Acta Psychologica 157 (2015) 195–199

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Is an attention-based associative account of adjacent and nonadjacent dependency learning valid?☆

Sébastien Pacton [a,b,c,*], Amélie Sobaco [a,b,c], Pierre Perruchet [d,e]

[a] Psychology Department, Université Paris Descartes, Paris, France
[b] Sorbonne Cité France, France
[c] INSERM, France
[d] Psychology Department, Université de Bourgogne, France
[e] CNRS, Dijon, France

ARTICLE INFO

Article history:
Received 20 January 2014
Received in revised form 7 March 2015
Accepted 9 March 2015
Available online xxxx

Keywords:
Statistical learning
Implicit learning
Associative learning
Nonadjacent dependencies
Attention

ABSTRACT

Pacton and Perruchet (2008) reported that participants who were asked to process adjacent elements located within a sequence of digits learned adjacent dependencies but did not learn nonadjacent dependencies and conversely, participants who were asked to process nonadjacent digits learned nonadjacent dependencies but did not learn adjacent dependencies. In the present study, we showed that when participants were simply asked to read aloud the same sequences of digits, a task demand that did not require the intentional processing of specific elements as in standard statistical learning tasks, only adjacent dependencies were learned. The very same pattern was observed when digits were replaced by syllables. These results show that the perfect symmetry found in Pacton and Perruchet was not due to the fact that the processing of digits is less sensitive to their distance than the processing of syllables, tones, or visual shapes used in most statistical learning tasks. Moreover, the present results, completed with a reanalysis of the data collected in Pacton and Perruchet (2008), demonstrate that participants are highly sensitive to violations involving the spacing between paired elements. Overall, these results are consistent with the Pacton and Perruchet's single-process account of adjacent and nonadjacent dependencies, in which the joint attentional processing of the two events is a necessary and sufficient condition for learning the relation between them, irrespective of their distance. However, this account should be completed to encompass the notion that the presence or absence of an intermediate event is an intrinsic component of the representation of an association.

1. Introduction

Nonadjacent dependencies refer to the cases where a statistical association exists between two events that are not immediately contiguous in space or time, due to the occurrence of one or several intervening events. This pattern is quite frequent in natural languages (e.g., between auxiliaries and inflectional morphemes, as in "is writing", irrespective of the verb stem). Nonadjacent dependencies are also present in other domains of high-level knowledge such as music. In Western music, for instance, two structurally important tones are often separated by other, less important tones (the ornaments). If the nonadjacent dependency between the two structurally important tones was not captured by the listener, the musical structure would not be perceived. Likewise, capturing the relationships between distant objects seems essential. As claimed by Turk-Browne, Jungé, and Scholl (2005), "People are constantly bombarded with noise in space and time that needs to be segregated in order to extract a coherent representation of the world, and people rarely encounter a sequence of relevant stimuli without any interruptions" (p. 562).

There is increasing evidence that the learning of nonadjacent dependencies is possible, but only under specific conditions (for a review: Perruchet, Poulin-Charronnat, & Pacton, 2012). Let us refer to a nonadjacent structure as AXC, where A and C stand for the associated events and *X* stands for a variable event, statistically independent from both A and C. A non-exhaustive list of conditions includes: (1) the high level of variability of the *X* event (Gómez, 2002, 2006; Onnis, Christiansen, Chater, & Gómez, 2003). (2) The high level of similarity between A and C. Similarity can be assessed on an acoustic dimension. Using musical tone sequences, Creel, Newport, and Aslin (2004) showed that nonadjacent dependencies were not acquired when all elements differed equally one another, whereas learning was successful when A and C were similar in pitch or timbre, and different from *X*. Likewise, Onnis, Monaghan, Richmond, and Chater (2005) showed that no

☆ This work has been supported by the Institut National de la Santé et de la Recherche Médicale (INSERM, UMR S 894), the Centre National de la Recherche Scientifique (CNRS, UMR 5022), the Université Paris Descartes, Sorbonne Paris Cité, the Université de Bourgogne, and the Institut Universtaire de France.

