Indirect Effects of Synthetic Grammar Learning in an Identification Task

Axel Buchner

Two experiments investigated the effects of incidental learning of an artificial grammar on an indirect measure. Ss memorized consonant strings and later identified these in a perceptual clarification procedure. The competitive chunking model by E. Servan-Schreiber and J. R. Anderson (1990) predicted faster identifications of grammatical as opposed to nongrammatical strings. Both experiments confirmed this prediction. Experiment 2 further investigated whether faster identification induces a feeling of familiarity, which should increase the probability of subjects responding "grammatical" to a string. Faster identifications were indeed related to "old" responses in a recognition judgment. However, there was no systematic relation between speed of identification and grammaticality judgments, which is inconsistent with the prediction of the competitive chunking model that familiarity exclusively mediates grammaticality judgments.

Although there is no need to include the concept of consciousness into scientific theories of human learning and memory (see, e.g., Velman, 1991; Wilkes, 1988), it appears to be this very topic that has stimulated much of the recent interest in, and debate about, "implicit" phenomena. For instance, in the context of implicit grammar learning, discussions emerged as to whether or not subjects can be said to be conscious of (fragments of) the rules underlying the generating grammar (Dulany, Carlson, & Dewey, 1984, 1985; Reber, Regan, & Allen, 1985), and whether or not subjects have conscious explicit knowledge about permissible letter bigrams (Mathews, 1990; Perruchet & Pacteau, 1990, 1991; Reber, 1990).

In a typical grammar-learning experiment, subjects start by memorizing strings of letters generated by a finite-state grammar. Subsequently, they are informed that the strings were generated by a complex set of rules. Subjects must then classify new strings as either grammatical or nongrammatical. Above-chance performance on these judgments of grammaticality is interpreted as the result of (a) the unconscious application of implicitly acquired tacit knowledge (Reber, 1989), which "is a valid, if partial, representation of the actual underlying rules of the language" (Reber & Allen, 1978, p. 181); (b) the deliberate application of explicit rules of correlated grammars (Dulany et al., 1984), defined on features that indicate grammatical status and lead to approximately the same classifications as the rules of the experimenter’s generating grammar; or (c) the use of fragmentary but explicit knowledge of permissible bigrams of letters (Perruchet & Pacteau, 1990).

Grammaticality judgments, by their very nature, must be regarded as direct, rather than indirect, measures of knowledge (Richardson-Klavehn & Bjork, 1988). That is, when subjects are asked to perform grammaticality judgments, the instructions to the subjects need to point out the regularities in the strings, and it is necessary to inform them about the actual rules of the language (Reber & Allen, 1978, p. 181); (b) the deliberate application of explicit rules of correlated grammars (Dulany et al., 1984), defined on features that indicate grammatical status and lead to approximately the same classifications as the rules of the experimenter’s generating grammar; or (c) the use of fragmentary but explicit knowledge of permissible bigrams of letters (Perruchet & Pacteau, 1990).

Correspondence concerning this article should be addressed to Axel Buchner, Department of Psychology, Trier University, Universitätsring 15, 54286 Trier, Germany. Electronic mail may be sent to buchner@cogpsy.uni-trier.de.
degree, the application of explicit rules during a test, independent of whether learning was incidental or not. This limits the use of direct grammaticality judgments in grammar-learning studies. The obvious alternative, therefore, would be to use indirect measures of knowledge that have been developed in the context of implicit-memory research and test whether such measures can indicate grammatical knowledge in grammar-learning experiments. This approach to measuring grammar learning has not yet been explored systematically (but see Gordon & Holyoak, 1983, Experiment 1, who asked their subjects to rate how much they liked certain strings), and it could also be one way to overcome the often-deplored mutual neglect of implicit-memory research, on the one hand, and implicit-learning research, on the other hand (see, e.g., Reber, 1989).1

The present experiments, therefore, use as a measure of grammatical knowledge the time needed to identify grammatical as opposed to nongrammatical strings when the strings are presented in a perceptual clarification procedure. By itself, this is a novel test of grammatical knowledge that has not been used before. However, this particular indirect measure has been chosen for yet another—and more important—reason. Identification time can be used to test an interesting model of the grammar-learning task that has recently been suggested by Servan-Schreiber and Anderson (1990). Their competitive chunking model explicates how processes relevant in the memorization phase can support later discriminations between new grammatical and nongrammatical strings. It is proposed in this process model that subjects form chunks of letters when forced to memorize letter strings.2 For instance, the string PTVPXVPS (generated by Grammar A, see Figure 1) may be memorized as consisting of five chunks, [PT] [V] [PX] [V] [PS]. These chunks can be said to capture some of the regularities inherent in the grammatical strings. As learning proceeds, longer chunks will be created from more elementary chunks (e.g., [PTV] [PX] [VPS]), resulting in a hierarchy of chunks that may then be used to perceive new strings generated by the same set of rules. The smaller the number of chunks needed to identify a string (i.e., the larger the chunks that can be applied to a particular string of a certain fixed length), the easier the string can be perceived and the more familiar it appears. This feeling of familiarity critically mediates between the grammatical status of a string and the subjects’ grammaticality judgments. The probability of a subject classifying a string as grammatical increases as a function of the feeling of familiarity associated with that string. Grammar learning, according to this model, can be said to be implicit in the sense that the chunking processes relevant during memorization are transfer appropriate (Kolers & Roediger, 1984; Morris, Bransford, & Franks, 1977; Roediger, Weldon, & Challis, 1989) for later string perception and, mediated by familiarity, for later discrimination performance.

It is common to both the competitive chunking model proposed by Servan-Schreiber and Anderson (1990) and the theoretical approaches of Dulany et al. (1984, 1985) and of Perruchet and Pacteau (1990) that grammatical knowledge is assumed to be captured by mental representations of permissible groups (chunks, features, or bigrams) of letters. The models deviate, however, in the way this knowledge is assumed to be turned into overt grammaticality judgments. Whereas Dulany et al. and Perruchet and Pacteau seem to have suggested that letter chunks relevant to a particular classification are searched and retrieved from memory, the competitive chunking model assumes that feelings of familiarity, which are based on implicit processes, mediate grammaticality judgments. Servan-Schreiber and Anderson successfully applied their model to predict patterns of grammaticality judgments under various conditions, but they did not assess the postulated familiarity component directly. It is a major goal of the present research to fill this gap by empirically exploring the role of familiarity in grammaticality judgments.

Both the search and the familiarity component, of course, are also found in models of recognition memory (e.g., Gillund & Shiffrin, 1984; Johnston, Dark, & Jacoby, 1985; Johnston, Hawley, & Elliot, 1991; Mandler, 1979, 1980; Mandler, Osen Hamson, & Dorfman, 1990). For instance, according to Johnston et al. (1985), two processes contribute to recognition judgments, one being memory search, the other being perception.

1 Note that Miller (1958; see also Reber, 1967) used the number of correctly recalled letter strings in a free-recall procedure as a dependent measure and found an advantage of grammatical strings over random sequences of letters. One might argue that free-recall performance is an indirect measure in that context. However, this conclusion is complicated by Miller’s “explicit encoding explanation” (Reber, 1967, p. 859) of the effect. Miller assumed that subjects use the systematicities they detect in the letter strings to recode the strings and remember them in coded form.

