THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2002, *55A* (2), 485–503

The formation of structurally relevant units in artificial grammar learning

Pierre Perruchet and Annie Vinter
Université de Bourgogne, Dijon, France
Chantal Pacteau
Université Paris V, Paris, France
Jorge Gallego
Université Paris VII, Paris, France

A total of 78 adult participants were asked to read a sample of strings generated by a finite state grammar and, immediately after reading each string, to mark the natural segmentation positions with a slash bar. They repeated the same task after a phase of familiarization with the material, which consisted, depending on the group involved, of learning items by rote, performing a short-term matching task, or searching for the rules of the grammar. Participants formed the same number of cognitive units before and after the training phase, thus indicating that they did not tend to form increasingly large units. However, the number of *different* units reliably decreased, whatever the task that participants had performed during familiarization. This result indicates that segmentation was increasingly consistent with the structure of the grammar. A theoretical account of this phenomenon, based on ubiquitous principles of associative memory and learning, is proposed. This account is supported by the ability of a computer model implementing those principles, PARSER, to reproduce the observed pattern of results. The implications of this study for developmental theories aimed at accounting for how children become able to parse sensory input into physically and linguistically relevant units are discussed.

In a typical implicit learning experiment, participants are first exposed to rule-governed material, without being asked to learn about the rules or even being informed about the structured nature of the material. A subsequent test is devised to reveal whether participants have learned about the structure. The usual outcome is that their behaviour is responsive to at least some structural aspects of the test material, whereas they lack conscious knowledge about the underlying rules (see review in Cleeremans, Destrebecqz, & Boyer, 1998; Reber, 1993). The mechanisms responsible for this phenomenon have been thought to underlie many of the

---

Requests for reprints should be sent to Pierre Perruchet, Université de Bourgogne, LEAD/CNRS, Faculté des Sciences, 6 Bd Gabriel, 21000, Dijon, France. Email: pierre.perruchet@u-bourgogne.fr

This research was supported by the Centre National de la Recherche Scientifique (UMR CNRS 5022), the University of Bourgogne, and the Region of Bourgogne (Contrat AAFE).

http://www.tandf.co.uk/journals/pp/02724987.html DOI:10.1080/02724980143000451

higher human acquisitions, such as the acquisition of first or second language, natural categories, reading and writing abilities, or sensitivity to musical structure. As a consequence of their potential generality, these mechanisms have been the topic of much theoretical speculation and many experimental studies, specially during the last decade. However, they still remain the object of lively debate and controversy.

In this paper, we focus on the learning of artificial grammar, which is one of the most thoroughly investigated paradigms in the implicit learning field. In a conventional situation, participants are first exposed to a set of consonant strings based on a finite-state grammar similar to that represented in Figure 1, and their performance is subsequently assessed through their judgements of the grammaticality of new strings. The usual result is that participants are able to classify the new strings as grammatical or ungrammatical with better-than-chance accuracy. Because, at the same time, they are unable to articulate the rules of the finite-state grammar, their performance was initially perceived as evidence of the unconscious abstraction and exploitation of these rules (Reber, 1967).

However, several alternatives to the abstractionist view have been proposed. One widely held view is that the strings are processed into successive fragments or chunks composed of a few letters (Gomez & Schvanevelt, 1994; Knowlton & Squire, 1994; Perruchet & Pacteau, 1990; Redington & Chater, 1996; Servan-Schreiber & Anderson, 1990). Only fragments are memorized, and grammaticality judgements during the test are assumed to be based on whether new strings can be perceived as a combination of old fragments. Let us assume, for instance, that participants are shown the study strings *SVDRT* and *HFVDX* (with * denoting string beginning and ending), which are both generated by the grammar represented in Figure 1, and then have to judge the test strings *SVDX*, *HFVDRT*, which are also generated by the grammar. If participants have coded the study strings as the four fragments *SV, DRT*, *HFV, and DX*, the grammaticality of new test strings such as *SVDX* and

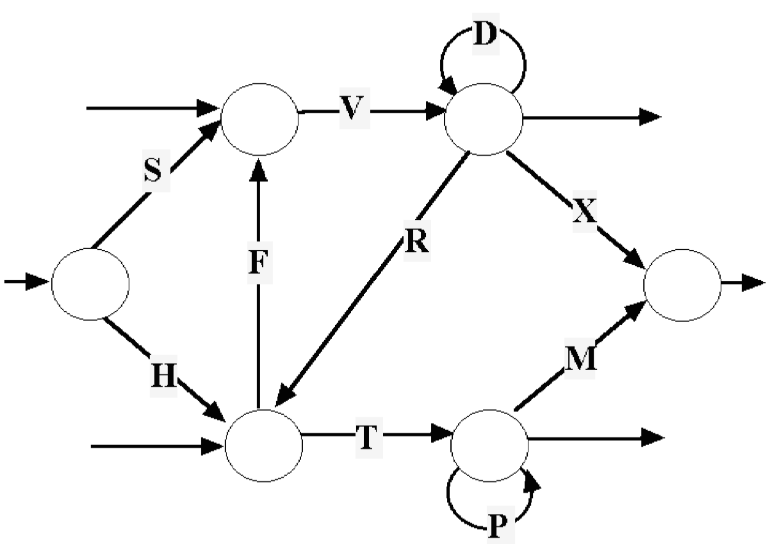

Figure 1. Schematic diagram of the grammar used in the present study.

HFVDRT follows. This possibility is highly relevant, because, given the structure of the finite-state grammars, test items are often a recombination of fragments of the study items. Thus coding the material as small units facilitates transfer and generalization to new strings.

This perspective has received considerable supportive evidence. Perruchet and Pacteau (1990) compared the performance of participants trained in standard conditions with those trained with the bigrams making up the strings. Both groups obtained similar results in a subsequent grammaticality test, suggesting that fragment knowledge alone is sufficient to account for improved performance. Such a result, however, is obtained only if the choice of the test strings by the experimenter makes the position of legal strings irrelevant for correct grammaticality judgements. Indeed, subjects in standard conditions have been shown to be sensitive to the positions of the fragments, a form of knowledge that subjects trained with isolated bigrams are obviously unable to acquire (Gomez & Schvanevelt, 1994). Redington and Chater (1996) have also demonstrated, by running simple models of performance, the surprising power of bigram and trigram knowledge in transfer situations in which this form of knowledge could be thought of as a priori insufficient. In addition, a fragmentary account has received strong support from studies involving other paradigms of implicit learning, such as the repeated-sequence tasks (e.g., Buchner, Steffens, & Rothkegel, 1998; Stadler, 1995).

