 1 and 2 (see above) as examples. Error bars denote standard error of the mean.

The mean RTs for each sentence region were comparable to those observed in previous studies of relative clause processing (e.g., Wells et al., 2009) and followed the same general trajectory, with the greatest RTs observed in Region 3 at the critical main verb.

The frequency of the multiword chunks that make up the relative clause itself has previously been shown to affect processing (Reali & Christiansen, 2007). We therefore initially focused on mean RTs across Regions 1 and 2 (e.g., SR: *attacked the senator* vs. OR: *the senator attacked*), following the hypothesis that chunking may serve to reduce the computational demands involved in processing the embedded clause material.

In order to test the relationship between participants' chunking performance in Exp. 1 and the self-paced reading RTs, we first sought to gain an overall measure of chunking ability. For this purpose, we focused on the difference in performance between experimental and control items in Exp. 1. This offered a means to control for a variety of factors, including working memory, attentional stability, and motivation: independent of chunking ability, each of these factors would be expected to impact experimental and control items equally. Thus, we adopted a measure which depended crucially on sensitivity to the *n*-grams appearing in the stimuli. For this reason, we refer to our measure of chunking ability as the *Chunk Sensitivity* score.

In calculating Chunk Sensitivity, we aimed to incorporate as much of the data from Exp. 1 as possible while still capturing strong individual differences. Because the low-frequency n-grams had the lowest variance in terms of scores (and were taken from the very bottom of the COCA frequency tables and thus the most difficult, with a mean chunk recall rate of under 50%), we focused on the high- and medium-frequency items. A stepwise analysis confirmed that excluding the low-frequency items from the correlations of interest (described below) explained more of the variance in the data. Chunk Sensitivity was then calculated as the difference in the mean proportion of correctly recalled chunks between experimental and control items (the COCA *n*-grams and the corresponding random subsequences in controls).

This measure was a significant predictor of relative clause RTs (the mean RT across Sentence Regions 1 and 2) for ORs (R=.34, β =-788.5, $t(68)=-3.0$, $p<0.01$) as well as SRs (R=.24, β =-465.3, $t(68)=-2.05$, $p<0.05$). Scatterplots depicting these correlations appear in Figure 2.

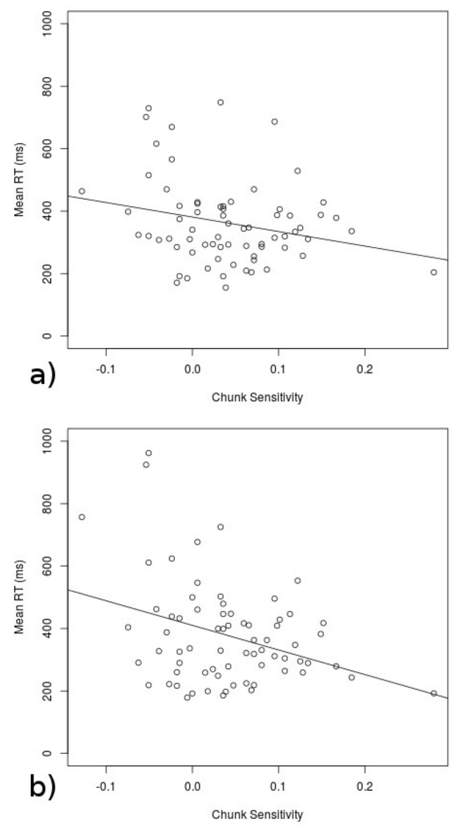

Fig 2: Correlation between the *Chunk Sensitivity* measure (derived from chunk recall scores from Exp. 1) and relative clause reading times for: **a)** SRs and **b)** ORs.

To further explore the role of chunking ability in relative clause processing, we analyzed the whole-clause reading time data using a linear mixed-effects model (LME), with *Clause Type* and *Chunk Sensitivity* as fixed effects and *Subject* as a random effect. This yielded a significant main effect of Chunk Sensitivity (β =-788.55, t=-3.18, $p<0.01$), a significant main effect of Clause Type (β=-29.04, t=-2.44, p<0.05), and a significant Clause Type x Chunk Sensitivity interaction (β=323.28, t=2.2, p<0.05).

