Two distinct origins of long-term learning effects in verbal short-term memory

Steve Majerus & Trecy Martinez Perez

University of Liege, Belgium & Fund for Scientific Research FNRS

Klaus Oberauer

University of Zurich, Switzerland, and University of Bristol, UK

Published in: Journal of Memory and Language

Author Note

Steve Majerus, Department of Psychology, Cognition & Behavior, University of Liege, and Klaus Oberauer, Department of Psychology, University of Zurich

This work was facilitated by a grant from the Fonds de la Recherche Scientifique – FNRS (Belgium), grant “Crédit aux Chercheurs 2009-2010 1.5.056.10”, to the first author, and a grant from the Economic and Social Research Council (ESRC), grant RES-062-23-1199 to the second author. We thank Franziska Fischer, Pauline Lanquetin, and Mahaut Sermonard for help with collecting the data.

Correspondence concerning this article should be addressed to Steve Majerus, Department of Cognitive and Behavioral Sciences, University of Liege, Boulevard du Rectorat B33, 4000 Liege, Belgium. E-mail: smajerus@ulg.ac.be

Abstract

Verbal short-term memory (STM) is highly sensitive to learning effects: digit sequences or nonword sequences which have been rendered more familiar via repeated exposure are recalled more accurately. In this study we show that sublist-level, incidental learning of item co-occurrence regularities affects immediate serial recall of words and nonwords, but not digits. In contrast, list-level chunk learning affects serial recall of digits. In a first series of experiments, participants heard a continuous sequence of digits in which the co-occurrence of digits was governed by an artificial grammar. In a subsequent STM test participants recalled lists that were legal or illegal according to the rules of the artificial grammar. No advantage for legal lists over illegal lists was observed. A second series of experiments used the same incidental learning procedure with nonwords or non-digit words. An advantage for legal *versus* illegal list recall was observed. A final experiment used an incidental learning task repeating whole lists of digits; this led to a substantial recall advantage for legal *versus* illegal digit lists. These data show that serial recall of non-digit words is supported by sublist-level probabilistic knowledge, whereas serial recall of digits is only supported by incidental learning of whole lists.

Keywords: short-term memory, phonological processing, statistical learning, Hebb learning

199 words

Many studies document a close interaction between short-term memory (STM) and long-term memory (LTM), leading a number of authors to consider that a substantial part of STM is based on temporarily activated long-term representations in the language system (Baddeley, Gathercole, & Papagno, 1998; Botvinick, 2005; Burgess & Hitch, 2006; Cowan, 1995; Gupta, 2003; Majerus & D'Argembeau, 2011; Martin & Saffran, 1992; Oberauer, 2009). Long-term learning effects on short-term recall provide the most direct evidence for these interactions, by demonstrating a close dependency between representations acquired over the long term and their subsequent use in STM tasks. The aim of the present study is to gain a deeper understanding of the mechanism that underlies these learning effects in STM. We focus our investigation on immediate serial recall of verbal materials because this is the area in which most research on the interplay between STM and LTM has been conducted so far.

A direct impact of verbal long-term learning on verbal STM has been demonstrated with two different paradigms, statistical learning and Hebb learning. In what follows we review the results obtained with these paradigms.

Effects of incidental statistical learning on verbal STM

Incidental learning of statistical regularities in sequences have been shown to improve immediate serial recall of lists conforming to these regularities, compared to lists not conforming to them. Karpicke and Pisoni (2004) asked participants to recall lists of color words by pressing appropriate buttons. In the acquisition phase, all lists were constructed according to an artificial grammar. In the test phases, lists constructed from the same grammar were mixed with lists constructed by an analogous alternative grammar. Serial recall was better for lists that conformed to the learned grammar than those that didn't.

Botvinick (2005) and Botvinick and Bylsma (2005) replicated the finding of improved serial recall of lists conforming to a learned artificial grammar, using one-syllable nonwords as items. Furthermore, Botvinick and Bylsma (2005) observed that the errors in the immediate serial recall task progressively adopted the phonological regularities of the artificial grammar.

Majerus, Van der Linden, Mulder, Peters, and Meulemans (2004) pushed this line of research one step further by demonstrating that merely listening to a continuous sequence of stimuli constructed from an artificial grammar improved serial recall of lists conforming to that grammar. The present study builds directly on their work, and therefore we summarize it in some detail here. Children and adult participants were exposed to a continuous stream of phonemes during 20 minutes. Phoneme successions were determined by an artificial phonotactic grammar (e.g., /b/ could be followed by /a/ or /o/, but not by /i/ or /u/). At the phoneme level, there were four consonants and four vowels that could be combined into eight legal syllables (which were determined by the phonotactic grammar). At the syllable level, each of the eight legal syllables could be followed by four of the eight legal syllables, leading to a total of 32 legal combinations. Participants were not required to pay attention to the phoneme stream; rather, they had to complete a complex coloring task while hearing the sequence. Afterwards, participants completed a test of immediate serial recall of nonwords. The nonwords were either legal or illegal according to the artificial phonotactic grammar that generated the continuous sequence of phonemes; legal and illegal phoneme combinations were in addition matched for native language phonotactic frequency. Participants showed significantly higher recall performance for legal nonwords versus illegal nonwords. This result shows that the participants had detected the regularities that determined the succession of the different phonemes, had acquired this knowledge, and used it in a STM task.

Taken together, these results show that statistical learning, that is, acquisition of knowledge about the transition probabilities or co-occurrence probabilities of successive elements in sequences, contributes to immediate serial recall of verbal material. This conclusion dovetails with evidence that statistical sublexical knowledge about people's spoken language has an effect on their verbal STM. Gathercole, Frankish, Pickering, and Peaker (1999) and Thorn and Frankish (2005) showed that phonotactic knowledge supports STM performance. Phonotactic knowledge refers to the phoneme-co-occurrence statistics of a given language: some phonemes co-occur very frequently (e.g., /ɪp/ for English) while others co-occur more rarely (e.g., /ɛz/ for English). Immediate serial recall performance for lists of nonwords of high phonotactic frequency leads to higher recall performance than immediate serial recall performance for nonwords of low phonotactic frequency (see also Majerus & Van der Linden, 2003).

Two explanations have emerged from the literature on statistical learning for how knowledge of transition probabilities is acquired (Perruchet & Pacton, 2006). One possibility is that the cognitive system gradually adjusts associations between adjacent elements (e.g., between phonemes in verbal sequences), such that the strength of associations reflect the transition probabilities. The alternative mechanism rests on the formation of chunks of small segments of a sequence, typically 2 to 5 elements of length, together with the gradual accumulation of knowledge of the relative frequencies of these chunks.

The Hebb effect

In the prototypical Hebb learning paradigm participants are repeatedly tested on immediate serial recall of short lists, usually consisting of digits. On every third trial, the same digit sequence is presented, leading to higher immediate serial recall performance for the repeated compared to unrepeated sequences. The effect has been initially observed by Donald Hebb (1961) and has led to extensive theoretical developments (e.g., Burgess & Hitch, 1992; Rumelhart, McClelland, & the PDP Research Group, 1986).

Burgess and Hitch (1992, 1999) proposed an influential computational model of short-term memory that also accounts for Hebb learning. The basis of this model is the language network, represented by local representations at the lexical and sublexical level. The lexical nodes are linked with a dynamic context layer, changing over time, and enabling the encoding of serial position, by associating each item activated in the lexical network with a distinct state of the context layer. According to this model, the Hebb effect occurs due to the strengthening of connections between the context layer and item nodes for items within repeated lists. In other words, this model assumes that a strengthening of associations between items and serial positions accounts for the Hebb learning effect. However, a study by Cumming, Page, and Norris (2003) obtained evidence against this account: when comparing non-repeated and partially repeated lists where items in some but not all positions were repeated (e.g., 7 8 5 1 6 4 9 2 3, 7 4 5 2 6 8 9 1 3, 7 1 5 8 6 2 9 4 3), no advantage for partially repeated over non-repeated lists was observed, suggesting that the Hebb affect does not arise from the strengthening of individual item-position associations.

A second possibility is that the learning system automatically detects item co-occurrences and strengthens inter-item connections for repeatedly co-occurring items. This possibility is questioned by results from Hitch, Fastame, and Flude (2005). They showed that learning in a Hebb paradigm did not transfer to lists that maintained the item-to-item transitions of the learned list but shifted them to new serial positions.

A third possibility, endorsed by Burgess and Hitch (2006) in a revision of their earlier model, is that repeated encoding of the same list into working memory contributes to the gradual strengthening of a unified representation of that list's sequence in long-term memory. A similar idea has been developed by Grossberg and Stone (1986). These unified long-term memory representations would assist recall of a new list to the degree that the new list is similar to the list represented in long-term memory, so that the latter is retrieved during encoding of the former. Evidence from a series of transfer experiments with the Hebb effect lends most support to this third theoretical option (Hitch et al., 2005).

The present study

The two paradigms discussed above suggest the existence of at least two types of long-term learning in STM tasks. In the Hebb paradigm participants are usually presented with the exact predetermined list, or at least substantial segments of a list, repeatedly, and the Hebb effect critically depends on the identical repetition of at least the initial part of the list (Hitch et al., 2005). Furthermore, demonstrations of the Hebb effect rely on learning of isolated lists with a distinct beginning and end. In this sense, the Hebb learning paradigm can be considered to be a kind of list-level learning task. In contrast, studies on statistical learning expose people to sequences governed by unequal transition probabilities between elements, such that certain pairs and triplets of elements occur more frequently than others, but identical repetitions of longer series are increasingly infrequent; this type of learning focuses on acquisition of local transition probabilities, spanning four to five elements at best. This situation is best described as probabilistic learning of sublist-level information.