From random segments to relevant units

However, a fragmentary account has its own limitations, which are made explicit in the following example. Let us assume again that participants see the study strings *SVDRT* and *HFVDX*, but that, instead of forming the fragments *SV, DRT*, *HFV, and DX*, they form the fragments *SVD, RT*, *HF, and VDX*. Any recombination of those fragments fails to generate the grammatical test strings *SVDX* and *HFVDRT*. Moreover, recombining those fragments generates non-grammatical strings, *HFRT* and *SVDVDX*. This example illustrates that the decomposition of the study strings into fragments may be more or less relevant. We define hereafter the level of relevance of a given segmentation of the study strings as the probability that a new recombination of the fragments leads to the building of grammatical, and only grammatical, new strings. Note that this definition is in keeping with the deep structure of the task: Indeed, the gist of the artificial grammar learning literature is the result of transfer to new strings, and our definition of relevance directly taps this aspect of the task.

How can the degree of relevance of a given segmentation be assessed? If well-formedness assessment of new test strings is performed thanks to the occurrence, in the test strings, of the same fragments as those in the study strings, a condition for relevance is that study and test strings share the largest number of identical fragments. Because the two subsets of strings are randomly drawn among the whole corpus of strings, this entails that fragments are all the more relevant the more frequently they are repeated throughout the strings. Thus a generally valid index of relative relevance is the mean number of repetitions of the same fragments throughout the corpus. We use here as a metric the inverse of this index, namely the number of different fragments needed to describe the corpus (the inverse relation follows from the fact that the size of the whole corpus is fixed—the fewer different fragments there are, the more they are repeated). Between two modes of segmentations, the more relevant is the one that generates the smaller number of different units.

TABLE 1
Two modes of parsing three strings of letters generated by the grammar shown in Figure 1

Segmentation	*Units*
(a)	S V R/F V D/X , S V/D/R F V , F V D/R F/V D X
(b)	S V/R F V/D X , S V/D/R F V , F V D/R F V/D X

Note: Although Segmentations a and b parse the material into the same number of units, they differ in economy because Segmentation a requires nine different units, whereas Segmentation b requires only five different units (SV, RFV, DX, D, and FVD).

To illustrate, let us consider Table 1, which represents two modes of segmenting three strings of letters generated by the grammar shown in Figure 1. The upper line (Segmentation a) illustrates a random segmentation. By contrast, the bottom line (Segmentation b) shows a more relevant segmentation. Although the total number of units (and hence the mean length of the units) does not differ in the two modes of parsing, the strings are composed of nine different units in the random segmentation, versus five in the relevant segmentation.[1]

It is worth stressing that this definition is correct only if certain conditions are fulfilled. Our definition of relative relevance holds only if the mean length of the units is the same between the modes of segmentation considered for comparison (because this condition is fulfilled in the context of the present study, only this case is considered here). Moreover, the fragments need to have a minimum size to convey some information, the value depending on the order of the sequential dependencies embedded in the grammar. For instance, the letters themselves

[1] This operational definition of relevance meets in fact another possible way of assessing relevance, starting from a careful analysis of the grammar itself. In the finite-state grammar represented in Figure 1, some transitions are unique: S is always followed by V and F by V. In most cases, a given letter can be followed by two letters: For instance, R can be followed by F and T. Finally, T has three possible successors: P, M, and "*out*", and V has four possible successors: D, X, R, and "*out*". In addition, it appears that the grammar includes one recursive loop (RFV), which is an obligatory pathway (with the repetition of the letters D and P) when long strings have to be generated. It could be argued that a segmentation is relevant whenever the segments include letters with unique transitions, and the boundaries between segments fall after a letter that may be followed by many successors. Also, a segmentation including a recursive loop as a segment looks more relevant than a segmentation in which a recursive loop is broken down. In keeping with this view, the sets of letters with a unique association (SV), and those composing a recursive loop (RFV), have been chunked together in the relevant segmentation displayed in Table 1.

However, this analytical way of assessing relevance is more complex than it appears at first glance, because the probability of occurrence of each transition must also be considered. For illustration, if a letter has two possible successors with one successor having a probability of occurrence of .99 and the other of .01, it may be relevant to form a fragment including the letter and its most frequent successor. The finite-state grammar displayed in Figure 1 certainly does not generate such striking disproportions, but there may be some related problems. For instance, although V can have four successors, it turns out that the successors X and "*out*" are certainly far less frequent than the other two, especially if long strings are allowed (because X and "*out*" end the strings). Moreover, assessing relevance in this analytical way turns out to be objectless, because it can be shown that a relevant segmentation as defined in the main text (i.e., including a minimum number of different fragments) is necessarily a segmentation in which the fragments are a good fit with the deep structure of the grammar (see, e.g., Brent, 1996; Servan-Schreiber & Anderson, 1990).

are the mode of coding generating the smallest number of different fragments. However, these fragments do not encode any regularity of the material. In the same way, forming fragments of two letters is informative only if the material is at least partially governed by first-order dependency rules (i.e., if only a subset of the possible pairs of letters is allowed).

Do actual participants in artificial grammar produce random or relevant segmentations (as instantiated by Segmentations a and b, respectively, in Table 1)? Servan-Schreiber and Anderson (1990) investigated the effect of inducing a more or less relevant parsing of the study strings. In order to constrain the nature of the units formed by the participants, they introduced spaces between small groups of letters within each string. These spaces were positioned in such a way as to induce a more or less relevant parsing of the material. Subjects' performance was sensitive to the mode of presentation. Relevant parsing induced a higher rate of correct grammaticality judgements on test strings. However, these results as such do not provide evidence that participants actually perform a relevant parsing of the letter strings when they are not constrained to do so. The fact that above-chance performance always occurs in standard grammaticality tests is not sufficient argument, because the level of performance usually reported is low enough to be easily explained by simpler processes. To our knowledge, the only evidence to date is provided by an additional aspect of the Servan-Schreiber and Anderson study. In effect, some participants performed the task in a standard condition—that is, without any extraneous spaces being introduced within the strings. It turns out that performance in this condition approximates to the performance of the participants who viewed the well-structured strings. This result suggests that participants in the standard condition naturally performed the relevant parsing that was artificially induced in the other participants. However. this evidence is indirect at best, because introducing spaces within the study strings may have altered the processing of the material in unpredictable ways, hence making the interpretation of between-groups differences hazardous.

The present study

The present study aims at assessing whether participants in artificial grammar learning experiments form random or relevant units. Accordingly, this issue appears *not* to be crucial with regard to the usually reported findings of laboratory studies of implicit learning. As claimed earlier, simpler interpretations turn out to be sufficient to account for better-than-chance grammaticality judgements. However, it remains the case that the issue is relevant for current theories of implicit learning, because any theory must be able to account for the whole pattern of results collected in the field in question. Now, as argued later in the paper, it would appear that most of the current theories are unable to account for an increasingly relevant parsing of the grammatical letter strings.