* Corresponding author at: Université Paris Descartes, Laboratoire Mémoire et Cognition, 71 av E. Vaillant, 92100 Boulogne-Billancourt, France.
E-mail address: sebastien.pacton@parisdescartes.fr (S. Pacton).

http://dx.doi.org/10.1016/j.actpsy.2015.03.002
0001-6918/

learning was obtained without some degree of phonological similarity between A and C syllables. (3) The membership of A and C to the same category, itself differing from the category of *X*. For instance, Newport and Aslin (2004) failed to observe learning with nonadjacent syllables (i.e., A, *X*, and C were syllables), whereas learning occurred when A and C were consonants and *X* was a vowel and, conversely, when A and C were vowels and *X* was a consonant. (4) The introduction of short pauses between the A*X*C sequences during the familiarization phase (Peña, Bonatti, Nespor, & Mehler, 2002). (5) The occurrence of an earlier training phase during which the to-be-associated pairs have been studied in adjacent conditions. Introducing structural complexity progressively during learning would meet the general learning principle known as the "starting small" effect (Lai & Poletiek, 2011).

Because adjacent dependencies are remarkably easy to learn in a large array of experimental settings, as shown throughout the associative learning literature, the restrictive conditions listed above have led some researchers to claim that learning adjacent and nonadjacent dependencies rely on different processes (e.g., Peña et al., 2002). In contrast with this dual-process view, Pacton and Perruchet (2008) proposed to account for both adjacent and nonadjacent dependencies within an integrated framework, grounded on the role of attention in associative learning (e.g., Hoffmann & Sebald, 2005; Hsiao & Reber, 1998; Jiménez & Méndez, 1999). Several authors have suggested that associative learning is an automatic process that links together all of the components that are present in the attentional focus at a given point (e.g., Frensch & Miner, 1994; Logan & Etherton, 1994; Perruchet & Vinter, 2002; Stadler, 1995; Treisman & Gelade, 1980) and that the joint attention given to a pair of events would be a necessary, but also a sufficient condition for the emergence of associative learning and memory. It is reasonable to postulate that, *by default*, the mental content composing the attentional focus at a given moment has a high chance of representing events that are close on spatial and/or temporal dimensions in the environment. This would account for the overt precedence to the condition of contiguity in the conventional associative learning literature. However, crucially, the attentional content may also encompass events that are not adjacent in the environment, although only if some specific conditions lead to pay joint attention to those events. In this regard, it is noteworthy that all the five experimental conditions listed above as beneficial for learning nonadjacent dependencies can be viewed as facilitating the attentional processing of the relevant events (i.e., A and C). Thus an attention-based account is seemingly able to retrospectively account for earlier data. More compellingly, this account is also able to generate testable predictions.

A straightforward prediction of the Pacton and Perruchet (2008) integrated framework is that conditions ensuring the very same amount of attentional processing for both adjacent and nonadjacent dependencies would result in a perfect symmetry between the two forms of learning. The authors reported a set of five experiments in which attentional processing was manipulated through the instructions given to the learners. Participants were faced with a set of problems, each comprising a sequence of digits embedding both adjacent and nonadjacent regularities. They were asked to perform an arithmetic task that involved the joint processing of two selected digits. These two digits were adjacent for a first group of participants, and nonadjacent for a second group. A subsequent recognition test explored how well participants from the two groups learned both adjacent and nonadjacent dependencies. The results were clear-cut. Participants who were asked to process adjacent elements learned adjacent dependencies but did not learn nonadjacent dependencies. Much more interestingly, participants who were asked to process nonadjacent elements learned nonadjacent dependencies but did not learn adjacent dependencies. It is noteworthy that the recognition score for nonadjacent dependencies was not significantly lower than the recognition score for adjacent dependencies reached by the participants who focused on adjacent dependencies. Thus, the objective adjacency of the events in the display played no role of its own when the attentional processing of A and C, as prompted by the task demand, was the same for each type of dependency (see also Jahn, 2012, for a replication).