The aim of the present study is to investigate whether the two kinds of learning can be explained by a single mechanism. To foreshadow, we obtained evidence for a dissociation between the two forms of learning. Specifically, we will show that learning of sublist-level, probabilistic regularities affects STM for phonemes, words, and nonwords, but not for digits. In contrast, list-level learning affects STM for digits. This pattern provides a challenge for an explanation of long-term learning effects on STM in terms of a single learning mechanism. In the General Discussion we will explore a theoretical option that could nevertheless provide a unified account of our results.

Our investigation of incidental learning effects with different kinds of verbal materials was motivated by two observations. First, the evidence summarized above has been obtained with different kinds of materials. As we have already noted, Hebb learning is most typically and frequently demonstrated using repeated and random strings of digits (e.g., Hitch, Flude & Burgess, 2009; Cumming, 2001; Hebb, 1961; Oberauer & Meyer, 2009). In contrast, so far, the experiments showing that learning of item co-occurrence regularities affects immediate serial recall have only used unfamiliar syllables.

Second, digits differ from other linguistic units (such as syllables, nonwords, and words) with regard to their familiarity and their typical use, which afford different kinds of learning. The combination of syllables and words in a given language is characterized by a number of probabilistic constraints and regularities, which can be acquired through probabilistic learning. By presenting a continuous sequence of unfamiliar verbal segments, the listener is put in a very similar condition as the prelinguistic child starting to construct phonological representations of words, but still having no lexical and lexical-semantic representations to map on these phonological representations. This situation might favor the learning of item co-occurrence regularities, as phonological information and underlying regularities are the only information available to the listener. Furthermore, the extraction of sublexical item-co-occurrence probabilities is important in everyday life speech processing as it allows to quickly identify uncommon phonological patterns and to adjust the phonological system to these new patterns, as is the case for instance when a listener is confronted for the first time with a speaker using a particular dialect (e.g., Cambridge-style English versus Scottish English) or with young children using phonologically deformed speech. In some way, we are well trained to process sublexical phonological regularities (and irregularities) and to use these properties in everyday life (e.g., Clopper & Pisoni, 2008; Johnson, 2005). In contrast, the combinations of digits are not subject to these regularities but are used in any possible combination to form numbers. Because in the everyday use of digits, every combination of digits into multi-digit numbers is equally likely to occur, every pairwise association or co-occurrence of digits is likely to be acquired to the same degree. There is no useful sublexical, probabilistic information to be extracted from digit co-occurrences in numbers. At the same time, some strings of digits might be used very often (e.g., a familiar phone number, a PIN). These multi-digit numbers could be acquired as unified representations, as proposed by Burgess and Hitch (2006) in their explanation of the Hebb effect. If our incidental learning mechanisms are tuned to the specific learning affordances of different kinds of verbal materials, then probabilistic learning of item co-occurrences might work for non-digit verbal units such as phonemes, syllables, and words, but not for digits. In contrast, learning of frequently repeated strings as unified chunks might work for all kinds of verbal units, including digits.

Experiment 1:

Incidental Learning of Item-Co-occurrence Regularities for Continuous Digit Sequences

All the experiments presented here used an incidental learning paradigm similar to the one used by Majerus et al. (2004). Experiment 1 explored whether the same type of incidental learning of item-co-occurrences could also be obtained for very familiar stimuli such as digits, the material most typically used in Hebb learning experiments. Given that digits are familiar syllables with a fixed phonological structure, the artificial phonotactic grammar succession rules used in the present experiment were restricted to the syllable-level. Each digit syllable could be followed by two other digit syllables, leading to a total of 18 possible combinations (see Figure 1). This artificial phonotactic grammar is simpler than the one used in the previous study by Majerus et al. (2004), and is closer to the Hebb learning situation where, for a sequence of 9 digits, 8 item successions have to be learned. Also, as in Majerus et al. (2004), it was ensured that learning was incidental by focusing the participants' attention on a complex coloring task during the presentation of the continuous sequence of digit syllables.

Method

Participants

Thirty University of Bristol undergraduates (between 17 and 35 years old) participated in a one-hour session for course credit or a remuneration of 7 GBP. All were fluent English speakers.

Material

The digits from 1 to 9 were recorded by a male voice and transformed into digital sound files. The sound file's durations varied from .52 to .91 s depending on speaking duration. During the learning sequence, 2200 digits were presented in an uninterrupted sequence with 10 ms in between successive sound files. The succession of the digits was determined by an artificial grammar, shown in Figure 1. Each arrow in the figure represents a possible transition. In generating legal sequences, the two transitions emanating from each digit were chosen at random with equal probability. To avoid any bias related to specific digits or pairs of digits, we created nine versions of this grammar (only one of which is shown in Figure 1) by rotating the assignment of digits to roles in the grammar: For each new participant the digit in each role was incremented by one, wrapping around from 9 to 1. For instance, participant 1 experienced transitions from 1 to 5 or 8 (each with probability 0.5), transitions from 2 to 3 or 5, and so on, as illustrated in Figure 1. For participant 2, the corresponding part of the grammar would generate transitions from 2 to 6 or 9, and transitions from 3 to 4 or 6. In this way, after 9 participants each digit had each role in the grammar once. The duration of each sequence was about 25 minutes. For the short-term memory task, the digit recordings were presented in sequences of 9 digits, played at the same pace as during the implicit-learning phase. There were two blocks of trials, one with lists that were entirely legal according to the grammar, and the other block consisting of lists that were illegal. The legal lists were constructed by randomly selecting a digit to initiate the list, and by deriving the remaining 8 digits according to the grammar (in the version chosen for the given participant). We furthermore ensured that each of the 18 possible legal digit successions occurred at a similar frequency throughout the 27 trials (each legal transition occurred between 9 and 16 times). To avoid any confound between the legality status of lists and their structure, we constructed the illegal lists by adding 3 to each digit in the legal lists, wrapping around from 9 to 1. Each block consisted of 27 trials. The presentation of learning and short-term memory sequences was controlled via the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) running on a MATLAB platform.

< INSERT FIGURE 1 ABOUT HERE >

Procedure

The incidental learning sequence was presented through headphones via a PC computer, at a comfortable loudness. At the same time, the participants were presented complex line drawings on paper sheets they had to color with crayons provided by the experimenter. The coloring task was presented to the participants as the main task and they were encouraged to color the drawings as creatively as possible without paying attention to the sound sequence, in line with the procedure developed by Saffran, Newport, and Aslin (1996) and Saffran, Newport, Aslin, and Tunick (1997). Immediately after the incidental learning phase, the short-term memory task was administered: each participant was presented the blocks of legal and illegal digit lists, the order of blocks being counterbalanced between participants. The participants were instructed to recall each digit list as accurately as possible after its presentation, and they were asked to guess when not remembering a digit in a given position. The responses were tape-recorded for later transcription and scoring. We determined the mean number of digits recalled in correct serial position for each condition. At the end of the experiment, the participants were informed of the primary goal of the study, according to APA ethical standards.

Results and Discussion

Mean proportion of digits recalled in the correct position was .65, with no discernible difference between legal and illegal lists, F (1, 29) = 0.20, MSE=.032, p=0.66, η^2_p= .007 (legal: mean = .65, SD = .17; illegal: mean = .64, SD=.14). In the experiment of Majerus et al. (2004), the effect size for the contrast between legal lists and illegal lists with violations on the syllable level, which is directly comparable to the present comparison of legal and illegal lists, was f = .72. The power to detect an effect of that size in our experiment was > .99. The conventional significance test only reflects the strength of evidence for the alternative hypothesis. The strength of evidence for the null hypothesis can be assessed by computing the likelihood ratio between the null hypothesis and the alternative hypothesis, penalizing the alternative hypothesis for its additional free parameter (i.e., the free estimate of the effect) by the correction formula for AIC or for BIC (Glover & Dixon, 2004; Wagenmakers, 2007). Applying this procedure to the findings from our experiment reveals a likelihood ratio of 2.8 in favor of the null hypothesis after penalizing according to the AIC formula, and a likelihood ratio of 5.1 in favor of the null hypothesis after penalizing according to the BIC formula. Therefore, the present data are not merely a failure to obtain evidence for an effect of incidental learning on serial recall of digits; they provide modest evidence for the null hypothesis that there is no such effect. Similar results were observed when analyzing performance as a function of serial position: a condition (2) by serial position (9) ANOVA showed no significant main effect of legality, F (1, 29) = 0.20, MSE=.032, p=0.66, η^2_p= .007, but a significant effect of serial position, $F(8, 232) = 62.7$, $MSE=.019$, $p<0.001$, $\eta^2_p= .68$; the interaction was not significant, $F(8, 232) = .30$, $MSE=.007$, $p<0.96$, $\eta^2_p= .01$; further exploration via planned comparison showed significant linear, $F(1, 29) = 81.5$, $MSE=.07$, $p<.001$, $\eta^2_p= .73$, and quadratic contrasts, $F(1, 29) = 54.9$, $MSE=.026$, $p<.001$, $\eta^2_p= .65$ for serial position (see Figure 2).