The first objective of the present study was to investigate whether participants in an artificial grammar learning experiment learn to parse the study strings into structurally relevant units. In a more operational form, the question is: Does familiarization with the letter strings decrease the number of different units needed to describe the whole corpus? Because the units of interest were assumed to shape the conscious coding of the material (e.g., Perruchet & Vinter, 1998a; Perruchet, Vinter, & Gallego, 1997), we simply asked participants what units they formed. In a first phase, participants were asked to read a set of strings generated by a finite-state grammar and, immediately after reading each string, to mark the natural segmentation positions with a slash bar. Next, the participants were submitted to a phase of familiarization with the letter strings, which consisted, depending on the group involved, of learning items by rote, performing a short-term matching task, or searching for rules. After the familiarization phase, the initial task of segmentation was repeated on the same letter strings with the same instructions. The design allowed a direct comparison of the participants' segmentation of the strings before and after familiarization with the material.

To anticipate, exposure to structured material modified the initial coding of the material in the expected direction. Our second objective consisted of testing a theoretical model devised to account for the formation of structurally relevant cognitive units by examining the ability of a computer-implemented version of this model, called PARSER, to simulate empirical findings.

EXPERIMENT

Method

Participants

The participants consisted of 78 first-year university students majoring in psychology. They participated in groups of 8–17. They were randomly assigned to experimental groups on an alternating basis. To make this possible, the instructions specific to each group were provided on separate sheets of paper.

Material

The letter strings were generated by the finite-state grammar represented in Figure 1. This grammar can generate 88 strings composed of 6, 7, or 8 letters. A total of 36 out of these 88 strings were selected for the experiment in such a way that all the paths of the grammar were equally represented (see list in Appendix A).

The letter strings were typed in upper-case Courier 16. They were arranged in three, four, or five items per line as shown in Appendix A. The start of the lines was slightly shifted in order to avoid the vertical correspondence of letters.

Procedure

The session consisted of three phases. The first and third phases consisted of segmenting the set of strings. These phases were identical and were common to all participants. The second phase was the familiarization phase, and it differed depending on which of three groups the participant was assigned to.

All participants received a booklet, whose front page began with the sentence:

"Do not turn the page before the signal."

Instructions for the first phase of the experiment were also printed on this page. They were:

"When we have to read a sequence of letters such as "QCNNWKVZ" aloud, it is usual to mark brief pauses at some places, and to spell for instance "QCN/NW/KVZ" or "QC/NNW/K/VZ". On the next page, you will see a set of sequences of letters. Your task is to segment these sequences, as you would segment them if you were reading aloud. To do that, insert a "/" where you would make a pause. Please work quickly; Don't try to think about it."

When the signal was given, participants turned the page to perform their task. The order of the 36 items for segmentation was random, but all the participants saw the same order. After all the participants in the group had completed the segmentation task, they were asked to turn the page and to read the instructions printed on the next page. These instructions differed depending on the groups.

The participants in Group 1 ($n = 25$) received rote learning instructions. They were simply asked to learn the letter strings, without any mention of a possible strategy that they could adopt, or about the nature of the memory test that they would be asked to perform. Those in Group 2 ($n = 27$) received instructions for a short-term matching task. These instructions were:

> *"The sequences of letters that you just saw are reprinted on the next page. In the upper half of the page, these sequences are numbered. The same sequences are reproduced in the lower half of the page, but in a different order and without an associated number. Your task is to write the number of each sequence in the empty brackets displayed above the unnumbered sequences."*

The instructions for the first two groups are usually termed "implicit". The participants in Group 3 ($n = 26$) received standard "explicit" instructions. They were asked to memorize the letter strings as participants in Group 1. However, the were informed that the letter string followed a complex set of rules and that discovering these rules would help them in memorizing the material.

When the signal was given, the participants turned the page and performed the task for 10 min. The letter strings for Groups 1 and 3 were typed in the same format as that used in the other phases. For Group 2, the letter strings were duplicated in the lower half of the page. A pair of brackets was displayed above each string. These brackets contained the number of the string in the upper list and were empty in the lower list. For all the groups, the order of strings was randomized, and it differed for each participant.

The participants then performed the last phase, which was identical to the first one. The only difference was that the instructions asked the participants not to try to remember the segmentation that they initially performed. The items were the same as those in Phase 1 and were displayed in the same order.

Results

The first stage of the analysis consisted of assessing the total number of units formed by the participants. The results are shown separately for each group in Figure 2. On average for all groups together, the participants segmented the material into 91.43 units before the familiarization phase and 91.90 units after familiarization. Because the entire material consisted of 258 letters, it can be deduced that the mean length of the units was 2.82 and 2.81 letters, respectively. An analysis of variance (ANOVA) was performed on the total number of units with groups (rote memory, matching task, and rule learning) as the between-subjects factor and phases (before vs. after familiarization) as the within-subject factor. Neither the main factors nor the interaction between them was significant (all Fs < 1). This means that participants formed units of the same length before and after familiarization, and that unit length did not differ as a function of the task that they performed during the familiarization phase.

Although we had no particular expectations with regard to the effect of the familiarization phase on the mean length of the units, this result has the advantage of making the assessment of structural relevance simpler. Indeed, under these conditions, the relevance of coding can be unambiguously assessed. Of two different segmentations, the one comprising the smaller number of *different* units is the more relevant.

Figure 2 shows the number of different units for each group. Averaged over groups, participants formed 38.45 different units before the familiarization phase and 35.47 different units