Although Pacton and Perruchet (2008) data provided a strong support for an attention-based, unitary account of adjacent and nonadjacent dependencies, a potential limitation could be that their conclusion was based on experiments relying exclusively on digits as stimuli. Using digits instead of syllables or visual shapes, as commonly exploited in the statistical learning literature, was dictated by the need for creating a task that allows to focus on either adjacent or nonadjacent elements in a meaningful way. An arithmetic task is especially well-suited for this objective because processing nonadjacent digits is what anyone does while performing the most basic arithmetic calculations in real-world settings. However, the downside is that the processing of digits could be quite specific. The perfect symmetry found between the processing of adjacent and nonadjacent dependencies in Pacton and Perruchet could be restricted to this material. In particular, a possibility is that when there is no specific reason to pay joint attention to either adjacent or nonadjacent events, as in most standard tasks of incidental learning, the processing of digits would be insensitive to the contiguity condition which has been shown to be so important throughout the associative learning literature, or at least, less sensitive to the contiguity of the items than the processing of other stimuli such as syllables, tones, or visual shapes. A control condition using standard incidental instructions, such as listening to oral language or tones, or watching visual shapes, was not implemented in Pacton and Perruchet, and as a consequence, there is currently no evidence that the natural processing of digits would exhibit a strong preference for adjacent relationships, as regularly found for other stimuli.

The first objective of the present study was to explore the pattern of performance in a condition using neutral instructions, which was missing in the study of Pacton and Perruchet (2008). Neutral instructions refer here to a task demand that does not require the selective processing of either adjacent or nonadjacent dependencies, as in most incidental learning tasks. Participants were exposed to sequences of digits embedding both adjacent and nonadjacent dependencies, as in Pacton and Perruchet, but they were simply asked to read aloud the items. To examine further whether the processing of digits is endowed with particular properties, half of the participants performed the very same reading task with syllables. Our hypothesis was that under neutral instructions, the usual asymmetry between adjacent and nonadjacent dependency should be found for digits as for syllables, with easier, if not exclusive learning of adjacent dependencies. If this hypothesis turned out to be wrong, then the unified attentional model proposed in Pacton and Perruchet, grounded on the exclusive use of digits as stimuli, should be reconsidered.

Whereas most studies on nonadjacent dependencies deal with the conditions making learning easier, the "what is learned?" issue has not yet been extensively explored. The second objective of the present study was to shed preliminary light on a particular aspect of this issue, namely the status of the intermediate event (*X*) within the A*X*C sequence. In a nutshell, the question is: While learning A*X*C, does the learner simply code that A is followed by C, or is *X* a mandatory component of the learner's representation? The framework of Pacton and Perruchet (2008) does not deal explicitly with this question, but its emphasis on the role of the attention paid to the target stimuli, A and C, suggests that the intermediate event, *X*, must receive no, or only a minimal amount of attentional processing during training.

A recognition test including the correct sequence A*X*C and a distractor like A*X*D has often been used (e.g., Gómez, 2002). However, such a test is inappropriate to investigate whether *X* is a mandatory component of the learner's representation. Indeed, participants could express a preference for A*X*C over A*X*D simply because A and C, contrary to A and D, have formed the mental content of their attentional focus at a given time during the study phase, without having learned whether an intermediate element is located between A and C. In order to address this issue, distractors must include a spacing violation (e.g., A*X*C vs.

ACX). Pacton and Perruchet (2008) used distractors of this kind along with distractors that include an order violation (e.g., AXC vs. CXA). However, they pooled them together for analyses. Thus, the score used by Pacton and Perruchet could be above chance because participants have coded the order of elements A and C (i.e., scores above chance for items with an order violation) even though they have not learned whether an intermediate element is located between A and C (i.e., scores at chance level for items with a spacing violation).

A huge amount of data collected in various subareas of the associative learning literature, such as classical conditioning, paired-associate paradigms, and serial learning, concur to show that the order of the elements in a pair or a sequence is actively coded and remembered. Although these studies investigated the order between adjacent elements, it seems reasonable to expect that the order is also coded in the case of nonadjacent elements and therefore that forced-choice recognition was above chance for items comprising an order violation in the study of Pacton and Perruchet (2008). Analyzing performance on recognition tests with distractors comprising a spacing violation alone is essential to investigate whether *X* is a mandatory component of the learner's representation. However, the comparison of the effect of a spacing violation with the better known effect of an order violation may be especially instructive because the stimuli including order violations can serve as a benchmark in evaluating the effect of spacing violations.