< INSERT FIGURE 2 ABOUT HERE >

Experiment 2:

Incidental Learning of Item Co-occurrence Regularities for Continuous Digit and Nonword Sequences

Experiment 1 revealed no evidence for incidental learning of item-co-occurrence regularities contained in a continuous sequence of digit syllables. It is unlikely that the artificial grammar was too complex to generate learning since the grammar used here only contained one level of rules (syllable-level), relative to a previous study which contained two levels of rules (syllable-level and phoneme-level) and yet produced significant effects of learning (Majerus et al., 2004). Furthermore, the syllable-level rules used here were simpler than those used by Majerus et al. The important difference seems to be that Majerus et al. used nonword syllables whereas here we used digit, raising the possibility that the incidental abstraction of item co-occurrence statistics does not work for digits as items. The aim of Experiment 2 was to replicate the null effect observed in the first Experiment, while showing that the grammar we used can actually lead to incidental learning when using nonwords rather than digits. Although Majerus et al. already showed that incidental learning of co-occurrence statistics is possible for nonwords, the two-level grammar used in that former study was different from the one-level grammar used in the present study. Furthermore, the syllables used in Majerus et al. were simple CV syllables, while the phonological forms of digits are characterized by more variable and complex syllable structures which might have interfered with the learning of syllable-co-occurrence regularities. In the following experiment, the syllable structure of nonwords was chosen to match the phonological structure of digits. Finally, because we expected serial recall of nonwords to be worse and more variable than that of digits, the short-term memory lists used in Experiment 2 were of increasing length, from two to ten syllables, to capture representative levels of short-term memory performance for individuals of varying short-term memory capacity.

Method

Participants

Ninety participants (79 women) aged between 18 years and 31 years participated in this experiment on a voluntary basis (no compensation fee nor course credit); they were randomly assigned to one of two experimental groups. All participants were native and monolingual French speakers with higher levels of education. Informed written consent was obtained for each participant.

Material

The nonwords were created by starting from the syllabic structure of French digit words, changing in each digit the minimal number of phonemes necessary to obtain a nonword (see Table 1). Whenever possible, the phonemes of two digit words were swapped when creating the nonwords, enabling us to use a similar pool of phonemes for digit and nonword stimuli. Furthermore, the phonotactic frequency of nonword diphones was matched to the phonotactic frequency of the diphones of digits, based on the phonetic corpus of French by Tubach and Boë (1990), $F(1,8)=.20$, $MSE=37398$, $p=.67$, $\eta^2_p=.02$. Both digits and nonwords were recorded by a female native French speaker and transformed into digital sound files. The learning sequences contained 3795 syllables presented at a monotonous regular rate, successive syllables being separated by a 200 ms silence period. The succession of the syllables was determined by the same type of artificial grammar as in Experiment 1 (see Figure 1), different variants of the same grammar being used for the different participants. The duration of each sequence was about 37 minutes. For the short-term memory task, the syllable recordings for legal and illegal lists were presented by increasing length, starting at list length 2 and ending at list length 10, with three trials per list length (27 lists per block). The construction of legal and illegal lists was identical to the procedure described in Experiment 1, except that illegal lists were constructed by sampling exclusively from illegal syllable successions; in Experiment 1, illegal lists could contain a small amount of legal syllable successions due to illegal lists resulting from the addition of 3 to legal lists which in a minority of instances led to a legal succession (seven of the eighteen legal transitions could actually occur in illegal lists, with the same average frequency as in legal lists). The presentation of learning and short-term memory sequences was controlled via the Cogent Toolbox (FIL, University College London, 2000) running on a MATLAB platform.

< INSERT TABLE 1 ABOUT HERE >

Procedure

The incidental learning sequence was presented though headphones via a PC computer, at mean output amplitude of 70 dB SPL. Half of the participants heard the digit sequence, and half of the participants heard the nonword sequence. During exposure to the sequence they worked on the same drawing task as in Experiment 1. After the incidental learning phase, the short-term memory task was administered: each participant was presented the blocks of legal and illegal digit or nonword lists, the order of blocks being rotated between participants. The participants were instructed to recall each list as accurately as possible after its presentation. The responses were tape-recorded for later transcription and scoring. We determined the mean number of digits recalled in correct serial position for each condition, pooling over the different list lengths. At the end of the experiment, the participants were informed of the primary goal of the study, according to APA ethical standards.

Results

For the digit incidental learning group, a repeated measures ANOVA on the number of digits correctly recalled in the short-term memory task showed an absence of list condition, legal and illegal lists leading to identical levels of recall, replicating the findings of Experiment 1, $F(1,44)=.83$, $MSE=42.0$, $p=.37$, $\eta^2_p=.02$ (Figure 3). As in Experiment 1, we applied the procedure proposed by Glover and Dixon (2004) to the findings from Experiment 2, revealing very comparable results, with a likelihood ratio of 2.07 in favor of the null hypothesis after penalizing according to the AIC formula, and a likelihood ratio of 4.40 in favor of the null hypothesis after penalizing according to the BIC formula. These results again provide modest evidence for the null hypothesis that there is no effect of incidental learning in the digit learning group. On the other hand, for the nonword incidental learning group, a repeated measures ANOVA on the number of nonwords correctly recalled in the short-term memory task revealed a significant effect of list condition, $F(1,44)=9.11$, $MSE=92.2$, $p<.01$, $\eta^2_p=.17$. This discrepant finding for the digit and nonword incidental learning groups was furthermore confirmed by a mixed ANOVA, with list condition as within subjects factor and learning group as between subjects factor: the interaction between learning group and list condition was significant, $F(1,88)=3.97$, $MSE=67.16$, $p<.05$, $\eta^2_p=.04$. By exploring the interaction via planned comparisons, we confirmed an effect of legal list condition for the nonword incidental learning group, $F(1,88)=12.51$, $MSE=67$, $p<.001$, but not for the digit incidental learning group, $F(1,88)=.52$, $MSE=67.16$, $p=.47$. This mixed ANOVA further showed a main effect of learning group, overall recall performance being higher in the digit incidental learning group as opposed to the nonword incidental learning group, $F(1,88)=99.01$, $MSE=632.0$, $p<.001$, $\eta^2_p=.53$.

< INSERT FIGURE 3 ABOUT HERE >

In a next set of analyses, we considered performance as a function of list length. For the digit incidental learning, a repeated measures ANOVA on the proportion of digits correctly recalled as a function of list lengths 4 to 10 (performance was invariably at ceiling for list lengths 2 and 3) revealed no main effect of legal *versus* illegal list condition, $F(1,44)=.58$, $MSE=.01$, $p=.45$, $\eta^2_p=.01$, but as expected, a main effect of list length, $F(6,264)=167.98$, $MSE=.03$, $p<.001$, $\eta^2_p=.79$. The interaction was not significant, $F(6,264)=1.99$, $MSE=.01$, $p=.07$, $\eta^2_p=.04$. Given that the interaction was nevertheless close to significance, planned comparisons were conducted, showing no significant effect of list condition (at $p<.05$) for any list length. As shown in Figure 4a, the marginal significant interaction was due to performance tending to be higher for the illegal condition at list lengths 6 and 8, and somewhat higher for the legal condition at list length 7. Hence, as for the previous analyses, there was no evidence for a reliable effect of list condition in the digit incidental learning group. The same analysis was conducted for performance in the nonword incidental learning group. In addition to the expected main effect of list length, $F(6,264)=296.20$, $MSE=.02$, $p<.001$, $\eta^2_p=.87$, a main effect of legal versus illegal list condition was observed for the nonword incidental learning group, $F(1,44)=11.61$, $MSE=.03$, $p<.001$, $\eta^2_p=.21$; both factors also significantly interacted, $F(6,264)=4.06$, $MSE=.01$, $p<.001$, $\eta^2_p=.08$. Planned comparisons (all p's <.05) showed that performance was higher for the legal condition as compared to the illegal condition most reliably for lists at length 5, 6 and 7 (see

Figure 4b). For shorter and longer list lengths, no reliable advantage for recall of legal lists is observed due to ceiling and floor levels of performance, respectively.

< INSERT FIGURE 4 ABOUT HERE >

A final analysis explored errors produced during recall of illegal digit and nonword lists, in order to determine whether these errors show a tendency towards regularization, reflecting regularities of the learned artificial grammar (Botvinick & Bylsma, 2005). We restricted our error analysis to incorrect syllable successions that were produced, excluding omission errors and extra-list intrusion errors (syllables different from those used in the short-term memory lists); we then determined for each syllable succession whether it was legal according to the grammar to be learnt. Finally, we determined the proportion of legal syllable successions relative to the overall amount of incorrect syllable successions produced, and compared this proportion to the chance probability of producing legal syllable successions (given that there were 9 syllables that could be associated in 72 different pairs, and given that our grammar only allowed for 18 of these pairs, the chance probability for legal successions in the absence of any learning is 18/72 = .25). We did not perform this analysis for legal list recall because errors in the legal condition are likely to include legal segments from the target sequence to be recalled due to target segments simply migrating in serial position (remember that we used a strict serial recall criterion); hence it is impossible to determine whether these errors just reflect serial position migration errors of syllables from the target short-term memory list or a real influence of sequence learning on error production. For digit sequence learning, we observed a mean proportion of .26 (SD: .08) legal successions from a total of 25.53 (SD: 14.7) incorrect syllable successions that were produced: this proportion did not significantly differ from a theoretical distribution centered on the expected proportion of .25, $t(44)=1.14$, $p=.26$. In contrast, for nonword sequence learning, the proportion of legal successions was .34 (SD: .18), from a total of 14.11 (SD: 7.98) incorrect syllable successions produced: this proportion was significantly higher than what could be expected from chance, $t(44)=3.16$, $p<.01$.