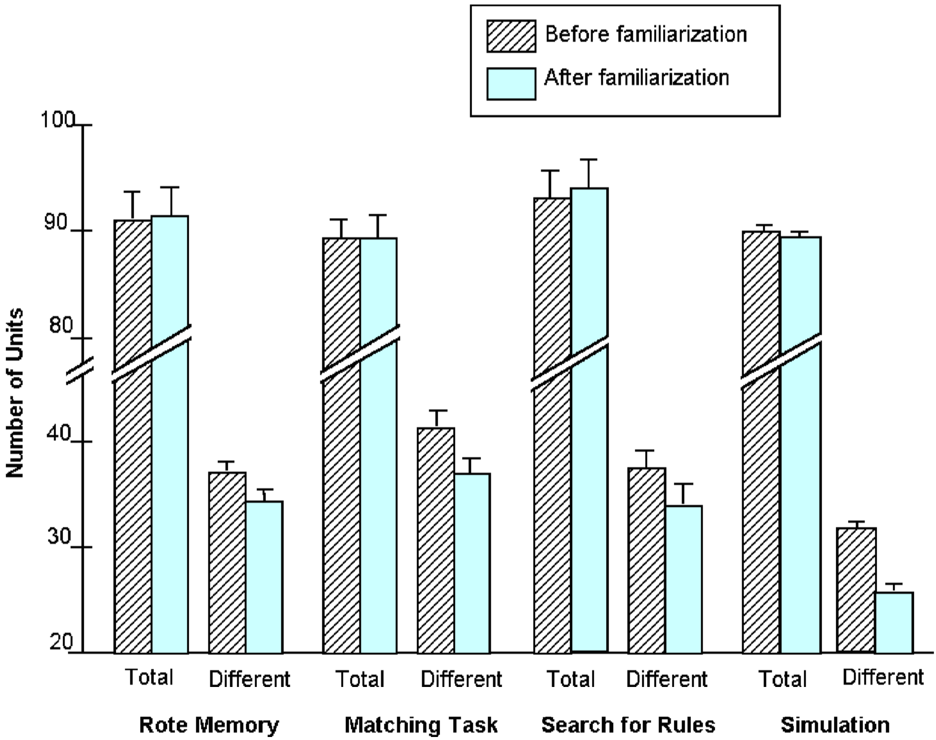

Figure 2. This figure reports (1) the total number of units formed by the participants, before and after a familiarization phase that consisted, depending on the group, of a rote memory task, a short-term matching task, or an explicit task of rule discovery; (2) the corresponding number of different units. The rightmost bars show the results generated by PARSER (parameters were set in such a way that the *total* number of units matches the values actually observed; results of interest concern the number of *different* units). Standard errors are represented at the top of each bar.

after familiarization. In other words, the percentage of different units fell from 42.05% before familiarization to 38.59% after familiarization, resulting in a gain of 3.46%. An ANOVA was performed on the number of different units with groups as the between-subjects factor and phases as the within-subject factor. The effect of phases, although modest in scale, was highly significant, $F(1, 75) = 30.65$, $p < .0001$. This effect was the same for the three groups, as evidenced by the non-significant Phases × Group interaction ($F < 1$). Planned comparisons showed that the effect of phases remained significant when each group was considered separately, all $F\text{s}(1, 24) > 9.48$, all $p\text{s} < .005$.

These results demonstrate that familiarization with the material leads to a decrease in the number of different units needed to perceive the material. Because the total number of units did not vary, this indicates that familiarization with the material leads participants to select those units that are the most often repeated, hence providing the best pieces of knowledge to process any new string generated by the same grammar. The level of improvement was the same, whether participants learned the items by rote, performed a matching task, or attempted to identify the structuring rules. This result may be viewed as an indication that the phenomenon is independent of the task actually performed by the participants. However, it is worth noting that all three tasks involved the same initial processing step, namely the attentional reading of the strings. Thus a tentative conclusion is that improvement in coding depends on some primitive aspects of string processing that are so fundamental that they are common to a wide variety of different tasks. The next section presents a theoretical interpretation, and its computer implementation, for this overall pattern of results.

PARSER: A COMPUTER MODEL OF UNIT FORMATION

We assume that the processing of the letter strings generated by a finite-state grammar proceeds through a succession of discrete, disjunctive attentional focuses, each focus embedding a few processing primitives (for a theoretical justification of this principle, see Perruchet & Gallego, 1997). In keeping with the literature on associative learning and memory, the primitives that are perceived within one attentional focus as a consequence of their experienced spatial or temporal proximity become the constituents of a new representational unit. If the association between the primitives that form a percept is not repeated, the representational unit created by this percept rapidly vanishes, as a consequence of both natural decay and interference with the processing of similar material. However, if the same percept re-occurs, the internal representation is progressively strengthened. The relevant units, in this sketch, emerge through a natural selection process, because (1) repeated percepts, as explained earlier, are more likely to match a structurally relevant unit than an irrelevant fragment, and (2) forgetting and interference lead the human processing system to select the repeated parts from all of the parts generated by the initial, presumably mostly irrelevant, chunking of the material.

Although the simplicity of this account is attractive, it might be argued that its reliance on some form of natural selection prevents it from explaining the quick formation of relevant units observed in the preceding experiment. To assess whether the theoretical principles underlying our account are powerful enough to explain observed behaviour, those principles were implemented in a computer model, PARSER. PARSER was originally developed (Perruchet & Vinter, 1998b; see also Brent, 1999, for an overview of the relevant literature including a comparison of PARSER with other computer models) to deal with the situations proposed by Saffran and co-workers (e.g., Saffran, Newport, & Aslin, 1996), in which participants have to find the words forming an artificial language presented as an unbroken sequence of syllables. The model works as follows: at each processing step, the model determines on a random basis the number of primitives embedded in the next percept; this leads to the formation of a new processing unit. This unit, of course, may be irrelevant. However, it is short-lived, and disappears from the system if not repeated in the near future. When the unit is perceived repeatedly, it is strengthened, and it serves to guide further perception. Appendix B summarizes the characteristics of the model, presented at length in Perruchet and Vinter (1998b).

Applying PARSER to a new domain requires us to specify the nature of the processing primitives relevant for this domain. Indeed, as emphasized by the proponents of the episodic-processing view of learning (e.g., Higham, 1997a) what is important is the nature of the processing carried out by the subjects, which depends only partly on the literal stimulus. In the original version of PARSER, the processing primitives were the syllables, a choice that was motivated by a vast amount of literature suggesting that syllables are natural units of speech processing (Jusczyk, 1997; see Perruchet & Vinter, 1998b, Footnote 1). In the present context, where the material is made up of consonant sequences, the individual letters were taken as processing primitives. This does not mean that letters are the basic units for the processing of printed material in any given situation: Units can certainly be larger, especially in the case of familiar items, or smaller, for instance if the task requires subjects to detect intra-letter features. However, the choice of letters seemed to us to be the most psychologically plausible, given the material used in the experiment and the nature of the orienting tasks presented to the participants.

The parameters of the model used in the present paper were the same as those used in Perruchet and Vinter (1998b), with two exceptions. These exceptions were introduced in order to fulfil a prerequisite making the measure of relevance simpler. In effect, our objective was to compare the degree of coding relevance achieved by the model and by actual participants. As claimed earlier, comparing the relevance of different codings is a simple matter when the total number of units is the same. In that case, the most relevant coding is the one that includes the smallest number of different units. However, the problem becomes quite complex when the total number of units varies. For the sake of simplicity, we constrained PARSER to produce the same *total* number of units as that generated by participants. This was achieved by adjusting two parameters in the model on a trial-and-error basis. As a result (1) the number of primitives incorporated in a unit was randomly distributed in the range 2 to 4 instead of 1 to 3, and (2) whatever the issue of the random distribution, the introduction of new primitives into an old unit was stopped as soon as the length of the unit was ≥ 3.