2 Note, however, that memorization of letter strings does not appear to be a necessary condition for chunking to occur. Reber and Allen (1978) had subjects simply “pay the utmost attention” (p. 196) to grammatical letter strings. They reported that subjects commented to have formed “‘groups’ or ‘chunks’ [of letters] which made ‘scanning easier’ ” (p. 200; see also Gordon & Holyoak, 1983, Experiment 1).
tual fluency, which, in turn, is supposed to be the “measurable underpinning of the feeling of familiarity” (p. 3; see also Jacoby, 1983; Jacoby & Dallas, 1981; Jacoby, Kelley, & Dywan, 1989; Jacoby & Whitehouse, 1989; Jacoby & Witherspoon, 1982; Jacoby, Woloshyn, & Kelley, 1988; Joordens & Merikle, 1992; Whittlesea, Jacoby, & Girard, 1990). If an item can be identified well under impoverished presentation conditions, then perceptual fluency and, hence, the feeling of familiarity for this item will be high, and vice versa.

Mandler (1980; for a recent review and extension, see Mandler, 1991) suggested that increasing degrees of intraitem organization affect the feelings of familiarity associated with an item. If one assumes that how well an item is identified depends on its intraitem organization, then it is relatively easy to see why the competitive chunking model predicts better and/or faster identifications of grammatical as opposed to nongrammatical strings under impoverished presentation conditions (e.g., in a perceptual clarification procedure). For grammatical strings, fewer chunks will be needed in the process of perception. The number of these chunks “is a measure of how compact the representation of a stimulus is” (Servan-Schreiber & Anderson, 1990, p. 601), which closely corresponds to Mandler’s (1980) concept of intraitem organization. Also, intraitem organization, according to Mandler (1980), is an automatic process that integrates the specific features of an item. Analogously, Servan-Schreiber and Anderson assumed that whenever an item is attended, chunks are formed automatically.

Note that the competitive chunking model is currently the only model of grammar learning that allows one to directly derive the prediction that grammatical strings should be identified faster than nongrammatical strings in a perceptual clarification procedure. The most obvious reason for this state of affairs is that other models (e.g., Dulany et al., 1984; Mathews, 1990, 1991; Perruchet & Pacteau, 1990) do not explicate that and how knowledge about permissible groups of letters, or about permissible relative spatial relations between letters, is used in the process of perceiving new strings. One could argue that Brook’s (1978, 1987) model comes closest to predicting faster identifications of grammatical as opposed to nongrammatical strings. In his model, it is assumed that instances of grammatical strings are stored in memory and used for later grammaticality judgments on the basis of the similarity between a given string and the stored exemplars. A similarity match would, on average, be reached more rapidly for grammatical as opposed to nongrammatical strings. However, this would seem to require that a complete representation of a string must be established before the comparison process could start. Hence, influences of grammatical status on perceptual identification do not seem to be a natural entailment of this theory.

In any case, the prediction of the competitive chunking model of faster identifications of grammatical as opposed to nongrammatical strings has not been put to an empirical test before. If this prediction turns out to be valid, then the perceptual fluency factor can be used to distinguish between different theoretical accounts of grammaticality judgments. Faster identifications should be accompanied by more intense feelings of familiarity than slower identifications. Thus, if greater familiarity increases the probability of a string being judged as grammatical, as is predicted by the competitive chunking model, then faster identifications in a perceptual clarification procedure should covary with “grammatical” responses, and slower identifications in a perceptual clarification procedure should covary with “nongrammatical” responses, irrespective of the objective grammaticality status of a string. If no such effect is observed, then in analogy to the rationale underlying two-process models of recognition memory, grammaticality judgments may be based on search processes, as is implicated by models that assume explicit knowledge as the basis for grammaticality judgments (Dulany et al., 1984, 1985; Perruchet & Pacteau, 1990).

Overview of the Present Experiments

The two experiments reported in this article consisted of a memorization and a test phase. During the memorization phase, subjects learned and reproduced a subset of the strings generated by a finite-state grammar. During the test phase, grammatical and nongrammatical strings were presented in a perceptual clarification procedure. Initially, a black square covered a letter string. Parts of the square were removed continuously, such that the string became more and more visible. Subjects pressed a key to indicate when they had identified a particular string. Immediately after the key was pressed, the string disappeared and subjects were asked to perform one of several direct judgments.

Experiment 1 was designed for two purposes. First, it was intended to test the prediction derived from Servan-Schreiber and Anderson’s (1990) competitive chunking model that grammatically correct strings would be identified faster than nongrammatical strings in a perceptual clarification procedure. Second, presenting strings in a perceptual clarification procedure and requiring their identification rather than grammaticality judgments can be conceived as an indirect measure of the knowledge subjects have acquired. It seemed, therefore, important to explore further whether instructions about the regularities underlying the strings and about the relation between study and test phases influenced the identification times. After memorization and before the test phase, one group of subjects (henceforth referred to as Group I) was informed that the strings they had memorized had been constructed regularly and that one half of the test-phase strings had been constructed by using the same rules, whereas the other half of the test-phase strings would violate the rules. Another group of subjects (henceforth referred to as Group N) was not given that information.

To control performance in the identification task, the perceptual clarification procedure was followed by a matching task for each item. On this matching task, subjects, confronted with a letter string, judged whether or not this string was the one they had just identified in the perceptual clarification procedure (i.e., they matched the string held in short-term memory to the new string). If subjects can identify grammatical and nongrammatical items equally well, then no difference for

---

3 I am grateful to Arthur S. Reber for pointing this out.
both discrimination performance and response bias should appear between item types.

Experiment 2 was designed both to conceptually replicate the findings obtained with the indirect test used in Experiment 1 and to test specifically the assumption of the competitive chunking model that familiarity mediates grammaticality judgments. For these purposes, two groups of subjects participated in this experiment, the groups differing with respect to the generating finite-state grammar (Grammar A vs. Grammar B, see Figure 1). Furthermore, either a recognition or a grammaticality judgment was required after identification. On the recognition task, subjects were asked simply to indicate whether or not the string they had identified had been memorized during the first phase of the experiment. This way, the recognition judgments were parallel in format to the grammaticality judgments that required the subjects to indicate whether or not an identified string was grammatical.

If perceptual fluency plays a role in this task, then "old" responses on the recognition judgments should covary with faster identification times and "new" responses on the recognition judgments should be accompanied by slower identification times, irrespective of the objective episodic status of an item. If this turns out to be the case, then the analogous analysis permits conclusions about whether grammaticality judgments are also influenced by perceptual fluency and, hence, familiarity on this task.

Experiment 1

Method

Subjects

Subjects were 51 female and 21 male undergraduate students at Bonn University, who received course credit for participating in the experiment. Subjects ranged in age from 19 to 34 years (M = 22.94; SD = 3.41); they were assigned at random to one of the two experimental conditions: Group I (subjects who were explicitly informed about the regular nature of the letter sequences before test and about the relation between study and test stimuli) and Group N (subjects who were not informed about this aspect of the task).