Discussion

On the one hand, these results replicate those observed in Experiment 1, by revealing a clear absence of incidental learning on digit short-term memory recall performance. Given that short-term memory recall procedures were different in Experiment 1 and 2, with a fixed length supra-span recall procedure in Experiment 1 and an increasing list length recall procedure in Experiment 2, the present experiment shows that the null effect in Experiment 1 was not caused by the specific recall procedure (which led to overall lower performance levels in Experiment 1). Furthermore, despite the use of different levels of analyses, considering either overall levels of performance or performance as a function of list length, there was not even a hint for the possibility of a digit sequence learning effect. This was further confirmed by an error analysis, showing that errors in the digit sequence learning group did not reflect the incidental grammar embedded in the learning sequence. On the other hand, Experiment 2 demonstrates that the incidental grammar embedded in the learning sequence can be learned for nonword stimuli. Learning was manifest not only as an effect on overall recall but also as a tendency of errors on illegal lists to conform to the learned grammar more often than chance, replicating the regularization effect of Botvinick and Bylsma (2005). These contrasting findings cannot be accounted for by differences at the level of phonological characteristics of the digit and nonword stimuli, because both were strictly matched for phonological structure and native language phonotactic frequency.

The comparison between the positive effect of learning for nonword sequences and the null effect of learning for digit sequences is difficult to interpret on the level of overall accuracy, because there was a large difference in overall performance levels for digit and nonword post-learning recall performance. However, the analyses as a function of list length showed that the most pronounced learning effect for nonword sequences was observed at list lengths 5 to 7, yielding performance levels ranging between .60 and .35 for illegal lists. Similar levels of performance were observed for digit recall at list lengths 8 to 10 in Experiment 2, and yet no evidence for learning was observed for these list lengths. Therefore, the difference in learning between digits and nonwords is unlikely to be caused by the different overall level of memory accuracy.

Experiment 3:

Incidental Learning of Item Co-occurrence Regularities for Continuous Word and Nonword Sequences

Experiment 3 further explores the item characteristics that do or do not lead to learning of item-co-occurrence regularities. Experiments 1 and 2 suggest that incidental learning of item-co-occurrence regularities for digit items is difficult to obtain, while the same regularities can be learned for nonword items. The goal of Experiment 3 is to investigate whether the null effect of incidental learning of item-co-occurrence regularities for digit sequences is due to the specific linguistic category of digits, or more generally due to the lexical status of digit items. In Experiment 3 we used the same incidental learning procedure as in the previous two experiments, but using familiar words and nonwords as learning material. Like digits, the words had a low age of acquisition and were of high lexical frequency, but they were chosen to have a less variable syllabic structure. The less complex syllabic structure of the words enabled us to construct a comparable set of nonword stimuli that, likewise, had a less complex structure than those used in Experiment 2, leading to a higher overall level of recall. Thereby, Experiment 3 created an opportunity to replicate the effect of learning for nonwords at a level of memory performance closer to that for digits.

Method

Participants

Sixty participants (47 women) aged between 18 years and 29 years participated in this experiment on a voluntary basis (no compensation fee nor course credit). All participants were native and monolingual French speakers with higher levels of education. Informed written consent was obtained for each participant. The participants were assigned at random to one of two groups; one was given the words and the other was given the nonword materials.

Material

Familiar, monosyllabic words were selected with a regular CV syllabic structure. The words are shown in Table 2. They were concrete, acquired early and were of moderate-to-high lexical frequency (mean frequency: 65.02; range: 7-196; Lexique 2, New, Pallier, Brysbaert, & Ferrand, 2004). The nonwords were constructed by changing in the word stimuli the minimal number of phonemes necessary to obtain a nonword (see Table 2). Whenever possible, the phonemes of different word forms were swapped when creating the nonwords. Furthermore, the phonotactic frequency of word and nonword diphones was matched, $F(1,8)=.70$, $MSE=7567$, $p=.43$, $\eta^2_p=.08$ (Tubach & Boë, 1990). In all other regards, construction and administration of the material followed the same procedure as in Experiment 2.

< INSERT TABLE 2 ABOUT HERE >

Procedure

The procedure was identical to Experiment 2.

Results

For the word incidental learning group, a repeated measures ANOVA revealed a robust effect of condition during the short-term memory test phase, legal word lists leading to higher recall performance than illegal word lists, $F(1,29)=12.59$, $MSE=57.80$, $p<.01$, $\eta^2_p=.30$ (see Figure 5). For the nonword incidental learning group, a robust effect of short-memory condition was also observed, with significantly higher recall performance for legal nonword lists than illegal nonword lists, $F(1,29)=12.75$, $MSE=86.40$, $p<.01$, $\eta^2_p=.31$ (see Figure 5). Furthermore, as shown in Figure 5, much higher nonword recall performance was observed in the present experiment relative to nonword recall performance in Experiment 2. Nonword recall performance and word recall performance were actually at comparable levels in Experiment 3. This was confirmed by a mixed ANOVA, with list condition as within-subjects factor and learning group as between-subjects factor: neither the main effect of learning group, $F(1,58)=0.84$, $MSE=593.00$, $p=.36$, $\eta^2_p=.01$, nor the interaction between list condition and learning group, $F(1,58)=0.27$, $MSE=72.00$, $p=.61$, $\eta^2_p=.005$, were significant.

< INSERT FIGURE 5 ABOUT HERE >

As with Experiment 2, we also analyzed performance as a function of list length. For the word incidental learning group, a repeated measures ANOVA on the proportion of words correctly recalled as a function of list lengths 4 to 10 (performance was invariably at ceiling for list lengths 2 and 3, as for the digit learning group in Experiment 2) revealed a main effect of legal *versus* illegal list condition, $F(1,29)=12.91$, $MSE=.02$, $p<.01$, $\eta^2_p=.31$, and a main effect of list length, $F(6,174)=148.45$, $MSE=.03$, $p<.001$, $\eta^2_p=.84$. The interaction was not significant, $F(6,174)=.82$, $MSE=.02$, $p=.56$, $\eta^2_p=.03$. Although there was an overall effect of legal list condition, Figure 6a shows that the effect was most reliable for the longest list lengths. The same analysis was conducted for performance in the nonword incidental learning group, revealing comparable results: a main effect of legal list condition, $F(1,29)=13.82$, $MSE=.02$, $p<.001$, $\eta^2_p=.32$, a main effect of list length, $F(6,174)=160.14$, $MSE=.02$, $p<.001$, $\eta^2_p=.85$, but no significant interaction, $F(6,174)=1.25$, $MSE=.02$, $p=.28$, $\eta^2_p=.04$. As for the word learning group, Figure 6b shows that the effect was most pronounced for the longest list lengths.

< INSERT FIGURE 6 ABOUT HERE >

Discussion

No error analysis was performed for Experiment 3 due to an insufficient number of incorrect syllable succession productions; 37% of participants produced no incorrect syllable successions at all. Instead, extra-list intrusion errors were very common. The high prevalence of extra-list intrusions can be explained by the larger lexical neighborhood (i.e., the number of familiar words that differ from the target item by a single phoneme) of the phonologically less complex items used in this Experiment, compared to the nonwords used in Experiment 2. The larger phonological neighborhood may also explain the lack of a recall advantage for words relative to nonwords in Experiment 3 (although we should note that the materials in Experiment 3 were specifically designed to lead to more comparable levels of performance for recall of lexical items and nonword items): nonword recall has been shown to benefit from the existence of lexical neighbors (Roodenrys & Hinton, 2002; Thorn & Frankish, 2005).

In sum, these results show that incidental learning of item-co-occurrence regularities is possible for material having only a phonological level of representation (i.e., nonwords) as well as for material having phonological and lexico-semantic levels of representation, such as words. Furthermore, performance levels obtained in Experiment 3 were very similar to those observed for the digit sequence learning group in Experiment 2, further indicating that the null effect of sequence learning for the digit sequence learning group in Experiment 2 cannot be explained by the specific range of performance levels observed in this group, relative to the nonword sequence learning group in Experiment 2.

Experiment 4:

Incidental Learning of Chunked Sequences for Digit Stimuli

The previous experiments show that incidental learning of item-co-occurrence regularities for digit items is difficult to obtain, whereas the same regularities can be learned for nonword and word items. This means that learning of digit sequences as evidenced by the Hebb learning effect must be supported by a different learning mechanism. Following Burgess and Hitch (2006), digit Hebb learning reflects learning of chunks, that is, unified representations of whole lists. In other words, repeated digit lists would be learned by laying down a long-term representation of that particular list as a whole, rather than acquiring knowledge about item-co-occurrence regularities. Experiment 4 tested this hypothesis directly, by using the same incidental learning paradigm as in the previous experiments, but by introducing sequence regularity in the form of repeated lists instead of probabilistic item-co-occurrence regularities. The sequence for the learning phase was constructed from three lists of seven digits, three lists of eight digits, and three lists of nine digits. These lists were concatenated in random order into a single sequence of 3816 digits, which was presented during the learning phase. To make the beginning of a list recognizable, the lists were separated by a short pause. These pauses should facilitate parsing the sequence into lists, which is a prerequisite for learning these lists as chunks.