The material was entered in PARSER in exactly the same way as it was displayed to actual participants. The set of 36 letter strings was first processed, in the same order as that for participants, for an assessment of the initial units created by the model. Then a second reading of the letter strings was performed, with the order of the strings varying randomly through simulations. In the test phase, the letter strings were shown again in the same order as that in the first phase for an assessment of the final units created by the model. Although each of the three successive phases were assigned different functions in the context of the simulations (initial parsing, familiarization, final parsing), the model performed exactly the same operations in the three phases. This solution was chosen for two reasons. First, learning cannot be easily "blocked" in PARSER, because the formation of relevant units is inherent to the simple processing of the material. Second, actual participants presumably also learn about the material throughout the entire session, irrespective of whether the experimenter has assigned the status of "test" or "training" to a particular phase.

Because the model includes the introduction of a random factor at several points, the reported results were averaged over 100 runs.

The parameters of the model were correctly adapted to reproduce those aspects of the actual data that were not of focal interest. Indeed, the total number of units used by PARSER to process the strings was 90.04 in the initial phase and 89.69 after familiarization. The respective mean lengths of the units were 2.865 and 2.876. These values closely approximate the values observed in actual participants (the mean length of the units was 2.82 and 2.81, respectively).

The crucial aspect concerned the number of *different* units. As shown in Table 1, PARSER simulated the key result of the experiment. Although the total number of units did not vary substantially, the number of different units used after familiarization (25.60) was notably lower than the number of units needed in the early phase (32.28). In other terms, the percentage of different units fell from 35.85% before familiarization to 28.54% after familiarization, representing a gain of 7.31%. One of the main differences from the performance of the actual participants was that the gain due to familiarization with the letter strings was larger (7.31% vs. 3.46%). Thus PARSER learned faster or better than actual participants. Another difference related to the initial level: PARSER used fewer units than did actual participants as early as the first test session (32.28 vs. 38.45). This latter result can also be accounted for by the superiority of PARSER in providing an economical coding after only limited practice. Indeed, it must be remembered that PARSER began to learn *during* the first test; it is likely that the mean performance reached during this test reflects the early effects of learning.

A strong indication of the relevance of the model would consist of a demonstration that, beyond the parallelism observed on the overall performances, the model generates the very same units as those generated by the actual participants. Unfortunately, such an analysis turned out to be unmanageable, because of the variability of the segmentations generated by both PARSER and the participants. It must be understood that the structure of the finite-state grammar does not provide compelling constraints on the nature of the relevant units. Several segmentations sharing approximately the same degree of relevance are possible. The choice between different segmentations certainly depends on the initial units that are created, presumably on a random basis. In other words, a given set of letters may or may not be a relevant chunk, depending on the nature of the other units that have already been created. This entails that the relevance of a segmentation can be assessed on a within-subject basis, but any analyses become meaningless when the units are pooled over subjects or simulations.

GENERAL DISCUSSION

Three major results emerge from the present study. First, after some familiarization with a set of letter strings generated by a finite-state grammar, the coding of the strings required a smaller number of different units, although the total number of strings (and hence the mean length of the units) remained unchanged. This was taken as evidence for the human ability to parse into structurally relevant units an arbitrarily structured material in which unit boundaries are not available to immediate perception.

A small number of prior studies had suggested the ability to parse arbitrary material into its relevant units. In several experiments, Saffran and co-workers (e.g., Saffran et al., 1996) used an artificial language composed of six trisyllabic words, such as *babupu* and *pupada*. The words were represented in random order. They were read by a speech synthesizer in immediate succession, without pauses or any other prosodic cues. Thus participants heard a continuous series of syllables without any word boundary cues. In a subsequent test phase, participants performed a forced-choice test in which they had to indicate which of two items sounded more like a word from the artificial language. One of the items was a word from the artificial language, whereas the other was a new combination of three syllables belonging to the language. Participants performed significantly better on this test than would be expected by chance. This result suggests that people are able to learn the relevant units of an arbitrary language without any unit boundary cues. However, the evidence was somewhat less compelling than in our own experiment because there was no direct indication that trisyllabic units were formed in the Saffran et al. experiments. Indeed, only trisyllabic items were displayed in the forced-choice test, and it is possible that the choice of one item over the other was guided by some sublexical knowledge, such as the probability of pairwise associations between contiguous syllables. In our experiment, participants were asked to directly segment the material, and therefore their performance provides a direct indication about the length of the units they formed.

The second result of our study is that structurally relevant coding was achieved irrespective of whether participants learned items by rote, performed a short-term matching task, or searched for the rules of the grammar during the familiarization phase. Thus learning to code the material into relevant units is relatively independent of specific task demands. Of special interest is the fact that instructions asking for rule searching are neither beneficial nor detrimental to improved parsing. This does not mean, however, that parsing is independent of the processing performed during the familiarization phase. First, the sampling of tasks is too limited to permit generalization. Second, it is worth noting that all three tasks involved the same initial processing step, namely the attentional reading of the strings. Thus a tentative conclusion is that improved coding depends on some aspects of string processing that are so fundamental that they are common to a wide variety of different tasks, at least during an earlier stage of practice.

This second result is also consistent with other data in the literature. Using the artificial language paradigm described earlier, Saffran and co-workers reported roughly similar performances irrespective of whether participants were asked to figure out where the words of the language began and ended (e.g., Saffran et al., 1996), or were simply exposed to the speech flow while having to draw an illustration with a colouring program (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). Also, the present finding fits well with the fact that, in standard artificial grammar learning studies, the discriminability of grammatical and nongrammatical test strings has been found independent of whether participants received implicit or explicit instructions during the study session (e.g., Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990), at least when the material is sufficiently complex (Reber, Kassin, Lewis, & Cantor, 1980). This relative insensitivity to task demands observed in both our segmentation task and in standard tasks of grammaticality judgements reinforces the assumption of a causal relation between participants' ability to parse the material into structurally relevant units and participants' grammaticality judgements.

The third main result of our study is that the formation of structurally relevant units can be accounted for by fairly simple processes. Our model assumes that, provided the task requirements involve the attentional reading of the letter strings, those strings are necessarily perceived as a succession of small and disjunctive units composed of a few primitives. This characteristic is thought to be inherent to the attentional processing of ongoing information. When a unit is repeatedly perceived, its components are associated and form a new representational unit as an automatic by-product of the joint attentional processing of the components. The structurally relevant units emerge thanks to a sort of natural selection process: Among all the units that are created, only those matching the relevant units are repeated sufficiently often to resist forgetting and interference. A computer-implemented version of this model, PARSER, demonstrated that so simple a mechanism was able to simulate observed performances: Where differences emerge between actual and simulated data, it is always the simulations that overachieve. This overachievement may hardly be considered to be detrimental to the theoretical validity of the model. Where the simulations exceed human performance, the addition of noise at different steps of the model would appear to be an obvious remedy.