Materials

Stimuli and apparatus. The finite-state grammar that was used to generate the letter strings for this experiment (Grammar A, see Figure 1) was equivalent to grammars used by Reber (1976) and by Reber et al. (1980). This grammar generated 43 grammatical strings of three to eight letters. Three randomly selected strings were discarded from the set of experimental strings. Twenty strings were randomly selected to be memorized during the first phase, and 20 strings were selected as new test strings for the second phase (see Appendix A). For each of the correct test strings, an incorrect test string was constructed by replacing one letter in the string with a letter from the alphabet of the grammar that was illegal at the given position. For example, the grammatical string TXTVTPS was changed to TXVTVPS, which is nongrammatical because of the V at the third position of the string. All five letters of the alphabet were used equally often in constructing nongrammatical strings. The distribution of incorrect letters over string positions was as follows: Incorrect letters occurred three times at Positions 1, 2, 4, 5, and 6; twice at Positions 3 and 7; and once at Position 8. Three of these incorrect letters were located at the terminal positions of strings.

In the memorization phase, strings were presented at the center of a 9-in. diagonal video screen controlled by a Macintosh SE microcomputer. At a viewing distance of about 40 cm, each letter was about 0.7° high and 0.4° wide. Strings were presented in seven sets of three at a time (except for the last set, which comprised only two strings). The string remained visible until subjects indicated that they had memorized them sufficiently for immediate recall of the entire set. Next, a "keyboard" appeared at the center of the screen. Although only five different letters were actually used by the grammar, the keyboard displayed all of the letters of the alphabet. Subjects used this keyboard to reproduce the current set of strings by clicking, with the computer mouse, into the symbols for individual letters. The selected letters were inserted into fields displayed below the keyboard. Corrections could be made for any letter at any time.

If subjects made one or more errors in recalling a set of strings, the entire set was presented again for memorization. If all of the strings of a given set were reproduced correctly, the next set of strings was presented for memorization. These next strings were selected randomly from the set of the remaining strings. On average, subjects needed about 4 min to memorize and recall one set of grammatical strings.

The strings to be displayed in the perceptual clarification procedure were selected at random from the set of the remaining strings. At the beginning of each trial, a black square appeared on the screen. The square was, at every edge, about 0.3° larger than the letter string it covered. Black pixels were removed continuously and at random locations within the square, such that a letter string covered by the square became more and more visible. When subjects indicated that they believed they had identified a particular string, the screen was cleared and a new string was displayed. Subjects indicated whether or not the new string was the one they had just seen during the identification task. Targets and distractors were selected randomly on this task, with the following restrictions: On 50% of the trials of the matching task, the correct string was displayed (i.e., the string that had appeared in the preceding identification trial). On the remaining trials, a different string was displayed. Both grammatical and nongrammatical strings were selected randomly from the set of the remaining grammatical and nongrammatical strings, respectively.

Each grammatical and nongrammatical test string was presented twice during the test phase. Therefore, the test phase comprised a total of 80 identification-plus-matching trials (2 • [20 grammatical + 20 nongrammatical strings]). For one half of the test strings from each set of 20 strings, the first presentation in the identification task was followed by a presentation of the correct string in the matching task, and the second presentation in the identification task was followed by a presentation of a different (grammatical or nongrammatical) string in the matching task. For the other half of the strings from each set, the sequence was reversed.

Anticipatory (< 0.4 s) and delayed responses (> 20 s) were scored as errors, and these trials were, unbeknownst to subjects, repeated later in the experiment. Reaction times were measured in milliseconds by using Drexel's millimeter (Westall, Perkey, & Chute, 1989).

Procedure. Subjects were tested one at a time. Standardized instructions were read to subjects. Relevant passages were repeated if this was requested. Subjects were told that their task was to memorize strings of consonants, in sets of three strings at a time (except for the last set, which comprised only two strings), and to reproduce them. The learning criterion was set to three correct reproductions for each set of strings. The subsequent test phase was not mentioned at this point. Before memorizing strings, subjects received computer-instructed training on how to use the "screen keyboard" for reproducing strings of letters.

After having completed the memorization task, subjects received
Design

The main dependent variables were (a) the reaction times in the identification task and (b) accuracy in the matching task. Independent variables were (a) grammatical status of the letter strings (grammatical vs. nongrammatical: within-subject), (b) whether an item was presented for the first or the second time during test (within-subject), and (c) whether subjects were explicitly informed about the regularities of the letter sequences before test and about the data presented during the identification task.

Given the total sample size and a = .05, it was possible to detect "large effects" (Cohen, 1977) for the type of instruction between-subjects comparisons. Therefore, a was set to .05 for all statistical tests reported, and individual p values are omitted. For reliable effects, partial squared multiple correlations ($R^2$) are reported as a measure of relative effect size, that is, the proportion of variance explained by the effect relative to the total variance not explained by other experimental variables (cf. Cohen, p. 412). A multivariate approach was used for all within-subject comparisons (O'Brien, & Kaiser, 1985). As a consequence, no mean square error values are reported for within-subject variables with more than two levels. The Pillai-Bartlett $V$ was used as test criterion, but $F$ approximations to $V$ are reported (Olson, 1976).

Results

On average, subjects needed 14.4 trials and about 28 min to memorize the 20 grammatical strings. There were no differences between Group I and Group N in the number of learning trials to criterion and in learning time (both $r < 1$).

The presentation of results from the test phase is organized into two main sections. The first section presents the analysis of subjects' matching performance during the test phase. The second section contains the analysis of the identification times in the perceptual clarification procedure.

Matching Performance

Discrimination ($P_r$) and response bias ($B_r$) indices were computed according to the two-high threshold model (Snodgrass & Corwin, 1988) for the matching task. $P_r$ ranges from 0 (indicating no sensitivity) to 1 (indicating perfect discrimination); for $B_r$ .50 indicates neutral response bias.

On average, subjects' discrimination performance on the matching task was quite good ($P_r = .78$, see Figure 2), and their response bias appeared to be approximately neutral ($B_r = .46$). A $2 \times 2 \times 2$ multivariate analysis of variance (MANOVA), with type of instruction (Group I vs. Group N) as the between-subjects variable and grammatical status (grammatical vs. nongrammatical) as well as order of presentation (first vs. second presentation of a string during test) as the within-subject variables on subjects' $P_r$ values for the matching task, revealed that there was no reliable difference between Groups I and N, $F(1, 70) < 2.03$; no reliable main effect of grammatical status, $F(1, 70) < 2.36$; and no reliable main effect of order of presentation, $F(1, 70) < 1.44$. Furthermore, there were no reliable interactions between these factors, all $Fs(1, 70) < 2.05$. Also, the $P_r$ values were reliably different from zero, $F(2, 70) = 2.348.05$. $MS_e = 0.14$, $R^2 = .99$, confirming the above-chance discrimination between old and new letter strings.