To rule out any possibility for item-co-occurrence learning, no digit pair was repeated within or between lists; in other words, each of the nine lists was composed of entirely new digit pairs, by sampling without replacement from all possible two-digit combinations (excluding digit repetitions). Because the majority of all 72 possible pairs of digits were used, and they were presented equally often, a learning mechanism sensitive to pairwise co-occurrence frequencies, or first-order transition probabilities, would find it much harder to learn from this sequence than from the sequences in the preceding experiments, where only 18 out of the possible 72 pairwise combinations were used. Thus, whereas in the previous experiments the regularities applied to pairwise combinations of successive elements, in Experiment 4 the regularities existed primarily across larger segments of the sequence. Learning mechanisms that are tuned to statistical regularities on the level of pairs of successive elements, as have been proposed for explaining statistical learning (Perruchet & Vinter, 1998), would find learning harder in Experiment 4 than in the preceding experiments. In contrast, a learning mechanism such as proposed by Burgess and Hitch (2006), which acquires unified representations of whole lists by chunking them, would have a better chance of acquiring the repeated lists (or sub-segments of lists) in the present experiment than in preceding experiments, because repeated sequences are marked by pauses.

At the same time, the present learning paradigm is more demanding than standard Hebb learning paradigms, where only a single list has to be learned at a time, and lists are learned intentionally at least for immediate recall. In contrast to typical Hebb learning experiments, Experiment 4 explores incidental learning of nine lists of digits simultaneously.

Method

Participants

Sixteen participants (12 women) aged between 18 years and 29 years participated in this experiment on a voluntary basis (no compensation fee or course credit). All participants were native and monolingual French speakers with higher levels of education. Informed written consent was obtained for each participant.

Material

Nine sequences containing 7 to 9 digits were created by randomly sampling without replacement from the 72 possible digit pairs combining the digits 1 to 9 (by avoiding pairs of identical digits such as 2 2). Given that we constructed three lists for each sequence length, 63 out of these 72 possible pairs were used, leaving only 9 unused digit pairs. The digit stimuli were presented following exactly the same procedure as described in Experiment 2, with the exception that a pause of 1.5 seconds was inserted between each list to mark the boundaries of the different lists. Each list was presented 53 times throughout the incidental learning phase, for a total of 3816 digit syllable presentations (note that roughly the same number of syllables was presented in Experiments 2 and 3: 3795). For the subsequent short-term memory task, the nine lists as well as nine illegal lists were presented in order of increasing length, starting at list length 7 and ending at list length 9. Illegal lists were constructed by increasing each digit of a legal list by 1, resulting in new lists never presented during the incidental learning phase. Furthermore, we ensured that the 8 digit pairs that were never used in the legal lists, also never occurred in the illegal lists. In other words, the illegal digit lists contained no digit pair that did not also occur in the legal lists; the legal and illegal lists only differed in terms of the higher-order combinations of pairs. Finally, as in the previous experiments, the sequences to be learned differed between participants, by incrementing the digits in each list by one from one participant to the next (wrapping around from 9 to 1) (see Table 3 for an example of the digit lists to be learned). The presentation of learning and short-term memory sequences was controlled via the Cogent Toolbox (FIL, University College London, 2000) running on a Matlab platform.

< INSERT TABLE 3 ABOUT HERE >

Procedure

The procedures for administration of the incidental learning task and the subsequent short-term memory task were identical to Experiments 2 and 3.

Results

A repeated measures ANOVA with condition (2) and list length (3) revealed a robust effect of list condition during the short-term memory test phase, legal digit lists leading to higher recall performance than illegal digit lists, $F(1,15)=50.99$, $MSE=.006$, $p<.001$, $\eta^2_p=.77$ (see Figure 7). In addition there was an effect of list length, $F(2,30)=33.85$, $MSE=.02$, $p<.001$, $\eta^2_p=.69$; the interaction was not significant, $F(2,30)=.10$, $MSE=.009$, $p=.90$, $\eta^2_p=.01$. As shown in Figure 8, the effect of legal list condition was robust for each of the three list lengths. Finally, we determined to what extent incorrect responses in the illegal list condition could have been influenced by the list chunks learned during the incidental learning condition, by determining the number of recalled digit segments larger than 2 digits that were part of the lists to be learned. We obtained no evidence that incorrect responses were influenced by recall of partial chunks stemming from the lists to be learned. Only 2 out of the 16 participants produced erroneous responses for illegal lists that matched sub-sequences of the legal lists that were larger than two digits. In these few instances the extent of overlap was minimal: one participant produced one 3-digit segment that coincided with a segment embedded in one of the legal lists, and the other participant produced a 4-digit segment that coincided with a segment of a legal list.

< INSERT FIGURES 7 AND 8 ABOUT HERE >

Discussion

These results indicate that incidental learning of digit sequences is possible, when the information to be learned relies on list-level information rather than on item-co-occurrence regularities. This is particularly remarkable given that the structure of the learning material contained very little regularities concerning pairwise co-occurrences: Of the 72 possible digit pairs (excluding repetitions), 63 pairs were used, and they were all used about equally often. In contrast, in the preceding experiments only 18 syllable pairs had to be learned. The strong learning effect for digit sequences in the present experiment, combined with the absence of such learning in the preceding experiments, implies that incidental learning of regularities in digit sequences cannot arise from acquisition of information about pairwise co-occurrences.

Whereas learning of pairwise co-occurrences was very difficult in Experiment 4, learning of higher-order regularities, spanning more than two successive elements was facilitated. Because every pair of digits occurred in only one list, each pair, with the exception of the final pair in a list, was perfectly predictive of the following pair. In fact, each pair at the beginning of a list perfectly predicted the remainder of the list. In contrast, the sequences used in the preceding experiments contained little higher-order structure – whereas individual elements were probabilistically predictive of the following element, pairs did not predict anything beyond individual elements. Thus, whereas the sequential structures of the preceding experiments afforded learning of pairwise co-occurrences but not higher-order regularities, the structure of Experiment 4 afforded the learning of higher-order regularities, up to the level of lists of 7-9 elements, but afforded little learning of pairwise relations.

The finding that digit sequences were incidentally learned very well in Experiment 4, whereas they were not learned at all in the preceding experiments, strongly implies that digit sequences are learned incidentally by a mechanism that acquires higher-order regularities without having to build on pairwise regularities of successive elements. Chunking of whole lists into unified representations is such a mechanism. Chunking of whole lists has been proposed to underlie the Hebb effect in serial recall (Burgess & Hitch, 2006). The most parsimonious explanation of the findings from Experiment 4, together with earlier findings on the Hebb effect, is that both effects are generated by the same mechanism. If we accept this explanation, it implies an expansion of the scope of the mechanism underlying the Hebb effect. Different from the Hebb paradigm, participants in our experiment did not encode the repeated lists for immediate serial recall. They were not even asked to attend to these lists. Moreover, whereas in experiments on the Hebb effect participants usually learn a single list, here they learned nine lists in parallel. It seems that there is a learning mechanism for acquiring chunks that is more powerful than previously revealed by the Hebb paradigm.

General Discussion

The four experiments in this article, together with the experiments by Majerus et al. (2004), establish three observations: First, incidental learning of sublist-level, item-co-occurrence regularities in a continuous auditory verbal sequence of items supports immediate serial recall of lists composed of items that follow the same regularities. Second, this particular instance of an effect of long-term learning on immediate serial recall works for short words and nonwords but not for digits. Third, incidental learning of list-level regularities, which requires chunking of whole lists (or at least of sub-segments longer than pairs), provides strong support for immediate serial recall of digit lists. These findings have important implications for our understanding of the influence of long-term memory on performance in verbal STM.

Learning of item co-occurrence regularities from continuous streams of verbal items, as demonstrated in our Experiments 2 and 3 and in Majerus et al. (2004), cannot easily be explained by gradual strengthening of item-to-position associations, because items are not consistently repeated in particular serial positions. It can also not be explained by the formation of unified list representations because the sequence is not segmented into lists. Even if participants spontaneously segmented the sequence into strings of the length of the later memory lists, these segments would rarely match a legal memory list in its entirety, so that there would rarely be a sufficiently large match between a memory list and a learned segment in LTM. We are left with two possibilities for explaining what is learned from continuous streams governed by an artificial grammar; they correspond to the two learning mechanisms that have been discussed for learning of regularities in continuous sequences (Perruchet & Pacton, 2006). Either people learn associations between successive items that capture their transition probabilities, or they segment the sequence into ad-hoc chunks of *n* successive items (with *n* ranging from 2 to about 4 or 5) and learn the relative frequencies of these chunks, thereby capturing the different probabilities of pair-wise transitions (and possibly of longer-range dependencies). Both learning mechanisms capture regularities in the transitions between successive items regardless of their position in a list relative to the beginning or the end.