Alternative models of implicit learning

Are alternative models of implicit learning able to account for the present results? The abstractionist view (e.g., Lewicki, Hill, & Czyzewska, 1992; Mathews, 1990; Reber, 1967) says nothing about the role that may be attributed to the chunking process.

The first historical alternative to the abstractionist position was the so-called instance-based or exemplar-based model proposed by Brooks (Brooks, 1978; Vokey & Brooks, 1992). In Brooks' model, participants store in memory the grammatical strings seen during the study phase, without any form of condensation or summary representation. During the test phase, they judge the grammaticality of test strings as a function of their similarity to the stored strings. In the earlier version of the position, similarity was assessed in terms of the surface features that a test string shares with a single study string, both strings being considered as holistic entities. This version, like the abstractionist view, says nothing about the issues at hand in this paper. However, the so-called "episodic account" inspired by the instance-based model (e.g., Higham, 1997b; Neal & Hesketh, 1997; Whittlesea & Dorken, 1993) exhibits several improvements. First, it now specifies that subjects do not necessarily process the whole string, but, depending on the mode of processing encouraged during the study phase, they may process larger or, interestingly, smaller entities, such as individual letters or sets of letters. In this context, the view that strings are processed into successive fragments or chunks becomes a particular case of a more general, episodic model. Second, similarity is not necessarily assessed with regard to a single study item, but may result from the pooling of several instances. In this way, the episodic model becomes able to encompass, at least in principle, the frequency effects that are essential within our own position. To the best of our knowledge, however, the proponents of the episodic view have never considered the possibility that the nature of the processed episodes may change during the course of learning in a way that makes them increasingly relevant for the task. We also fail to find in the episodic model a mechanism able to account a posteriori for such a phenomenon.

Let us consider now the connectionist models of implicit learning. The most popular is the simple recurrent network, or SRN, initially proposed by Elman (e.g., Elman, 1990; see also Cleeremans, 1993). An SRN is a network that is designed to learn to predict the next event of a sequence. Elman (1990) presented such a network with a continuous stream of phonemes one phoneme at a time, the task being to predict the next phoneme in the sequence. The accuracy of prediction was assessed through the root mean square error for predicting individual phonemes. After training, the error curve had a strikingly marked saw-tooth shape. As a rule, the beginning of any word coincided with the tip of the teeth. This means that after a word, the network was unable to predict the next phoneme. However, as the identity of more and more phonemes in a word was revealed, the accuracy of prediction increased up to the last phoneme of the word, and the error curve therefore fell progressively. The start of the next tooth indexed the beginning of the next word. Therefore, an SRN provides at least a part of the information needed to parse a continuous speech flow into words (for more recent models, see Aslin, Woodward, LaMendola, & Bever, 1996; Christiansen, Allen, & Seidenberg, 1998). Generalizing from these data, there is high likelihood that an SRN is also able to help in discovering the relevant units of the material generated by a finite-state grammar. However, it is worth stressing that the relevant units can only be inferred from the graded distribution of errors after learning is completed. They have no genuine existence, and as such they are devoid of any causal function in the learning process. This feature radically distinguishes connectionist models from PARSER, in which the new units serve to shape the encoding of incoming information.

By contrast, models based on the notion of partial or fragmentary knowledge (e.g., Dulany et al., 1984; Gomez & Schvanevelt, 1994; Knowlton & Squire, 1994; Perruchet & Pacteau, 1990) give chunking a central position. Most of these models do not possess procedures that make it possible to account for an improvement in the way the material is coded. However, an exception can be found in Servan-Schreiber and Anderson (1990) who propose an account that shares many points with the present one. They also assert that "the main learning process in the Reber task [is] some sort of chunking and that grammatical discrimination [is] based on the degree to which compact representations of strings could be built from the collection of learned chunks" (p. 600). Their "competitive chunking model" also leads to the formation of structurally relevant units. However, we fail to see how this model can account for the performance revealed in the present experiment. Indeed, Servan-Schreiber and Anderson understand learning as the formation of longer and longer chunks, from the individual letter to the whole strings, with the chunks being structured within a hierarchical network. It should be remembered that we did *not* observe a progressive lengthening of the segments formed by the participants, but rather an improvement in their structural relevance while their mean length remained constant. This pattern of results is more consistent with the process of progressive selection of the relevant units from possible candidates, as implemented in PARSER, than with the assumption of a hierarchical network of chunks.[2]

Implications for other research domains

Parsing the world into its relevant units is essential for adaptation. For instance, the task of segmenting the speech flow into words is construed as the primary task that infants have to perform when acquiring their mother tongue (e.g., Brent & Cartwright, 1996). The crucial importance of performing a relevant segmentation of the speech signal has been emphasized for other aspects of language, such as reading and spelling (e.g., Muter, Hulme, Snowling, & Taylor, 1998). Likewise, the task of segmenting the visual environment into its actual units has also been construed as one of the essential tasks for infants adapting to their environment (e.g., Markman, 1990) or for adults performing complex categorizations (e.g., Schyns,

[2] According to a reviewer of an earlier version of the paper, the result that the mean length of the chunks did not increase with training could be due to the fact that participants remembered their pre-training segmentation during the final test. We believe it to be somewhat implausible that participants may remember the way they segmented specific strings, given the interference introduced by the training phase. In any case, the fact that the *nature* of the chunks changed does not support this hypothesis. However, it remains possible that participants remembered that they had segmented the strings into chunks of, say, 2 or 3 letters in the initial phase, and that, for some unknown reason, they tried to keep the same chunk size in the final phase. To rule out such a hypothesis would require further studies involving a between-subjects design. However, even if this factor is at least partially responsible for the stability of the chunk size, it remains the case that the Servan-Schreiber and Anderson (1990) model does not contain any mechanisms that account for our specific result, namely the increasing relevance of chunks whose size does not increase.

Goldstone, & Thibaut, 1998). Thus the ability to parse the sensory input correctly appears to be fundamental to development and learning, because it allows people to perceive and represent the world in a way that corresponds to certain aspects of its structure.

Now, simple observation of everyday behaviour indicates that humans usually manipulate structurally relevant units. Everyone codes the continuous speech stream of a speaker as a succession of words and structured sentences, and not as a succession of segments each involving a random number of individual auditory signals. Similarly, one perceives and represents unitary and meaningful objects, and not a set of randomly juxtaposed features. Does this mean that our study only replicates in a laboratory setting a phenomenon that has always been known? Our argument is that the present findings reveal at least two aspects of the phenomenon that can not be inferred from the simple observation of everyday behaviour, both of them providing worthwhile information about the mechanisms underpinning the target ability.