The same analyses for the bias index revealed that $B_r$ values did not differ between groups ($F < 1$), and that the average bias was not different for grammatical and nongrammatical items ($F < 1$) or for both presentations of a string, $F(1, 70) < 3.56$. However, the interaction between type of instruction and order of presentation was reliable, $F(1, 70) < 5.57$. $MS_e = 0.03$, $R^2 = .07$. This interaction is due to the fact that although $B_r$ did not vary with order of presentation for Group I, the response bias became more conservative for Group N. At present, I do not have a ready explanation for this effect. No other interaction in this analysis was reliable (all $Fs < 1$). The test of whether the overall $B_r$ was reliably different from $.50$ was not reliable, $F(2, 70) = 1.99$, indicating neutral average response bias (see Figure 2).

To summarize, subjects showed relatively good performance on the matching task. Instructions informing subjects about the regularities did not influence matching accuracy. Also, matching performance was not affected by the grammatical status of the string presented in the identification task.

Identification Task

Grammatical status of strings was predicted to influence the speed of identification. In addition, instructions informing subjects about the regularities should not interact with performance if the identification task is to be conceived as an indirect measure of grammatical knowledge. Reaction times for the identification task are displayed in Figure 3. The number of anticipatory and delayed responses was negligible ($<1\%$).

---

4 All $P_r$ and $B_r$ values have been adjusted, as suggested by Snodgrass and Corwin (1988, p. 35).
INDIRECT EFFECTS OF SYNTHETIC GRAMMAR LEARNING

A 2 x 2 x 2 MANOVA, with type of instruction (Group I vs. Group N) as the between-subjects variable and grammatical status (grammatical vs. nongrammatical) as well as order of presentation (first vs. second presentation of a string during test) as the within-subject variables on subjects' mean reaction times, revealed reliable main effects of grammatical status, $F(1, 70) = 12.07, MS_e = 136,426.07, R^2_p = .15$, and of order of presentation, $F(1, 70) = 75.73, MS_e = 438,788.76, R^2_p = .52$. The difference in reaction times between Groups I and N was not reliable ($F < 1$). In addition, no interaction in this analysis was reliable, all $Fs(1, 70) < 3.53$. This latter result indicates, first, that the critical finding of faster identifications of grammatically correct as opposed to grammatically incorrect strings did not change during the test phase and, second, that this difference was not influenced by whether or not the instructions explicitly informed subjects about the regularities in the strings and the study-test relation.

For exploratory purposes, an additional analysis was performed to see whether the differences between grammatical and nongrammatical strings varied as a function of string length. As a result of the small number of strings with a length of three to five letters, these strings were combined into a single category. A 4 x 2 MANOVA, with string length (three to five vs. six vs. seven vs. eight letters) and grammatical status (grammatical vs. nongrammatical) as the within-subject variables, revealed reliable main effects of string length, $F(3, 69) = 66.43, R^2_p = .74$, and grammatical status, $F(1, 71) = 11.75, MS_e = 295,615.93, R^2_p = .14$, as well as a reliable interaction between these two variables, $F(3, 69) = 3.34, R^2_p = .13$. Figure 4 illustrates that this interaction is due to the lack of a difference in reaction times between grammatical and nongrammatical strings of a length of three to five letters.

To summarize, reaction times for identifications in the perceptual clarification procedure were affected by the grammatical status of the string to be identified. In contrast, although less time was required to identify grammatical as opposed to nongrammatical strings, both types of strings were
recognized equally well immediately after identification, and response bias was approximately neutral. Thus, the same information about the identity of a string was accumulated in less time for grammatical strings in contrast to nongrammatical strings.

Furthermore, whether or not the instructions explicitly stated that the strings from the memorization and test phases were constructed according to the same set of rules did not influence performance. This effect did not diminish during the test.

Strings were identified faster the second time they were presented in the clarification procedure. This is most readily explained by subjects getting acquainted with the test procedure, but there may have been additional item-specific repetition effects (cf. Kolers & Roediger, 1984). In any case, it is important for the present purposes that this speedup did not influence the difference in identification times between grammatical and nongrammatical strings.

Discussion

The results of Experiment 1 were rather clear-cut. First, subjects performed well on the matching task. This indicates that they performed as instructed on the identification task and took the time necessary to identify a string properly. Grammatical status of the string to be identified did not influence the matching performance with respect to both grammatical discrimination and response bias. It did, however, influence reaction times in the identification task. Subjects took less time to identify a grammatical string as opposed to a nongrammatical string. This is exactly what Servan-Schreiber and Anderson’s (1990) competitive chunking model predicted. A new grammatical string appears less complex and more integrated than a nongrammatical string because its perceptual representation is composed of larger chunks created during memorization. Hence, the grammatical string is easier to perceive and identification is more fluent. In contrast, strings that violate the rules of the generating grammar appear more complex and less integrated because more elementary subunits have to be applied to identify nongrammatical letter strings. As a consequence, identification is more effortful and less fluent. The effect appears to be absent for relatively short strings. The next step, then, is to test whether perceptual fluency is used implicitly for direct recognition and grammaticality judgments.

Experiment 2

Experiment 2 was designed for two purposes. First, it was intended to conceptually replicate the finding from Experiment 1 that grammatical strings are identified faster than nongrammatical strings when presented in a perceptual clarification procedure, under the conditions of a different generating grammar and when different judgments are required after identification. Second, Experiment 2 directly tested the claim that subjective familiarity of strings mediates grammaticality judgments. Faster identification of grammatical strings should lead to the impression of greater perceptual fluency and, hence, greater familiarity of the identified string. To investigate the influence of identification speed on later direct judgments, subjects were required to answer one of two questions after each identification. Their task was to indicate either whether the identified string had occurred during memorization (recognition judgment) or whether the string was grammatical or nongrammatical (grammaticality judgment). The former judgments can be used to test whether the often-reported influence of familiarity on recognition judgments can be found for the present task (i.e., whether faster identifications are found with both correct and incorrect “old” responses, and whether slower identifications covary with correct and incorrect “new” responses). If this validity check turns out to be positive, then subsequent analyses can be
performed to test whether there is a similar influence of familiarity on grammaticality judgments (i.e., whether faster identifications covary with correct and incorrect “grammatical” responses and whether slower identifications covary with correct and incorrect “nongrammatical” responses).

**Method**

**Subjects**

Subjects were 40 female and 21 male undergraduate students at Bonn University, who received course credit for participating in the experiment. Subjects ranged in age from 19 to 47 years (M = 24.7; SD = 6.2); they were assigned at random to one of the two experimental conditions: Group A (subjects who memorized and later identified strings generated by Grammar A; see Figure 1) and Group B (subjects who memorized and later identified strings generated by Grammar B, see Figure 1). None of the subjects had participated in Experiment 1.

**Materials**

**Stimuli and apparatus.** Two different finite-state grammars were used to generate the letter strings for this experiment: Grammar A, which had already been used in Experiment 1, and Grammar B (see Figure 1). Subjects in Group A first memorized and later judged strings generated by Grammar A, whereas subjects in Group B performed the same tasks with Grammar B strings. Grammar B was constructed so that it could generate approximately the same number of strings (42) with a length of three to eight letters, as Grammar A (43). Yet the number of initial and the number of terminal letters was larger in Grammar B than in Grammar A, and Grammar B could not generate strings containing salient runs. This manipulation was aimed at exploring whether the basic findings about differences in identifications of grammatical as opposed to nongrammatical strings can be generalized to a different finite-state grammar and to assess globally whether salient features, such as initial and terminal letters or runs, make a difference in either learning or test performance, or both.