The Hebb effect, in contrast, does depend on the repetition of items in the same positions relative to the beginning of a memory list, as demonstrated by Hitch et al. (2005). They showed that learning in a Hebb paradigm did not transfer to lists that maintained the item-to-item associations of the learned list but shifted them to new serial positions. Burgess and Hitch (2006) as well as Grossberg and Stone (1986) considered that repeated encoding of the same list into short-term memory contributes to the gradual strengthening of a unified representation of that list's sequence in long-term memory. These unified long-term memory representations would assist recall of a new list to the degree that the new list is similar to the list represented in long-term memory, so that the latter is retrieved during encoding of the former. We hypothesized that this learning mechanism would also operate incidentally on sequences of verbal items, if the sequences contained repeated lists, with clearly marked beginnings and ends.

In Experiment 4, we provide direct evidence for this hypothesis, by showing a capacity in human beings to learn repeated digit sequences as unified chunks. Given that

Experiments 1 and 2 have shown that sublist-level chunks based on pairwise item-co-occurrences probabilities cannot be learned for digit sequences, we can interpret the learning effect in Experiment 4 as reflecting learning of higher-order regularities, that is, the acquisition of chunks that unify the whole repeated lists, or at least segments of these lists longer than two digits. This conclusion is further supported by the fact that there was no evidence for transfer of partial segment knowledge in recall errors for illegal digit sequences in Experiment 4, in contrast to Experiment 2, where we found transfer of pairwise co-occurrence regularities to recall errors for nonword lists.

Thus, the present study directly shows the existence of two types of learning that both serve the acquisition of long-term knowledge of serial order, and both assist immediate serial recall. One type, which we might call Hebb learning, builds and gradually strengthens unified representations of short lists, as proposed by Burgess and Hitch (2006); this type underlies the learning effect in Hebb learning paradigms using digit sequences, because the present study clearly shows that this is the only possible type of learning for digit sequences. The other type, often referred to as statistical learning (Perruchet & Pacton, 2006), acquires knowledge of transition probabilities or co-occurrence probabilities of successive items in sequences of arbitrary length. The present study shows that this type of probabilistic learning works for words and nonwords, but not for digits; other studies have demonstrated statistical learning for a variety of other materials shapes (Fiser & Aslin, 2002) and tones (Creel, Newport, & Aslin, 2004).

Although the present study shows that incidental learning for digit sequences and word/nonword sequences stem from different sources and types of knowledge (deterministic, list-level knowledge versus probabilistic knowledge about item-co-occurrences), this does not necessarily mean that the underlying learning mechanisms as such are fundamentally different. As we have noted in the Introduction, digits are linguistic units that we frequently experience in arbitrary combinations. If we continuously learn about the transition probabilities between linguistic units, statistical learning of digit-to-digit transitions might have saturated for young adults. Saturation could be reached if, for instance, learning rests on implicit counting of frequencies of chunks, each chunk representing a short sequence (i.e., pair, triplet, …) of adjacent elements (Perruchet & Vinter, 1998): After experiencing several ten thousands of instances of any possible pair of digits, adding a few hundred more to some of them but not others in an experiment like ours makes very little difference for the overall relative frequencies of these pairs. In contrast, nonwords have never been experienced before, and even the elements of a small set of frequent words, as used in Experiment 3, have rarely been experienced in immediate succession (for instance, how often has a French speaker experienced the pair “chou-lit”? How often has an English speaker experienced the sequence “cabbage-bed”?). Therefore, learning is arguably still far from saturation for pairs of words in our set, so that additional experimentally induced learning still makes a difference for the memory strength of learned compared to non-learned pairs.

Statistical learning could be assumed to rest on chunks reflecting relatively small strings of units (i.e., mostly pairs and triplets). Generation and gradual strengthening of such small chunks is feasible in continuous streams of input because the stream can be initially parsed into chunks in any arbitrary way, or in multiple ways in parallel, without creating an enormous set of different chunks (Perruchet & Vinter, 1998). In contrast, learning for digit sequences rests on chunks that represent whole memory lists of 7 to 9 items, or at least large parts of such memory lists, in a unified fashion. Sequences of about 7 to 9 successive digits are still novel enough, even after 20 years of experience with numbers, for some additional exposure to make a difference in their relative frequency. Generation of such large chunks is feasible without explosion of the number of chunks, if the input is already parsed into lists, as in the Hebb learning paradigm and in our Experiment 4. With digits as material, learning cannot be observed in an experiment using un-parsed continuous input with only statistical regularity, as in our experiments, because the structure in the input affords only learning of small chunks of short sequences, which have all been over-learned for digits already. In contrast, as demonstrated by Experiment 4, larger chunks of whole memory lists can be learned when they are repeated as a whole, and are identifiable by segmentation of the input.

ACKNOWLEDGMENTS

This work was facilitated by a grant from the Fonds de la Recherche Scientifique – FNRS (Belgium), grant “Crédit aux Chercheurs 2009-2010 1.5.056.10”, to the first author, and a grant from the Economic and Social Research Council (ESRC), grant RES-062-23-1199 to the second author. We thank Franziska Fischer, Pauline Lanquetin, and Mahaut Sermonard for help with collecting the data.

REFERENCES

Baddeley, A.D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173.

Botvinivk, M. (2005). Effects of domain-specific knowledge on memory for serial order. *Cognition, 97,* 135-151.

Botvinick, M., & Bylsma, L. M. (2005). Regularization in short-term memory for serial order. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31,* 351-358.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436.

Burgess, N. & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31,* 429-460.

Burgess, N. & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551-581.

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.

Clopper, C., & Pisoni, D.B. (2005). Perception of dialect variation. In D. Pisoni, & R.E. Remez (Eds.), The handbook of speech perception (pp. 312-337). Oxford, UK: Blackwell Publishing.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1119-1130.

Cumming, N. (2001). *The Hebb effect: Investigating long-term learning from short-term memory.* Unpublished doctoral thesis, Cambridge University, Cambridge.

Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory*, 43-63.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Human Learning and Memory, 25,* 84-95.

Glover, S., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review, 11*, 791-806.

Grossberg, S., & Stone, G. (1986). Neuronal dynamics of attention switching and temporal order information in short term memory. *Memory & Cognition, 14,* 451-468.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56A,* 1213–1236.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 458-467.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J.E.Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37-46). New York: Oxford University Press.

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

Hitch, G., Flude, B., & Burgess, N. (2009). Slave to the rhythm: experimental tests of a model for verbal short-term memory and long term sequence learning. *Journal of Memory and Language, 61,* 97-111.

Johnson, K. (2005). Speaker normalization in speech perception. In D. Pisoni, & R.E. Remez (Eds.), The handbook of speech perception (pp. 363-389). Oxford, UK: Blackwell Publishing.

Karpicke, J.D., & Pisoni, D.B. (2004). Using immediate memory span to measure implicit learning. *Memory & Cognition, 32,* 956-964.

Majerus, S., & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language, 64,* 181-197.

Majerus, S. & Van der Linden, M. (2003). The development of long-term memory effects on verbal short-term memory : A replication study. *British Journal of Developmental Psychology, 21,* 303-310.

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language, 51,* 297-306.

Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language, 43,* 240–274.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers, 36,* 516-524.

Oberauer, K. (2009). Design for a working memory. *The Psychology of Learning and Motivation, 51*, 45-100.

Oberauer, K. & Meyer, N. (2009). The contributions of encoding, retention, and recall to the Hebb effect. *Memory, 17,* 774-781.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233-238.

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

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 29–33.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35,* 606-621.

Saffran, J. R., Newport, E. L., Aslin, R. N., & Tunick, R. A. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8,* 101-105.

Thorn, A. S. & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 729-735.

Tubach, J. L. & Boe, L. J. (1990). *Un corpus de transcription phonétique*. France: Telecom.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*, 779-804.

Table 1. Stimulus characteristics for stimuli used in Experiment 2.

Stimuli	Phonetic transcription	Syllabic structure	Phonotactic frequency (mean)
Digits			
un	[œ̃]	V	1635
deux	[dø]	CV	258
trios	[tʀwa]	CCvV	2001
quatre	[katʀ]	CVCC	1752
cinq	[sɛ̃k]	CVC	275
six	[sis]	CVC	1800
sept	[sɛt]	CVC	1568
huit	[ɥit]	vVC	1258
neuf	[nœf]	CVC	339
Nonwords			
eu	[ø]	V	1367
nan	[nɑ̃]	CV	428
stoi	[stwa]	CCvV	1916
daste	[dast]	CVCC	1356
linde	[lɛ̃d]	CVC	341
leul	[lœl]	CVC	1973
lère	[lɛʀ]	CVC	1892
yeure	[jœʀ]	vVC	811
zim	[zim]	CVC	436

Table 2. Stimulus characteristics for stimuli used in Experiment 3.

Stimuli	Phonetic transcription	Syllabic structure	Phonotactic frequency (mean)
Words			
dent	[dɑ̃]	CV	1043
choux	[ʃu]	CV	2
feu	[fø]	CV	8
veau	[vo]	CV	118
pain	[pɛ̃]	CV	29
riz	[ʀi]	CV	1479
chat	[ʃa]	CV	119
nez	[ne]	CV	1231
lit	[li]	CV	1771
Nonwords			
na	[na]	CV	1102
chon	[ʃɔ̃]	CV	1
zin	[zɛ̃]	CV	138
bi	[bi]	CV	206
leuh	[lø]	CV	19
ko	[ko]	CV	1305
chu	[ʃy]	CV	4
ti	[ti]	CV	1447
reu	[rœ]	CV	1886

Table 3. Example of digit lists to be learned in Experiment 4.