First, our study shows that a structurally relevant coding can be acquired with *arbitrary* material. Developmental researchers have provided various explanations of how children become increasingly sensitive to the structural constraints embedded in their environment. However, it appears that the explanations proposed by a few distinguished authors do not apply to arbitrarily structured material, because they rely on some form of nativist assumption. Relevant units are generally thought not to be identifiable as such at birth. However, their discovery is thought to be strictly guided by innate constraints and domain-specific knowledge (Bower, 1979; Karmiloff-Smith, 1992), assumptions (Markman, 1990), presuppositions, or intuitive theories (Spelke, Breinlinger, Macomber, & Jacobson, 1992) about the structure of the world. Similar explanations have evolved in other domains. For instance, a prevalent view in object perception is that all objects are perceived as a combination of a few basic elements. In the Biederman (e.g., Biederman, 1987) recognition by component theory, the recognition of a very large set of objects is construed as the combination of 36 geometric elements, called "geons". Geons are the relevant units permitting the perception and representations of external objects, and, in this way, they play the role of the units that PARSER extracts from the linguistic material. However, in Biederman's theory, geons are a priori fixed units.

Second, material generated by artificial grammar provides no extraneous indications on the relevant units other than those embedded in the material itself. Words have an associated meaning, and many of them match discrete objects. This referential function of words can help their discovery within the speech flow. Likewise, objects have a material consistency, and they can be apprehended through senses such as touch, a feature that can also help their isolation within the visual field. By contrast, the relevant units of the letter strings generated by a finite-state grammar can only be discovered thanks to the intrinsic properties of the grammar.

Needless to say, our results do not rule out the possible implication of innate knowledge in the segmentation issue. Likewise, they do not rule out the idea that multiple sources of information are used to parse the world into its relevant units. However, the present findings point out the importance of aspecific learning mechanisms able to extract the structure of complex material. In those respects, our study supports recent approaches emphasizing the role of statistical or distributional learning mechanisms (e.g., Bates & Elman, 1996; Redington & Chater, 1997).

To summarize, the present study shows that participants learned to code material generated by a finite-state grammar into increasingly relevant units, and that this result is independent of whether or not the participants were asked to search for the rules of the grammar. A theoretical account of this phenomenon, based on general principles of associative memory and learning, was proposed. This account was supported by the ability of a computer model implementing those principles, PARSER, to reproduce the observed pattern of results.

REFERENCES

Aslin, R.N., Woodward, J.Z., LaMendola, N.P., & Bever, T.G. (1996). Models of word segmentation in fluent maternal speech to infants. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 117–134). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, *274*, 1849.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Bower, T.G.R. (1979). *Human development*. San Francisco: Freeman.

Brent, M.R. (1996). Advances in the computational study of language acquisition. *Cognition*, *61*, 1–38.

Brent, M.R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Sciences*, *3*, 294–301.

Brent, M.R., & Cartwright, T.A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, *61*, 93–125.

Brooks, L.R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–215). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Buchner, A., Steffens, M.C., & Rothkegel, R. (1998). On the role of fragmentary knowledge in a sequence learning task. *Quarterly Journal of Experimental Psychology*, *51A*, 251–281.

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, *13*, 221–268.

Cleeremans, A. (1993). *Mechanisms of implicit learning: A connectionist model of sequence processing* (p. 227). Cambridge, MA: MIT Press, Bradford Books.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406–416.

Dulany, D.E., Carlson, A., & Dewey, G.I. (1984). A case of syntactical learning and judgement: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Gomez, R.L., & Schvanevelt, R.W. (1994). What is learned from artificial grammar? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 396–410.

Higham, P.A. (1997a). Chunks are not enough: The insufficiency of feature frequency-based explanations of artifician grammar learning. *Canadian Journal of Experimental Psychology*, *51*, 126–138.

Higham, P.A. (1997b). Dissociations of grammaticality and specific similarity effects in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1029–1045.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: Bradford/MIT Press.

Knowlton, B.J., & Squire, L.R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 79–91.

Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, *47*, 796–801.

Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*, 57–77.

Mathews, R.C. (1990). Abstractness of implicit grammar knowledge: Comments on Perruchet and Pacteau's analysis of synthetic grammar learning. *Journal of Experimental Psychology: General*, *119*, 412–416.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 71*, 3–27.

Neal, A., & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review*, *4*, 24–37.

Perruchet, P., & Gallego, J. (1997). A subjective unit formation account of implicit learning. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 124–161). Oxford, UK: Oxford University Press.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264–275.

Perruchet, P., & Vinter, A. (1998a). Learning and development: The implicit knowledge assumption reconsidered. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495–531). Thousand Oaks, CA: Sage Publications.

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

Perruchet, P., Vinter, A., & Gallego, J. (1997). Implicit learning shapes new conscious percepts and representations. *Psychonomic Bulletin and Review*, *4*, 43–48.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Reber, A.S., Kassin, S.M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 492–502.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A re-evaluation. *Journal of Experimental Psychology: General*, *125*, 123–138.

Redington, M., & Chater, N. (1997). Probabilistic and distributional approaches to language acquisition. *Trends in Cognitive Science*, *1*, 273–277.

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., & Barrueco, S. (1997). Incidental language learning. *Psychological Science*, *8*, 101–105.

Schyns, P.G., Goldstone, R.L., & Thibaut, J.P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences*, *21*, 1–53.

Servan-Schreiber, D., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592–608.

Spelke, E.S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605–632.

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

Vokey, J.R., & Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328–344.

Whittlesea, B., & Dorken, M.D. (1993). Incidentally, things in general are incidentally determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General*, *122*, 227–248.

Original manuscript received 28 February 2000
Accepted revision received 16 August 2001

APPENDIX A

HFVDRFVX, FVRFVRTM, FVDRFVDX, SVRTPPM
SVDDRTPP, FVDDRTPM, FVDDRFVD
HFVRTPP, HFVDDRTM, SVRFVRTP, HFVRFVRT
FVRFVX, FVRFVDDX, FVDRFV, SVDRFVDD
SVDRTPP, HFVRTP, FVDRTPM, SVRTPP, SVDRFV
FVDRFVD, SVRFVD, SVDDRT, SVRFVDRT
HFVDDX, HFVDDRFV, HFVRFVD, SVDDRFV
FVDDRTP, HFVDRTP, SVRFVDD, FVRFVDX
HFVRTPPM, HFVDDRT, FVDRTM, SVDRTPPM

Set of letter strings generated by the grammar represented in Figure 1 and used in the present study. They are ordered as they were displayed in Phases 1 and 3.