In the present study, as in that of Pacton and Perruchet (2008), the distractors used in the recognition tests are composed of two kinds of violations, namely spacing (e.g., AXC vs. ACX) and order (e.g., AXC vs. CXA). We analyzed forced-choice recognition for each kind of violation for both the new collected data and Pacton and Perruchet's data that we reanalyzed. If learning proceeds from the joint attentional focus on A and C, as proposed in the Pacton and Perruchet's framework, the presence or absence of the unattended event *X* in the test items should be relatively inconsequential. In any case, order violations, which involve the relevant events A and C, should be easier to detect than spacing violations, which involve the unattended event *X*.

2. Experiment

2.1. Method

2.1.1. Participants

Forty-eight undergraduate students, 39 females, aged 17–24 years ($M = 20$ years and 2 months) from the Université Paris Descartes, Paris, France, participated in the experiment in partial fulfillment of a course requirement. All participants were native French speakers. Participants were assigned to one of the four experimental conditions defined below, with $n = 12$ for each group.

2.1.2. Material

The stimuli were digits for half of the participants, and syllables for the other half, everything else being equal. For the sake of concision, this section deals only with the digits. For the syllables, digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9, were simply replaced by the CVC syllables *moc*, *guz*, *fib*, *pem*, *zil*, *lur*, *cad*, *rog*, *dap*, and *bef*, respectively.

The sequences used in the study phase consisted of 20 digits. In every sequence, (a) each of the digits (from 0 to 9) occurred twice, (b) the target pair of adjacent digits occurred twice and (c), the target pair of nonadjacent digits also occurred twice (see Appendix A). The target pairs were identical in all of the sequences presented to a given participant. For half of the participants assigned to the digit condition, the adjacent digits were *64* and the nonadjacent digits were *3X7*, and for the other half, the adjacent digits were *37* and the nonadjacent digits were *6X4*. Within each sequence, the position of the target pairs of digits and the position of the residual digits were determined randomly. A set of 20 sequences was generated for each participant.

Stimuli for the two-choice recognition test consisted of 16 pairs of three-digit strings. Each pair was composed of a legal string and an illegal string. Eight pairs were used to assess the learning of adjacent dependencies. For the participants who were exposed to *64* as the adjacent dependency during the study phase, the legal string was *64X* for four pairs and *X64* for four other pairs. The illegal string resulted from a spacing violation in four pairs *6X4* and from an order violation in four other pairs *46X* or *X46*. Eight other pairs were used to assess the learning of nonadjacent dependencies. The legal string was *3X7* for the eight pairs. The illegal string resulted from a spacing violation in four pairs *37X* or *X37* and from an order violation in four other pairs *7X3*. For the participants who were exposed to the counterbalanced combination during the study phase (i.e., *37* as the adjacent dependency and *6X4* as the nonadjacent dependency), the test pairs were built following the same principles. The digit *X* was selected randomly among digits 0, 1, 2, 5, 8, and 9. This digit was the same for the legal string and for the illegal string of a pair (e.g., 640–604). All the stimuli were displayed in Chicago font, size 24, with PsyScope for Mac OS X (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macbook (Apple Inc., Cupertino, CA) 13.3-in. (39.1-cm) wide-screen display.

2.1.3. Procedure

The participants were tested individually in a sound-attenuated room. The experiment consisted of a study phase and a test phase. Before the study phase, participants were instructed that sequences of digits (or syllables) will be presented on the screen of the computer and that their task was simply to read aloud the whole sequence of items from left to right. The 20 study sequences were presented one at a time in random order, and participants were told to press the space bar to skip from a sequence to the next one.