List length	List
7	4 2 6 5 8 7 1
7	9 3 1 6 8 5 4
7	2 7 9 8 1 4 6
8	5 2 9 6 7 3 8 4
8	7 2 5 6 3 9 4 8
8	1 5 3 4 7 6 1 7
9	6 4 9 2 8 3 5 1 8
9	8 9 1 2 4 6 7 5 9
9	3 2 1 9 7 4 5 7 8

FIGURE LEGENDS

Figure 1. Examples of artificial grammars used in Experiments 1, 2 and 3.

Figure 2. Means and standard errors for recall performance in the post-learning short-term memory task, as a function of list condition and serial position in Experiment 1 (digit co-occurrence learning).

Figure 3. Means and standard errors for recall performance in the post-learning short-term memory task, as a function of list condition and item type in Experiment 2 (digit and nonword co-occurrence learning).

Figure 4. Means and standard errors for recall performance in the post-learning short-term memory task, as a function of list condition, item type and list length in Experiment 2 (4a: digit co-occurrence learning; 4b: nonword co-occurrence learning).

Figure 5. Means and standard errors for recall performance in the post-learning short-term memory task, as a function of list condition and item type in Experiment 3 (word and nonword co-occurrence learning).

Figure 6. Means and standard errors for recall performance in the post-learning short-term memory task, as a function of list condition, item type and list length in Experiment 3 (6a: word co-occurrence learning; 6b: nonword co-occurrence learning).

Figure 7. Means and standard errors for recall performance in the post-learning short-term memory task in Experiment 4 (digit chunk learning).

Figure 8. Means and standard errors for recall performance in the post-learning short-term memory task as a function of list length in Experiment 4 (digit chunk learning).

Figure 1.

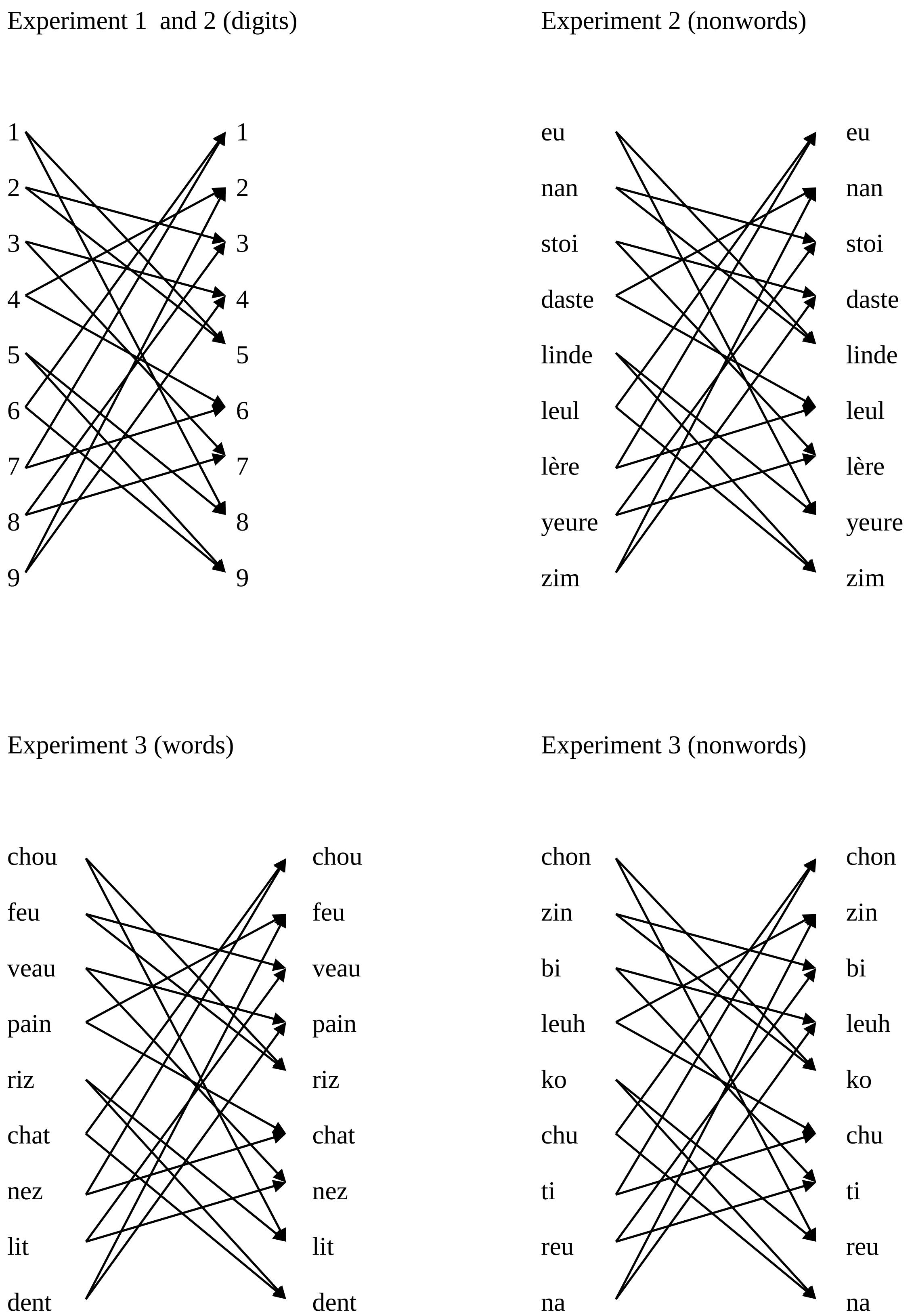
Experiment 1 and 2 (digits)
1 2 3 4 5 6 7 8 9
1 2 3 4 5 6 7 8 9
Experiment 2 (nonwords)
eu nan stoi daste linde leul lère yeure zim
eu nan stoi daste linde leul lère yeure zim
Experiment 3 (words)
chou feu veau pain riz chat nez lit dent
chou feu veau pain riz chat nez lit dent
Experiment 3 (nonwords)
chon zin bi leuh ko chu ti reu na
chon zin bi leuh ko chu ti reu na

Figure 2.

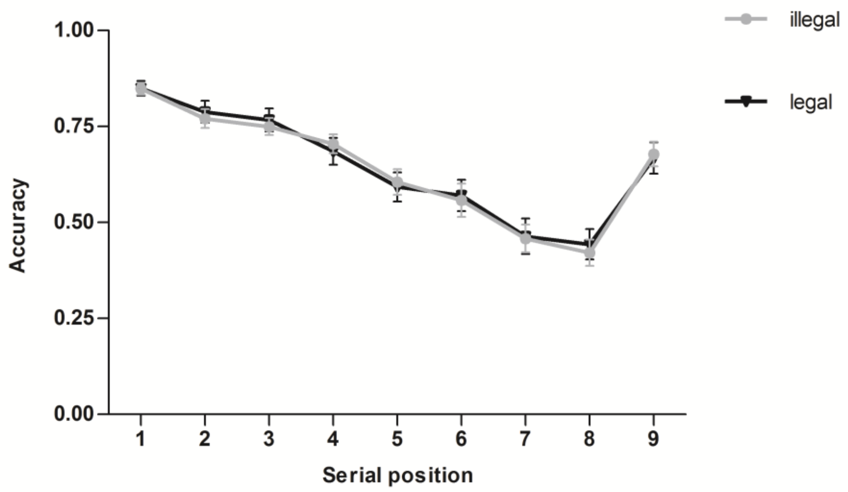
illegal
legal
1.00
0.75
0.50
0.25
0.00
Accuracy
1
2
3
4
5
6
7
8
9
Serial position

Figure 3

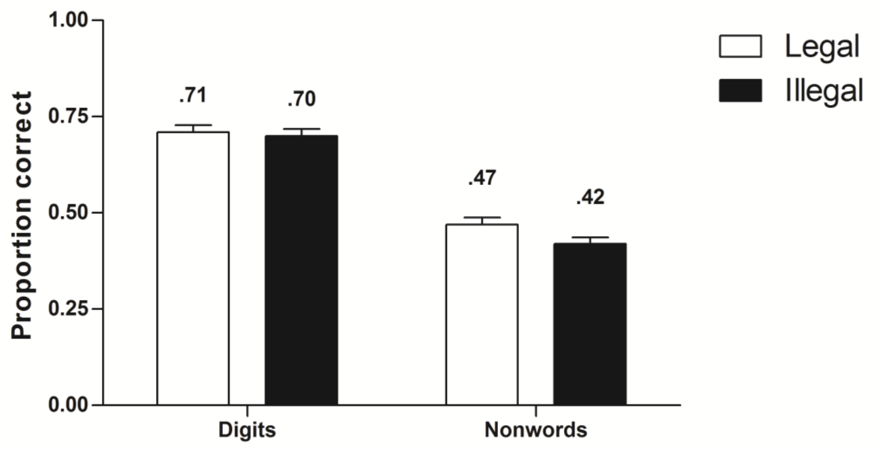
1.00
0.75
0.50
0.25
0.00
Proportion correct
.71
.70
.47
.42
Digits
Nonwords
Legal
Illegal

Figure 4.