As in Experiment 1, 40 strings were selected randomly from each grammar and assigned to the set of the 20 study strings and to the set of 20 correct test strings (in fact, Grammar A strings were identical in Experiments 1 and 2; see Appendixes A and B). Incorrect strings were generated as in Experiment 1. The distribution of incorrect letters over string positions for Grammars B and A were, respectively, as follows: Incorrect letters occurred three times at Positions 1, 2, 3, 4, and 6; twice at Positions 5 and 7; and once at Position 8. For Grammar A, incorrect letters occurred three times at Positions 1, 2, 4, 5, and 6; twice at Positions 3 and 7; and once at Position 8. As for Grammar A, three incorrect letters were located at the beginning, and three at the terminal position of the strings generated by Grammar B.

The memorization phase was identical to Experiment 1. The identification task differed, however, in that (a) both the strings from the memorization phase and the new grammatical and nongrammatical strings were presented for identification and (b) every identification of a letter string was followed either by a recognition judgment or by a grammaticality judgment. After subjects had indicated identification, they were asked either whether the identified string had occurred during the memorization phase or whether it was grammatical or whether it had occurred during the memorization phase. Subjects indicated, by hitting a “yes” or “no” button in a screen dialog, what they thought to be the correct answer.

Every 10 trials, subjects were asked to take a rest break for as long as they wished. During the rest break, the computer displayed both the subject's mean reaction times and an adjective describing the overall performance on the recognition and grammaticality judgments along a scale ranging from fair to excellent. The entire experiment lasted about 83 min (between 53 and 137 min) for both experimental groups.

**Design**

The main dependent variables were (a) reaction times in the identification task, (b) accuracy in the recognition judgments, and (c) accuracy in the grammaticality judgments. Independent variables were (a) grammatical status of the letter strings (grammatical vs. nongrammatical, within subject), (b) episodic status of the letter strings (previously memorized vs. new at test, within subject), (c) whether an item was presented for the first or the second time during test (within subject), and (d) whether subjects memorized strings generated by Grammar A (Group A: 30 subjects) or by Grammar B (Group B: 31 subjects).

Given the total sample size and (a) = .05, it was possible to detect "large effects" (d = 0.85; see Cohen, 1977) for the type of grammar between-subjects comparisons. Therefore, (a) was set to .05 for all statistical tests reported, and individual p values are again omitted.

**Results**

The presentation of results is organized into four main sections. In the first section, results from the memorization phase are reported. The second section contains the analysis of the identification times in the perceptual clarification procedure, whereas the third section presents performance on the recognition and grammaticality judgments. Finally, in the fourth section, the relationship between identification times and the judgment tasks is investigated.
Identification Task

New grammatical and nongrammatical strings, on the one hand, were not possible (see Appendixes A and B) as the between-subjects variable and grammatical status (new grammatical vs. nongrammatical strings) as well as order of presentation (first vs. second presentation of a string during test) as the within-subject variables on subjects’ mean reaction times, revealed reliable main effects of grammatical status, \(F(1, 59) = 9.98, MS_e = 328,192.26, R^2_p = .14\), and of order of presentation, \(F(1, 59) = 53.26, MS_e = 363,388.28, R^2 = .47\). The difference in reaction times between Groups A and B was not reliable, \(F(1, 59) = 2.97\). In addition, no other interaction in this analysis was reliable (all \(Fs < 1\)). Thus, grammatical strings are identified faster than nongrammatical strings, independent of the type of grammar used. In addition, the strings were identified faster the second time that they were presented in the clarification procedure, but this speedup did not influence the critical difference between grammatically correct and incorrect strings. The speedup is—as in Experiment 1—most readily explained as a general training effect, although there may have been additional item-specific repetition effects.

For exploratory purposes, an additional analysis was performed to see whether the differences between grammatical and nongrammatical strings varied as a function of string length. As a result of the small number of strings with a length of three to five letters, these were combined into a single category, as was done in Experiment 1. A 4 x 2 MANOVA, with string length (three to five vs. six vs. seven vs. eight letters) and grammatical status (new grammatical vs. nongrammatical strings) as the within-subject variables, revealed reliable main effects of string length, \(F(3, 58) = 78.64, R^2 = .80\), and grammatical status, \(F(1, 60) = 10.39, MS_e = 694,375.48, R^2 = .15\), as well as a reliable interaction between these two variables, \(F(3, 58) = 11.28, R^2 = 0.37\). Figure 6 illustrates that the interaction was due to larger differences in reaction times for longer strings. This result is similar to the finding from Experiment 1, in which differences in reaction times between grammatical and nongrammatical strings were present for strings with a length of six to eight letters, but not for shorter strings. This difference in reaction times is discussed in the General Discussion section.

Figure 6 also displays the mean identification times for strings that had to be memorized during the first phase of the experiment. It appears that subjects took about as long to respond to these previously memorized strings as it took for them to react to new strings that were grammatically correct. Note, however, that a direct comparison between the previously memorized strings, on the one hand, and the grammatical and nongrammatical new strings, on the other hand, was not possible because they formed disjunctive sets. In contrast, the new grammatical and the new nongrammatical strings were identical, with the exception of one single letter that was replaced by a rule-violating letter in each nongrammatical string.

Recognition and Grammaticality Judgments

For both recognition and grammaticality judgments, discrimination (\(P_r\)) and response bias (\(B_r\)) indices were again computed according to the two-high threshold model (Snodgras &
Corwin, 1988). One complicating factor in the present situation was that recognition and grammaticality judgments were not necessarily independent. Obviously, old items (i.e., strings from the memorization phase) must be grammatical, and nongrammatical items must be new (because only correct strings were memorized). Therefore, performance measures for recognition judgments were based exclusively on grammatical strings, and performance measures for grammaticality judgments were based exclusively on strings that had not been memorized during the first phase of the experiment. Figure 7 displays the average performance of both groups, separately, for the first and the second presentation of strings during the test phase.

A 2 × 2 MANOVA, with type of grammar (Grammar A vs. Grammar B) as the between-subjects variable and order of presentation (first vs. second presentation of a string during
test) as the within-subject variable on subjects’ $P_r$ values for the recognition judgments, revealed that there was no reliable difference between groups ($F < 1$) but a reliable main effect of order of presentation, $F(1, 59) = 7.31$, $MS_e = 0.020$, $R^2_s = .11$, indicating lower discrimination performance when a particular string was presented for the second time. However, a reliable interaction between type of grammar and order of presentation, $F(1, 59) = 6.08$, $MS_e = 0.020$, $R^2_s = .09$, points to the fact that this performance decrease was present primarily in Group A. The $P_r$ values were reliably different from zero, $F(2, 59) = 44.78$, $MS_e = 0.080$, $R^2_s = .60$, indicating above-chance discrimination between old and new letter strings. The same analyses for the bias index revealed that $B_r$ values did not differ between groups, $F(1, 59) < 1.34$, but the bias became more neutral during the test phase, $F(1, 59) = 4.80$, $MS_e = 0.023$, $R^2_s = .08$. The interaction between these two variables was not reliable ($F < 1$). Overall, $B_r$ was not reliably different from .50, $F(2, 59) < 1.75$, indicating overall neutral response bias for the recognition judgments.