Once the sequences of the study phase were read, participants performed the forced-choice recognition test. The 16 pairs of three-digit strings (or syllables) were presented one at a time at the center of the screen, with the two items of each pair being separated with a slash mark. The order of the items within a pair (left or right) and the order of the pairs in the sequence were randomized for each participant. For each test pair, participants were instructed to select which of the two strings was seen in the study phase of the experiment. They were told that they had to make a choice even if they felt that they were guessing. They were to press the Q key of the (French) AZERTY keyboard if they felt that the correct answer was the string on the left and the M key if they felt that the correct answer was the string on the right. The screen was cleared immediately after the participant's keystroke, and there was a 2-s delay before the presentation of the following pair.

2.2. Results

Fig. 1 shows the mean percentage of correct responses on the recognition test for items with adjacent dependencies and for items with

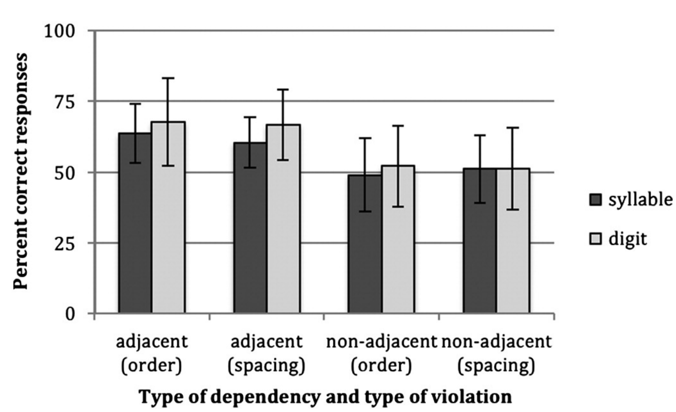

Fig. 1. Mean percentage of correct responses (and standard deviation of the mean) as a function of the type of dependency and the type of violation in the illegal string.

nonadjacent dependencies, instantiated with digits or syllables, according to whether the illegal string included a spacing or an order violation.

An ANOVA was performed with the type of materials (digits vs. syllables) as a between-subject variable, and the type of dependency (adjacent vs. nonadjacent) and the type of violation (order vs. spacing) as within-subject variables. There was a main effect of the type of dependency, with better performance for items involving adjacent dependencies than for items involving nonadjacent dependencies (64.58% vs. 50.78%, $F(1, 46) = 12.98$, $p = .001$, $\eta_p^2 = .22$). The other main effects were not significant and there was no reliable interaction between variables ($ps > .35$).

For the items involving adjacent dependencies, t tests revealed significantly above chance level performance for any combination of materials and type of violation (digit and order violation, $t(23) = 2.81$, $p = .01$; digit and spacing violation, $t(23) = 3.24$, $p = .004$; syllable and order violation, $t(23) = 3.19$, $p = .004$; and syllable and spacing violation, $t(23) = 2.85$, $p = .009$). For the items involving nonadjacent dependencies, performance never differed from chance level (all $ps > .72$).

The percentages of correct responses shown in Fig. 1 were very similar to those obtained by Pacton and Perruchet (2008, Experiment 1) for participants who focused on adjacent elements (adjacent dependencies and order violations: $M = 69.79\%$, $SD = 25.11\%$; adjacent dependencies and spacing violations: $M = 73.70\%$, $SD = 25.23\%$; nonadjacent dependencies and order violations, $M = 51.56\%$, $SD = 25.60\%$; nonadjacent dependencies and spacing violations: $M = 50.00\%$, $SD = 27.86\%$). An ANOVA combining these values with those of the present experiment, with the experimental conditions (Pacton & Perruchet, 2008, Experiment 1 for participants who focused on adjacent elements; the present experiment with digits; and the present experiment with syllables) as a between-subject variable, and the dependency type (adjacent vs. nonadjacent) and the type of violation (order vs. spacing) as within-subject variables, confirmed the main effect of dependency type, with better scores for the items involving adjacent dependencies than for the items involving nonadjacent dependencies ($F(1, 141) = 31.96$, $MSE = 1.26$, $p < .001$, $\eta_p^2 = .19$). Crucially, there was no effect of the experimental condition, no effect of the type of violation, and no interaction (all $p_s > .25$), suggesting that under neutral instructions, participants lend attention to adjacent dependencies as a default.