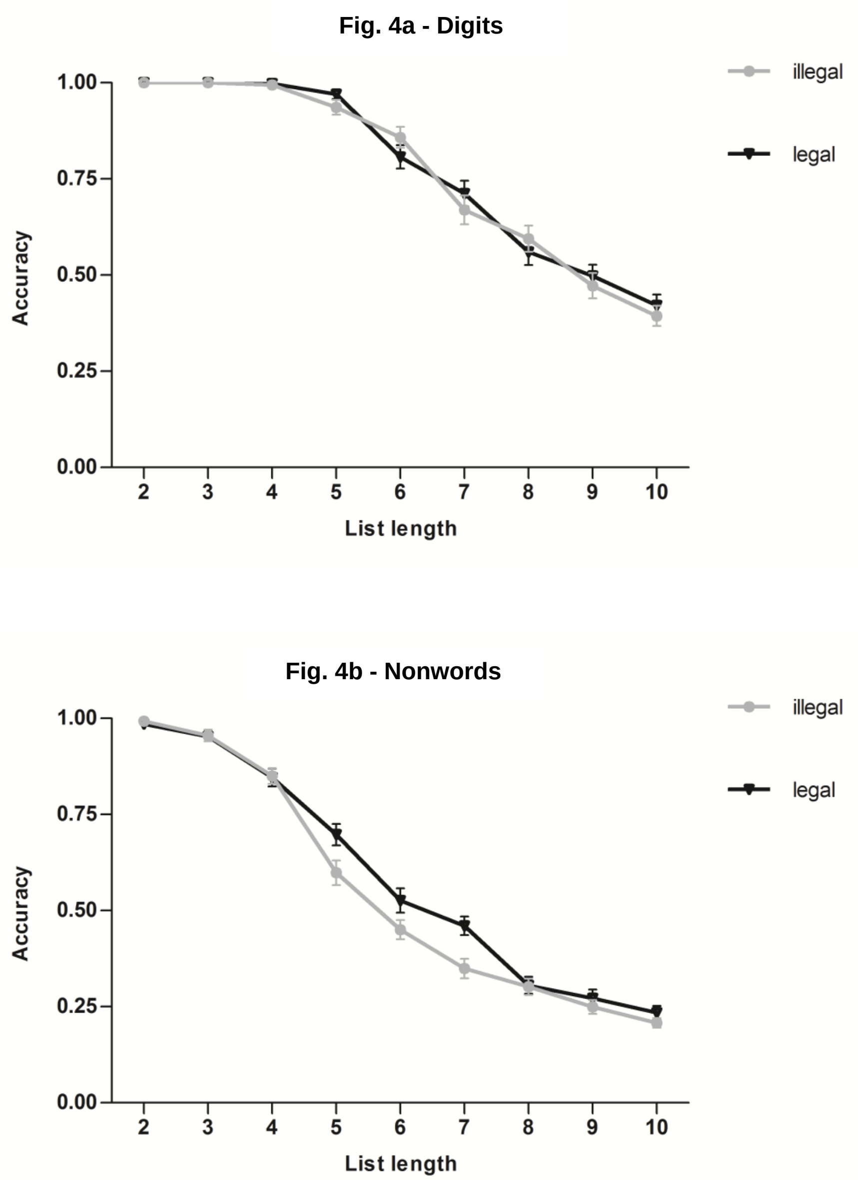
Fig. 4a - Digits
illegal
legal
Accuracy
1.00
0.75
0.50
0.25
0.00
2
3
4
5
6
7
8
9
10
List length
Fig. 4b - Nonwords
illegal
legal
Accuracy
1.00
0.75
0.50
0.25
0.00
2
3
4
5
6
7
8
9
10
List length

Figure 5.

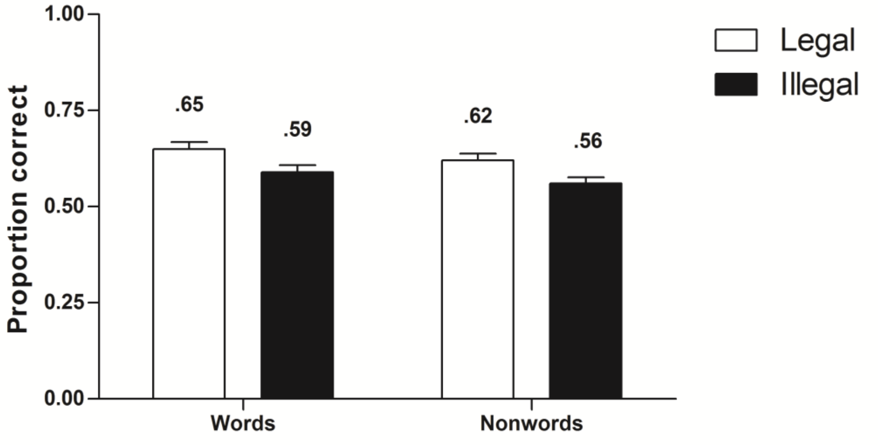
1.00
0.75
0.50
0.25
0.00
Proportion correct
.65
.59
.62
.56
Words
Nonwords
Legal
Illegal

Figure 6.

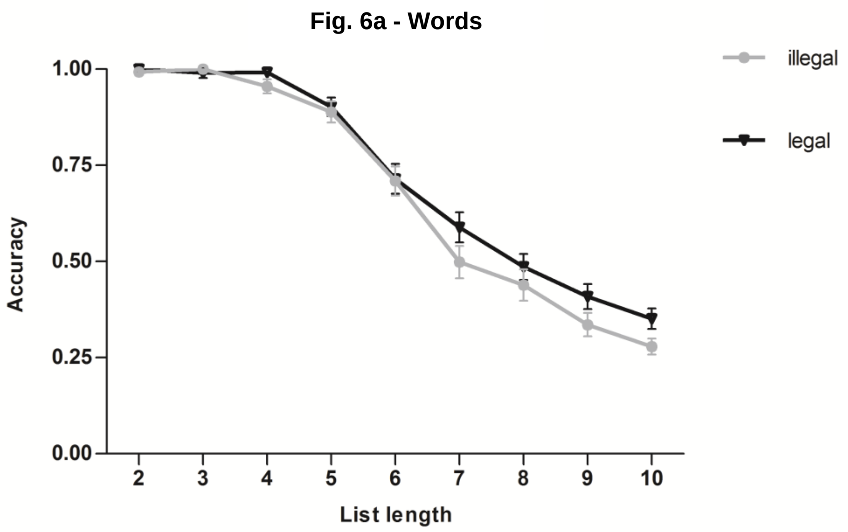
Fig. 6a - Words
illegal
legal
Accuracy
1.00
0.75
0.50
0.25
0.00
2
3
4
5
6
7
8
9
10
List length

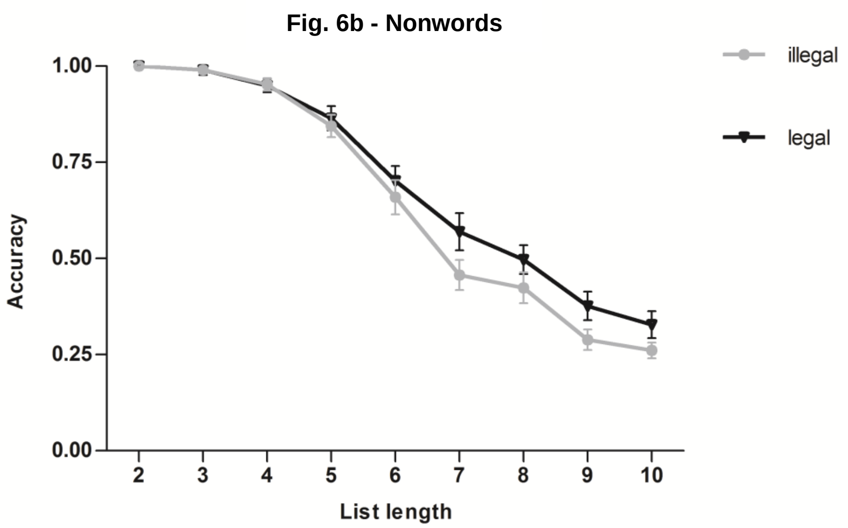
Fig. 6b - Nonwords
illegal
legal
Accuracy
1.00
0.75
0.50
0.25
0.00
2
3
4
5
6
7
8
9
10
List length

Figure 7.

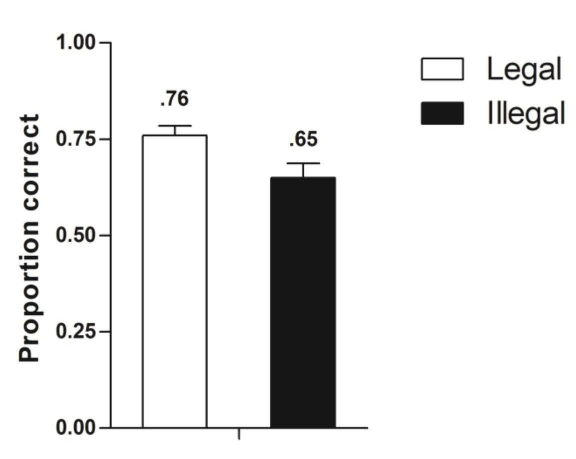

Figure 8.