Parallel analyses were performed for the grammaticality judgments. A 2 x 2 MANOVA, with type of grammar (Grammar A vs. Grammar B) as the between-subjects variable and order of presentation (first vs. second presentation of a string during test) as the within-subject variable on subjects’ $P_r$ values for the grammaticality judgments, showed that none of the main effects or the interaction were reliable, $F(1, 59) < 2.25$. Again, the $P_r$ values were reliably different from zero, $F(2, 59) = 52.80$, $MS_e = 0.11$, $R^2_s = .64$, which reflects above-chance discrimination between grammatical and non-grammatical letter strings. The same analyses for the bias index revealed that $B_r$ values did not differ between groups ($F < 1$) but that, parallel to the recognition judgments, bias for the grammaticality judgments became more neutral during the test phase, $F(1, 59) = 7.42$, $MS_e = 0.012$, $R^2_s = .11$. The interaction between type of grammar and order of presentation was not reliable, $F(1, 59) < 2.47$. As with the recognition judgments, $B_r$ was not reliably different from .50 for the grammaticality judgments, $F(2, 59) < 2.22$, indicating, again, neutral average response bias.

To summarize, subjects showed above-chance performance on both the recognition judgments and the grammaticality judgments. Thus, it was possible to analyze the relationship between these judgment tasks and the identification times from the perceptual clarification procedure preceding the recognition and grammaticality judgments.

**Relationship Between Identification Times and Judgment Tasks**

First, it is important to establish, for the present experimental situation, that faster reaction times in the identification task can be interpreted as indicating perceptual fluency, which is known to influence subsequent recognition judgments (e.g., Johnston et al., 1985, 1991). If this test turns out positive under the current conditions, then it makes sense to investigate the postulated relation between reaction times and grammaticality judgments. Therefore, the first part of this section contains the results for the identification-time/grammaticality-judgment patterns, and the second part contains the results for the identification-time/grammaticality-judgment patterns.

To compare only homogeneous sets of strings, the analysis involving the recognition judgments was first performed for false alarms and correct rejections (the “new” items), then for hits and misses (the “old” items from the memorization phase). If there was an influence of fluency of identification, it should have been strongest for items that had not previously been memorized (i.e., for the grammatical and nongrammatical new items). In contrast, old items might occasionally be recognized directly as a consequence of a successful memory search process, which should reduce the overall influence of perceptual fluency. Therefore, the comparison between false alarms and correct rejections (i.e., reactions to strings from the set of the new items) is most critical.

In all subsequent analyses, string length was entered as a separate within-subject variable. This was necessary because identification times differed considerably as a function of string length, such that averaging reaction times across different string lengths for subjects with missing data for certain judgment status by string length combinations would have caused unpredictable distortions. For example, if for a particular subject there were no hits for strings with a length of eight characters, then the mean reaction time associated with hits would have underestimated this subject’s true reaction times for that category. As a consequence, subjects with missing data for certain judgment status by string length combinations had to be excluded from further analyses.

Mean reaction times for identifications that were followed by a recognition judgment are presented in Figure 8 for both groups of subjects. A 2 x 2 x 4 MANOVA ($N = 43$), with type of grammar (Grammar A vs. Grammar B) as the between-subjects variable and judgment status (false alarms vs. correct rejections) as well as string length (three to five vs. six vs. seven vs. eight letters) as the within-subject variables on subjects’ mean reaction times, showed no reliable main effect of type of grammar ($F < 1$), but the main effect of judgment status was reliable, $F(1, 41) = 11.70$, $MS_e = 2.450,507$, $R^2_s = .32$. “Old” responses are found with faster identifications, which is what one would expect if fluency of identification influenced later recognition judgments as predicted. The main effect of string length was also reliable, $F(3, 39) = 57.66$, $R^2_s = .82$, as was the interaction between string length and judgment status, $F(3, 39) = 7.15$. The latter result is probably due to the reversed judgment status effect for strings with a length of three to five letters in Group B. No other interaction in this analysis was reliable.

A 2 x 2 x 4 MANOVA ($N = 25$), with type of grammar (Grammar A vs. Grammar B) as the between-subjects variable and judgment status (hits vs. misses) as well as string length (three to five vs. six vs. seven vs. eight letters) as the within-subject variables on subjects’ mean reaction times, showed, as before, no reliable main effect of type of grammar ($F < 1$). The main effect of judgment status also failed to reach the preset $\alpha$ criterion, $F(1, 23) = 3.21$, $p > .086$, although six out of eight mean reaction times were in the predicted direction. However, it should be mentioned that the reduced sample size in this analysis implies an increase in the size of the effects that can be detected with this analysis as
compared with the previous analysis. Finally, there was again a reliable effect of string length, $F(3, 21) = 31.83, R^2_p = .82$, and no interaction in this analysis was reliable.

To summarize this first part, speed of identification seems to have an influence on subsequent recognition judgments. The findings are consistent with existing research that emphasizes the role of indirect effects of perceptual fluency and familiarity on direct judgments of episodic status (Johnston et al., 1985, 1991).

Turning now to the analysis of the relationship between identification times and subsequent grammaticality judgments, corresponding analyses were first performed for the nongrammatical items (false alarms vs. correct rejections) and then for the grammatical items (hits vs. misses). A $2 \times 2 \times 4$ MANOVA ($N = 18$), with type of grammar (Grammar A vs. Grammar B) as the between-subjects variable and judgment status (false alarms vs. correct rejections) as well as string length (three to five vs. six vs. seven vs. eight letters) as the within-subject variables on subjects' mean reaction times, showed no reliable main effect of type of grammar ($F < 1$). Again, the main effect of judgment status failed to be reliable ($F < 1$), indicating no influence of fluency on grammaticality judgments. Finally, there was a reliable effect of string length, $F(3, 35) = 51.67, R^2_p = .82$, and no interaction in this analysis was reliable.

Discussion

The main results of Experiment 2 can be summarized as follows. First, it was possible to replicate the basic finding from Experiment 1 that, as predicted by the competitive chunking model, grammatical strings are identified faster than nongrammatical strings when presented in a perceptual clarification
procedure. The effect is reliable both when different direct judgments are required after identification and when a different finite-state grammar is used to generate strings lacking some of the salient features of Grammar A.

Second, consistent with the existing literature on the influence of perceptual fluency on judgments of episodic status, faster identifications were related to "old" responses of a subsequent recognition judgment. The effect was most pronounced for new items (i.e., false alarms were related to shorter reaction times than correct rejections), which is plausible because these strings had not been encoded during memorization and, hence, feelings of familiarity could exert maximal influence relative to memory search processes. For strings that had been presented for memorization in the first phase of the experiment, the effect failed to reach the present α level, although, at a descriptive level, there appeared to be a consistent difference between "old" and "new" responses with respect to reaction times. This result may be (partly) due to the somewhat lower power of the relevant statistical test. However, such a finding is also plausible because, according to two-process theories of recognition judgments (Gillund & Shiffrin, 1984; Johnston et al., 1985, 1991; Mandler, 1980), memorized strings may be retrieved by search processes. In the present experiments, the learning criterion was relatively strict, requiring three perfect recalls in succession of each group of strings to be memorized. Under such conditions, even material as artificial as consonant strings may later be retrievable by memory search processes on a delayed test. In any case, these results are important because they confirm that identification times can be regarded as a valid indicator of feelings of familiarity in the present experimental situation.