Performances of participants who focused on nonadjacent elements in Pacton and Perruchet (2008, Experiment 1) were obviously different from those of the present experiment, given they learned exclusively nonadjacent dependencies, as described in the earlier paper. However, reanalyzing the data as a function of the types of violations included in the test shows that performances were above chance for both order ($M = 66.41\%$, $SD = 29.43$, $t(95) = 5.46$, $p < .001$) and spacing violations ($M = 75.26\%$, $SD = 26.03$, $t(95) = 9.51$, $p < .001$). Unlike what is observed for adjacent dependencies, the scores for spacing violations turned out to be significantly better than for order violations ($t(95) = 2.23$, $p = .028$).

3. Discussion

Pacton and Perruchet (2008) proposed an attention-based single-process account of adjacent and nonadjacent dependency learning. They showed that adjacent and nonadjacent dependencies were learned to the same extent when the current focus of attention is oriented towards the relevant events by the task demand. However, this account is based on experiments relying on digits as stimuli, and it is possible that the processing of digits differs from the processing of syllables or pictures with regard to the effect of contiguity. Indeed, by contrast with other stimuli, processing nonadjacent digits is common in real-world tasks, such as arithmetic calculation. A possibility is that in conditions where attention is not drawn away from adjacent events due to either the nature of the materials or the task instructions, the processing of digits would be relatively insensitive to the spatial or temporal adjacency of the items. This neutral condition was not implemented in Pacton and Perruchet, and as a consequence, this possibility, which would be devastating for their framework, cannot be ruled out.

In the present study, which used Pacton and Perruchet's (2008) materials with tasks instructions that did not require focusing on one type of relations at the expense of the other, only adjacent dependencies were learned. Moreover, the very same pattern was observed irrespective of whether digits or syllables were used as stimuli. These results strongly suggest that digits do not differ from other stimuli regarding their sensitivity to the condition of contiguity. In addition, the recognition scores for adjacent events of the participants who simply read the whole sequence of digits from left to right were not statistically lower than those collected in conditions where participants were asked to process adjacent dependencies in Pacton and Perruchet's study. This is consistent with the hypothesis that Pacton and Perruchet proposed to account for the indisputable preeminence of adjacent dependency learning in most experimental settings. This hypothesis stipulated that by default, it is mostly, if not only, adjacent events that are attentionally processed.

The second objective of the present study was to shed some light on the question of what is learned. The distractors used in the recognition tests were composed of two kinds of violations, namely spacing (e.g., A*X*C when the correct item was AC*X*) and order (e.g., CA*X* when the correct item was AC*X*). The results were straightforward: insofar as adjacent dependencies are concerned, spacing violations were detected as well as order violations. This effect was also observed in the condition in Pacton and Perruchet (2008, Experiment 1) in which participants were asked to deal with adjacent dependencies. In addition, a reanalysis of the data obtained in the condition in Pacton and Perruchet (2008) drawing attention on nonadjacent dependencies shows that both kinds of violations were learned, with even better performance with spacing violations than order violations.

This pattern of results suggests that what is learned is not only “A is followed by C”, but more precisely, either “A is immediately followed by C”, or “A is followed by C after some intervening item *X*”. In other words, the presence or absence of an intervening item seems to be an intrinsic component of the representation of the AC association. This conclusion departs from the prevalent conceptions, in which the intervening item tends to be construed more as a difficulty in the learning of AC than as a constitutive part of the AC relation. For instance, the fact that an earlier training phase during which the to-be-associated pairs have been studied in adjacent conditions facilitates the subsequent learning of nonadjacent dependencies (Lai & Poletiek, 2011) seems to be more consistent with the view that the intervening event is only a source of difficulty when learning the relations between two events, with a continuous transition from AC to A*X*C being possible by way of transfer or generalization.