APPENDIX B

PARSER is centred on a single vector, known as Percept Shaper (PS). PS is composed of the internal representations of the displayed material and may be thought of as a memory store or a mental lexicon. A weight, which reflects the person's familiarity with the item, is assigned to each element in PS. At the start of the familiarization session, PS contains only the primitives needed for the processing of the material (i.e., the letters). At the end, it should contain, in addition, the structurally relevant units that form the material.

The way the words are built in PS during training is described in the flowchart in Figure 3. Let us consider how the flowchart works for Simulation 1 (over 100) of the reported study. The first string was HFVDRFVX. The string was first segmented into small and disjunctive parts. In PARSER, the multiple determinants of this initial parsing were simulated by a random generator, which selected the size of the next percept within a range of 2 to 4 units (Figure 3, Step a). For Simulation 1, the random generator provided 3 and 3 in the first two trials. In consequence, the first percepts were *HFV* and *DRF*. Because none of the first four percepts was present in PS (Step b), they were created as new units (Step c) and assigned a weight (Step d). Also, the weights of the components, *H*, *F*, *V*, *D*, and so on, were incremented.

At each time step (a time step is defined by the processing of a new percept—that is, by one cycle in the flowchart in Figure 3), the units forming PS are affected by forgetting and retroactive interference (Figure 3, Step f). Forgetting is simulated by decreasing all the units by a fixed value. Interference is simulated by decreasing the weights of the units in which any of the letters involved in the unit currently processed are present. In the case described here, interference occurred while *DRF* was perceived. Indeed, *F* was already present in HFV. In consequence, the weight of *HFV* was decremented (in addition to the decrement due to forgetting).

In this early phase, perception was driven by the initial primitives of the system, namely the letters. However, the psychological principles implemented by the model stipulate that a representation created during learning may become able to guide perception, as the initial primitives were. The condition for an element of PS to shape perception is that its weight is at least equal to a threshold value, Wt. In contrast, when the frequency of perceiving a given element is not high enough to counteract the effects of forgetting and interference, this element is removed from PS when its weight becomes null.

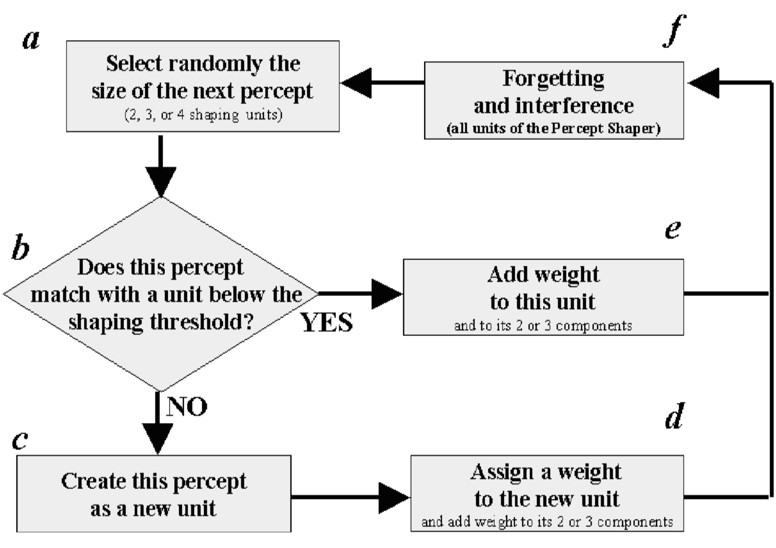

Figure 3. Operations performed by PARSER at each time step.

In the reported simulation, as in Perruchet and Vinter (1998b), the starting weight given to any created unit was set to 1. The primitives that formed PS before training were also assigned a weight of 1. The increment received by an old unit in PS when this unit serves to shape perception was set to 0.5. The decrements due to forgetting and interference were set to 0.05 and 0.005, respectively. The last parameter, namely the threshold value above which a unit is able to shape perception, was set to 1. Although the absolute values of these parameters were arbitary, their relations, which were the only relevant aspects for the model, possesses at least a rough behavioural likelihood. For instance, the parameters were set in such a way that a unit, when just created in PS, is only able to shape the immediately subsequent percept. Indeed, its initial weight (1) quickly decreases due to forgetting (0.05), and it therefore no longer meets the threshold value for shaping perception (1). If this unit is not perceived again within the next 20 (1/0.05) percepts, its weight becomes null, and it will be eliminated from PS (note that interference may speed the process). A unit that has gained a weight of 3, for instance, which means that it has been perceived from 4 to 10 times or more (each new percept involving this unit is accompanied by a gain of 0.5, but the effects of forgetting and interference are unpredictable because they depend on the number and the nature of the intervening percepts), will lose its ability to shape new percepts within 40 time steps and will disappear from PS within 20 further time steps (again when interference is ignored).

Let us return now to the performances observed in Simulation 1 to examine how the system works quantitatively. Table 2 shows the content of PS and the weight of each unit (Columns 1 and 2) after the processing of 49 syllables (PS also included units with a weight lower than 1, which are not reproduced here). The continuation of the sequence was *DRFVD*. The random generator (Figure 3, Step a), determining the number of units present in the next percept, provided a value of 3. Perception was shaped by the three units *D*, *R*, and *FVD*. Because *DRFVD* did not match a unit the weight of which was below the shaping threshold (Step b), this percept was created as a unit in PS (Step c) and was assigned a weight of 1 (Step d; see Table 1, Column 3). In addition, the components forming *DRFVD* received an additional weight of 0.5 (Table 2, columns 4, 5, and 6). Note that *FVD* was incremented, in keeping with the fact that it was an autonomous component of the percept *DRFVD*, but not its primitive parts *F*, *V*, and *D*, which did not share this status. On the contrary, *FVD* interfered with *F*, *V*, and *D*, and, in consequence, the weights of these units were decremented by .005. Finally, all the units in PS were decremented by 0.05 to simulate forgetting (Figure 1, Step f; see Table 2, column 7). The right-hand column of Table 2 displays the state of PS after *DRVD* has been processed.

TABLE 2

Changes in constituents of PS during one time step. In this example, the changes are due to the perception of *DRFVD*, when *D, R*, and *FVD* are existing units of the system

Unit	*Initial weights*	*Creation*	*Processing components*[a] *D*	*R*	*FVD*	*Forgetting*	*Final weights*
V	4.60				–.005	–.05	4.545
R	3.65			+.5		–.05	4.10
D	3.645		+.5		–.005	–.05	4.09
F	3.60				–.005	–.05	3.545
P	3.10					–.05	3.05
FVD	2.875				+.5	–.05	3.325
M	2.85					–.05	2.80
T	2.35					–.05	2.30
DRT	1.745					–.05	1.695
S	1.55					–.05	1.50
X	1.20					–.05	1.15
DRFVD	–	1					1

[a]Add weight; interference.