Third, analogous relations to speed of identification have not been found for the grammaticality judgments. This is so in spite of the fact that grammatical strings are identified faster than nongrammatical strings. Thus, the prediction of the competitive chunking model that feelings of familiarity mediate grammaticality judgments could not be confirmed in the present experiment.

It is interesting to note that recognition and grammaticality judgments show equal levels of discrimination performance (mean $P_r = .18$, $SD = 0.17$, and mean $P_r = .23$, $SD = 0.14$, for the recognition and grammaticality judgments, respectively), $t(1, 60) = 1.93$. This renders it unlikely that differences between these judgments are due to differences in the difficulty of performing these tasks. However, one must be cautious in
drawing such comparisons because both judgments are not based on identical item samples.

General Discussion

The present experiments show that an indirect test can be used to discriminate between strings generated by a finite-state grammar and strings that violate the rules of that grammar. When presented in a perceptual clarification procedure, grammatical strings are identified faster than nongrammatical strings. In contrast to direct grammaticality judgments, the identification task does not require that subjects are informed about the regularities in the strings and about the relationship between the study and test situations.

More important, the reported differences in identification times for grammatical as opposed to nongrammatical letter strings are consistent with predictions of the competitive chunking model proposed by Servan-Schreiber and Anderson (1990). Moreover, it is the only model of grammar learning that directly predicts these effects for an identification task, in that it relates processes relevant during memorization of strings to later perception of new strings. According to this process model, letter chunks of growing lengths are formed during memorization of grammatical strings. The compactness (or intratext organization in terms of two-process models of recognition memory) of the mental representation of any given string is assumed to be a monotone function of the number of strings needed in the process of perception. Grammatical strings, which should have, on average, higher degrees of intratext organization, should be perceived more easily and, hence, faster under impoverished presentation conditions. This is exactly what has been found in the present experiments.

Notice that the critical comparison was between new grammatical and nongrammatical strings in both experiments. As mentioned before, these two sets of strings were directly comparable because each string in the set of grammatical strings was identical to one string in the set of nongrammatical strings, with the exception of one single letter, and none of these strings had been experienced by the subjects before they were presented for identification. Furthermore, the critical effects did not depend on the type of grammar that generated the strings. However, to empirically support the presupposition that there were no a priori differences between grammatical and nongrammatical strings that could potentially influence identification performance, an additional group of subjects was run and confronted with the test phase of Experiment 2 only. These subjects were told that a different group of participants in the experiment had begun by memorizing grammatical strings and that the task was to perform as well as possible on the grammaticality and the recognition task without previous string memorization. A $2 \times 2 \times 2$ MANOVA, with type of grammar (Grammar A vs. Grammar B) as between-subjects variable and grammatical status (grammatical vs. nongrammatical) as well as order of presentation (first vs. second presentation of a string during test) as within-subject variables on subjects' mean reaction times, revealed no reliable main effects of grammatical status ($F < 1$). Average response times were 7,338 ms for grammatical and 7,379 ms for nongrammatical strings. A reliable effect of order of presenta-

The differences in reaction times tended to be larger for longer than for shorter strings. At least two factors could explain this finding. First, it could simply be that, for short strings in the present experiments, identification is relatively fast, so that a ceiling effect prevents the grammaticality manipulation to be reflected in the reaction times. Second, one could assume that chunks do not only contain information about the elements they are composed of, but also provide positional order information (Johnson, 1970; Tulving, 1962). For instance, each chunk could contain information about the probability of occurrence of a particular subsequent chunk. That subsequent chunk would then benefit from the expectation induced by the previous chunk. If the grouping structure is violated at a particular position, identifying subsequent elements will not be facilitated by such expectations. On average, this effect should be more pronounced the longer the sequence of elements after the violated position.

Notice that the critical comparison was between new grammatical and nongrammatical strings in both experiments. As mentioned before, these two sets of strings were directly comparable because each string in the set of grammatical strings was identical to one string in the set of nongrammatical strings, with the exception of one single letter, and none of these strings had been experienced by the subjects before they were presented for identification. Furthermore, the critical effects did not depend on the type of grammar that generated the strings. However, to empirically support the presupposition that there were no a priori differences between grammatical and nongrammatical strings that could potentially influence identification performance, an additional group of subjects was run and confronted with the test phase of Experiment 2 only. These subjects were told that a different group of participants in the experiment had begun by memorizing grammatical strings and that the task was to perform as well as possible on the grammaticality and the recognition task without previous string memorization. A $2 \times 2 \times 2$ MANOVA, with type of grammar (Grammar A vs. Grammar B) as between-subjects variable and grammatical status (grammatical vs. nongrammatical) as well as order of presentation (first vs. second presentation of a string during test) as within-subject variables on subjects' mean reaction times, revealed no reliable main effects of grammatical status ($F < 1$). Average response times were 7,338 ms for grammatical and 7,379 ms for nongrammatical strings. A reliable effect of order of presenta-

The differences in reaction times tended to be larger for longer than for shorter strings. At least two factors could explain this finding. First, it could simply be that, for short strings in the present experiments, identification is relatively fast, so that a ceiling effect prevents the grammaticality manipulation to be reflected in the reaction times. Second, one could assume that chunks do not only contain information about the elements they are composed of, but also provide positional order information (Johnson, 1970; Tulving, 1962). For instance, each chunk could contain information about the probability of occurrence of a particular subsequent chunk. That subsequent chunk would then benefit from the expectation induced by the previous chunk. If the grouping structure is violated at a particular position, identifying subsequent elements will not be facilitated by such expectations. On average, this effect should be more pronounced the longer the sequence of elements after the violated position. In fact, studies conducted by Dienes, Broadbent, and Berry (1991) and by Mathews et al. (1989) have presented evidence that subjects have available such positional information. However, a direct test of this hypothesis would require string length and position of the incorrect letter to be manipulated orthogonally, which was not the purpose of the present experiments.

The competitive chunking model assumes that, in addition to facilitating perceptual processes, intratext organization also affects how familiar a string appears. The probability of a string to be classified as grammatically correct is said to be a function of its familiarity. The data from Experiment 2 are not consistent with the assumption that familiarity mediates grammaticality judgments. Speed of identification was used as an indicator of perceptual fluency, which should translate into feelings of familiarity. Consistent with existing research, faster identifications were related to subsequent "old" responses for the recognition judgment (see, e.g., Jacoby & Dallas, 1981; Jacoby et al., 1989; Jacoby & Whitehouse, 1988; Jacoby & Witherup, 1982; Johnston et al., 1985, 1991; Joordens & Merikle, 1992; Whittlesea et al., 1990). However, parallel findings were not obtained for the grammaticality judgments. Rather, these appeared to be unrelated to the speed of identification, which is inconsistent with the prediction of the competitive chunking model.