Thus, participants with greater chunking ability processed relative clause material faster overall, as evidenced by the main effect of Chunk Sensitivity. As expected, ORs yielded longer readings times overall, as indicated by the main effect of Clause Type. Importantly, participants with greater chunking ability processed the two clause types more similarly, experiencing fewer difficulties with object relatives than subjects with lower chunking abilities, as shown in the interaction between Clause Type and Chunk Sensitivity. To further visualize this interaction, we divided participants into *good chunkers* and *poor chunkers* using a median split across the Chunk Sensitivity measure. Each group consisted of 35 subjects, the mean Region 1-2 RTs for which are depicted in Figure 3.

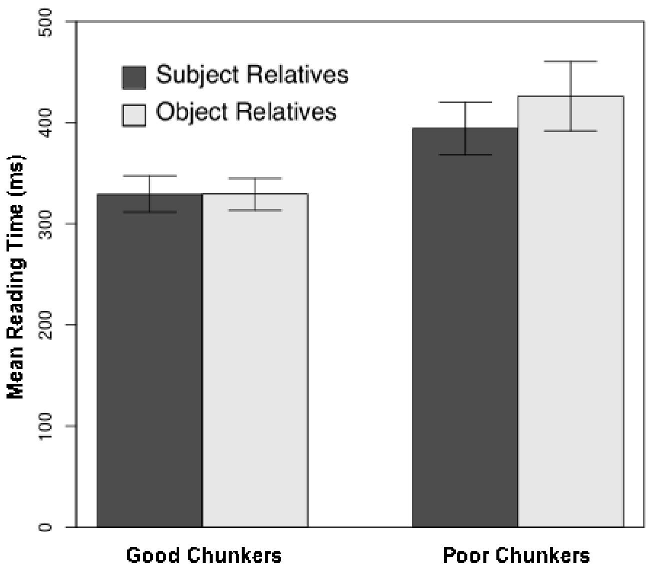

Fig. 3: Mean reading times (RTs) across subject and object relative clauses for individuals measured to have good and poor chunking ability in Experiment 1.

As can be seen in Figure 3, the difference between whole-clause RTs for SRs and ORs was greater for poor chunkers than for good chunkers, as confirmed by the significant Chunk Sensitivity x Clause Type interaction in the LME. This finding provides a qualitative fit with patterns from previous studies, in which good statistical learners (e.g., Misyak et al., 2010), individuals measured to have high verbal working memory (King & Just, 1991), and high-experience individuals (Wells et al., 2009) showed little difference between SR and OR processing, whereas differences were greater for lower-performing individuals.

Crucially, however, previous studies have examined RTs at the critical main verb. In the present study, correlations between Chunk Sensitivity and RTs for the critical main verb did not reach significance for either clause type. However, good chunkers exhibited faster RTs at the critical main verb for both clause types: a *Clause Type* (SR vs. OR) x *Chunking Ability* (Good vs. Poor) ANOVA yielded a significant main effect of *Chunking Ability* (F(1,68)=4.16, p<0.05) alongside the expected effect of *Clause Type* (F(1,68)=8.08,p<0.01). Unlike previous individual differences studies, there was no significant interaction with clause type: the chunking advantage was not significantly greater for the critical main verb in ORs, as might be predicted on the basis of previous work.

The finding of a significant effect of chunking ability for the main verb is noteworthy, as the main verb involves a long-distance dependency with the head noun: that greater chunking ability is tied to lower RTs at the main verb supports the hypothesis that better chunking of the relative clause material can reduce the computational demands imposed by long-distance dependencies.

In sum, our measure of chunking skill predicted reading times for relative clauses, consistent with the notion that chunking at higher levels may reduce the computational demands involved in processing embeddings. Because success in our chunking task requires sensitivity to consonant clusters in written text, it seems reasonable to assume that more experienced readers will fare better on this task than less experienced individuals. That chunking ability more reliably predicted reading times for ORs than for SRs is therefore consistent with the view that increased language experience may reduce processing difficulties more for ORs than for SRs (Wells et al., 2009). Further work will be necessary in order to tease apart the differential impact of chunking ability on clause-internal regions, the focus of the present study, versus the main verb region, which has been the focus of previous individual differences work.

General Discussion

In the present study, we have shown that individual differences in chunking ability predict on-line sentence processing abilities. In Experiment 1, we tested a novel twist on a paradigm previously used to study chunking: the serial recall task. The results revealed considerable variation in participants' ability to successfully generalize previous knowledge of sub-lexical chunks of letter consonants to novel contexts. In Experiment 2, subjects processed SR and OR sentences in a self-paced reading task. Chunking performance from Experiment 1 was then used to predict RTs within the embedded clause and at the critical main verb for both relative clause types. Chunking ability successfully predicted RTs for both OR and SR sentences.