More importantly for our concern, this pattern of results was not expected by the Pacton and Perruchet's (2008) attentional model of learning. The claim that only the joint attention paid to the to-be-associated items is relevant tacitly implies that the presence or absence of an intervening event is incidental. To comply with the present data, a model of adjacent and nonadjacent dependency learning must incorporate the idea that the whole structure of the stimuli is coded, including the possible presence of variable intervening events. A priori, there are two main ways in which this objective may be fulfilled. The crucial role of attention may be left unchallenged, with the corollary that attentional coding must include the intervening event whenever this event is present. An alternative interpretation would be that attention focuses exclusively on the to-be-associated events, as originally argued, and that possible intervening events are automatically processed. The literature on contextual cueing (e.g., Jiang and Chun, 2003) shows that contextual regularities are implicitly coded and allow an optimization of subsequent attentional processing. Studies on nonadjacent dependencies learning could take profit of this conceptualization, by considering the intervening event as an automatically processed contextual element. Irrespective of whether the processing of intervening events is construed as an intrinsic component of attentional coding or as an automatic by-product of the attentional processing of the neighboring elements, how the presence of intervening events can be coded in memory raises an important challenge for further research.

Appendix A

An example of 20 sequences of 20 digits used in the study phase and an example of 16 items used in the recognition test. In this example, the adjacent dependency is 64 and the nonadjacent dependency is 3X7.

Study	Test			
	Adjacent dependency		Nonadjacent dependency	
	Order violation	Spacing violation	Order violation	Spacing violation
51950643873970122648	648–468	640–604	317–713	307–037
90864852307327156419	642–462	645–654	397–793	317–137
26406485253973179801	964–946	264–624	357–753	387–378
26431718503570642998	264–246	864–684	327–723	397–379
58120649564397032718				
93871064813275645290				
03970645853872119264				
56415864003272839791				
15192564086432793078				
80926439731780152645				
39715982206438756401				
99643175530786481022				
56493875012931786420				
96409138783275021645				
06482864939750213571				
16493572182050964387				
15938726459208307164				
13079317564822648059				
81397016426425593078				
09882643572151064397				

References

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers*, *25*, 257–271.

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

Frensch, P. A., & Miner, S. C. (1994). Effects of presentation rate and of individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory and Cognition*, *22*, 9–110.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Gómez, R. L. (2006). Dynamically guided learning. In Y. Munakata, & M. H. Johnson (Eds.), *Attention & performance XXI: Processes of change in brain and cognitive development* (pp. 87–110). Oxford, En- gland: Oxford University Press.

Hoffmann, J., & Sebald, A. (2005). When obvious covariations are not even learned implicitly. *European Journal of Cognitive Psychology*, *17*, 449–480.

Hsiao, A. T., & Reber, A. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In M. Stadler, & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 471–494). Thousand Oaks, CA: Sage Publications.

Jahn, G. (2012). Replication of Study 4b by Pacton & Perruchet (2008, JEP:LMC). Retrieved from https://osf.io/vmz2e/files/FinalReport_Replication_Study4b_ Pacton Perruchet_2008_JEPLMC.pdf/

Jiang, Y., & Chun, M. M. (2003). Contextual cueing: reciprocal influences between attention and implicit learning. In L. Jimenez (Ed.), *Attention and Implicit Learning* (pp. 277–296). John Benjamins Publishing Company.

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 236–259.

Lai, J., & Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition*, *118*(2), 265–273.

Logan, G. D., & Etherton, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1022–1050.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Onnis, L., Christiansen, M. H., Chater, N., & Gómez, R. (2003). Reduction of uncertainty in human sequential learning: Preliminary evidence from artificial grammar learning. *Proceedings of the 25th Annual Conference of the Cognitive Science Society, Boston, MA*.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in speech processing. *Journal of Memory and Language*, *53*, 225–237.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 80–96.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.

Perruchet, P., Poulin-Charronnat, B., & Pacton, S. (2012). Learning nonadjacent dependencies. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1941–1944). New-York: Springer-Verlag.

Perruchet, P., & Vinter, A. (2002). The self-organized consciousness. *Behavioral and Brain Sciences*, *25*, 297–388.

Stadler, M. A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 674–685.

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564.