It is important to notice two points. First, in the present article, familiarity was assessed globally, as indicated by the average speed of identification of all items that fell into a certain judgment category (hits, false alarms, etc.). This seems to deviate somewhat from Servan-Schreiber and Anderson's (1990) simulation of grammaticality judgments, in which familiarity was computed locally for each individual string relative to a fixed number of strings that were judged previously (i.e., the number of chunks needed to perceive an individual string was not transformed into some absolute familiarity value, but familiarity was relative in that it took into account the number

$$F(1, 8) = 20.90, MS_e = 302.017.42, R^2 = .72.$$
of strings needed to perceive the preceding test strings; for details, see Servan-Schreiber & Anderson, 1990, p. 604). The assumption of context-dependent feelings of familiarity and, hence, judgments of grammaticality, is very plausible. Unfortunately, the level of analysis could not be so detailed in the present study. However, the average familiarity of all strings judged grammatical in different contexts in the simulation should be higher than the average familiarity of all strings judged nongrammatical. At this level, the simulation and the present study seem comparable. For instance, in Experiment 2, the hypothesis was tested that the average familiarity (as indicated by speed of identification) of the new and incorrect strings that were judged grammatical should be higher than the average familiarity of the new and incorrect strings that were judged nongrammatical. Again, the critical finding is that although the data are consistent with the assumption that familiarity influences recognition judgments, a similar relation does not seem to be present between familiarity and grammaticality judgments.

Second, the present empirical evidence does not mean that familiarity never influences grammaticality judgments—although this remains a possibility. It does, however, argue against the idea that the perceived familiarity of a string is an exclusive basis for grammaticality judgments, as is assumed in the original formulation of the competitive chunking model, which "was able to simulate [grammatical] discrimination behavior . . . without recourse to any of the explicit rules of the form proposed by Dulaney et al." (Servan-Schreiber & Anderson, 1990, p. 605). There was no indication of familiarity influences in the present task, but it is certainly conceivable that familiarity may play a role in different tasks that require judgments about the regularity of events.

Subjects were able to explicitly discriminate, to a certain degree, grammatical from nongrammatical letter strings in the absence of evidence suggesting feelings of familiarity as the basis of grammaticality judgments. This points to the fact that the grammaticality judgments in the present task may have been based on explicit knowledge of permissible groups (chunks, features, or bigrams) of letters, as has been suggested and empirically supported by Dulaney et al. (1984, 1985) and by Perruchet and Pacteau (1990). The assumptions of these authors imply that the chunks relevant to a classification are retrieved from memory and applied to a given letter string. Thus, an immediate expansion of the competitive chunking model would include memory-search processes operating in parallel to the feeling-of-familiarity basis for grammaticality judgments. Grammaticality judgments could then be achieved by simultaneously assessing the familiarity of a letter string and searching for a decomposition into letter chunks that are known. If perceptual fluency associated with an item exceeds a certain threshold (as compared with other items) or if the decomposition is successful given a certain set of retrievable chunks, then the item is classified to be grammatical. If perceptual fluency is below a certain threshold or if the decomposition yields a chunk that is judged not to have been encountered before, then the item is classified to be nongrammatical. Whichever process finishes first would, according to this qualitative conceptualization, determine the direction of the grammaticality judgment.

Such a reformulation of the original model would, of course, be parallel to dual-process models of recognition memory (Gillund & Shiffrin, 1984; Johnston et al., 1985, 1991; Mandler, 1980) and permit implicit grammar-learning research to be integrated into a large body of existing memory research. In addition, the controversy as to whether either familiarity or explicit rules determine judgments of grammaticality could be converted into the empirical question as to what degree feelings of familiarity as opposed to memory-search processes determine grammatical discrimination in a particular task situation.

References


(Appendixes follow on next page)
### Appendix A

Grammatical Strings Generated by Grammar A
and Nongrammatical Strings as Used in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Study Phase: Memorized Strings</th>
<th>Test Phase: Grammatical Strings (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TXS TSSSXS TXXTTTVV TSXXTTVV</td>
<td>PTTV TXXVPS TSXXTTVV TSSXXTVV</td>
</tr>
<tr>
<td>PVV PVFXVPS TXXVPXVV PVFXTPTVS</td>
<td>TXXVV PTTTIVPP TXXTVPS TSSXXVPS</td>
</tr>
<tr>
<td>PTVPSS PTFTVPS TSSSSXS PTVPXTVV</td>
<td></td>
</tr>
<tr>
<td>PTTTAVV TSSXXVV TSSXXVV PTTTTTVV</td>
<td></td>
</tr>
<tr>
<td>PVFXVAVV TSSXXVPS TSXXTVPS PTVPXVPS</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Phase: Grammatical Strings</th>
</tr>
</thead>
<tbody>
<tr>
<td>TXS TSSSXS PTVPXV TXXTTTVV</td>
</tr>
<tr>
<td>PVV PVFXVPS TXXVPXVV</td>
</tr>
<tr>
<td>PTVPSS PTFTVPS TSSSSXS</td>
</tr>
<tr>
<td>PTTTAVV TSSXXVV TSSXXVV</td>
</tr>
<tr>
<td>PVFXVAVV TSSXXVPS TSXXTVPS</td>
</tr>
</tbody>
</table>

Test Phase: Nongrammatical Strings

| TXS TSSSXS PTVPXV TXXTTTVV |
| PVV PVFXVPS TXXVPXVV |
| PTVPSS PTFTVPS TSSSSXS |
| PTTTAVV TSSXXVV TSSXXVV |
| PVFXVAVV TSSXXVPS TSXXTVPS |

### Appendix B

Grammatical Strings Generated by Grammar B and
Nongrammatical Strings as Used in Experiment 2

<table>
<thead>
<tr>
<th>Study Phase: Memorized Strings</th>
<th>Test Phase: Nongrammatical Strings</th>
</tr>
</thead>
<tbody>
<tr>
<td>PXX TSVSX SPXVTPS TTPXVSP</td>
<td>SPT SPXSP TSVSVST PXVSSTPS</td>
</tr>
<tr>
<td>SPXX TSVTXPS TSVTXPS PXVTXPS</td>
<td>TPX TSSTT TSVTXPS TSVPTXPS</td>
</tr>
<tr>
<td>TTPS PXVTXPS PXVSTXSP TSVTXPS</td>
<td>VSP SXXVTT TTTXVTP SPXSVSPT</td>
</tr>
<tr>
<td>PXVTVT SPXVSPX TSXVSPX SPXVTSPX</td>
<td>PTV PTTPVP SPXSVSP TSVSVXP</td>
</tr>
<tr>
<td>TTPXX PXVSPX TPTXVTT PXVSXSP</td>
<td>XXVSX TSXSP SPXVSPX PTPXVTPS</td>
</tr>
</tbody>
</table>

Test Phase: Grammatical Strings

| SPX PXVTSP TSVSTT PXVSTTPS |
| TSX TSVTTT PXVSTST TSVTPSTP |
| TSVP SPXVTT TTPXSTP SPXSVSTT |
| PXVP TTPXVP SPXVSTP TSVSVSTP |
| PXVXS TSVSTP SPXVSTSP TTPXVTPS |

Received February 25, 1993
Revision received May 13, 1993
Accepted June 7, 1993

axel buchner