These findings suggest that chunking is relevant for understanding language processing, in line with the notion that chunking takes place at multiple levels: low-level chunking of sub-lexical letter sequences successfully predicted complex sentence processing abilities, consistent with the notion that chunking may reduce the computational burden imposed by embeddings and long-distance dependencies during sentence processing.

This work is also of relevance to understanding language acquisition: as described in the introduction, the Now-or-Never bottleneck requires that language learning take place in an incremental, on-line fashion, suggesting an integral role for chunking. This is consistent with previous computational modeling work showing that chunking can account for key findings relevant to children's phonological knowledge and word learning abilities (e.g., Jones, Gobet, Freudenthal, Watson, & Pine, 2014) as well as work which has sought to model the role of chunking in language learning during on-line processing (McCauley & Christiansen, 2011, 2014). Future behavioral work will examine individual differences in chunking ability in a developmental context, attempting to trace the impact of chunking on specific aspects of acquisition, including the early development of complex sentence processing.

The need for further individual differences work with adults is underscored by the finding that good chunkers had fewer difficulties in relative clause processing, while poor chunkers were shown to have greater difficulties in OR processing relative to SR processing, consistent with previous findings from individual differences studies on statistical learning (Misyak et al., 2010) and verbal working memory (King & Just, 1991). This raises the intriguing possibility that chunking may partly mediate the relationship between those more nebulous constructs and aspects of sentence processing, consistent with the finding that individual differences in language experience are tied to similar SR/OR effects (Wells et al., 2009). Future work will seek to gauge the relative importance of chunking for language processing in individual differences studies which examine chunking ability alongside measures of working memory and statistical learning.

Acknowledgments

We wish to thank Jess Flynn, Scott Goldberg, Yena Kang, Jordan Limperis, Sam Reig, and Sven Wang for assistance with running participants. Thanks to Erin Isbilen and Nick Chater for helpful comments and discussion.

References

Arnon, I. & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62,* 67-82.

Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning. *Psychological Science*, *19*, 241-248.

Christiansen, M.H. & Chater, N. (in press). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-114.

Culicover, P.W. & Jackendoff, R. (2005). *Simpler syntax*. New York: Oxford University Press.

Davies, M. (2008) *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at http://corpus.byu.edu/coca/.

Ericsson, K.A., Chase, W.G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208,* 1181-1182.

Janssen, N. & Barber, H.A. (2012). Phrase frequency effects in language production. *PloS ONE 7*: e33202.

Jones, G., Gobet, F., Freudenthal, D., Watson, S. E., & Pine, J. M. (2014). Why computational models are better than verbal theories: the case of nonword repetition. *Developmental Science, 17*, 298-310.

Just, M.A., Carpenter, P.A. & Woolley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

King, J. & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580-602.

McCauley, S.M. & Christiansen, M.H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1619-1624). Austin, TX: Cognitive Science Society.

McCauley, S.M. & Christiansen, M.H. (2014). Acquiring formulaic language: A computational model. *Mental Lexicon, 9*, 419-436.

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

Miller, G.A. & Taylor, W.G. (1948). The perception of repeated bursts of noise. *Journal of the Acoustic Society of America, 20,* 171-182.

Misyak, J.B., Christiansen, M.H. & Tomblin, J.B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1,* 31.doi:10.3389/fpsyg.2010.00031.

Perruchet, P. & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language,39*, 246-263.

Reali, F. & Christiansen, M.H. (2007). Word-chunk frequencies affect the processing of pronominal object-relative clauses. *Quarterly Journal of Experimental Psychology, 60,* 161-170.

Remez, R.E., Ferro, D.F., Dubowski, K.R., Meer, J., Broder, R.S. & Davids, M.L. (2010). Is desynchrony tolerance adaptable in the perceptual organization of speech? *Attention, Perception, & Psychophysics*, *72*, 2054-2058.

Studdert-Kennedy, M. (1986).Some developments in research on language behavior. In N.J. Smelser & D.R. Gerstein (Eds.), *Behavioral and social science: Fifty years of discovery: In commemoration of the fiftieth anniversary of the "Ogburn Report: Recent Social Trends in the United States"* (pp. 208-248). Washington, DC: National Academy Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Wells, J.B., Christiansen, M.H., Race, D.S., Acheson, D.J. & MacDonald, M.C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58*, 250-271.