

Implicit Learning Out of the Lab: The Case of Orthographic Regularities

Sébastien Pacton and Pierre Perruchet
Université de Bourgogne

Michel Fayol
Université Blaise Pascal

Axel Cleeremans
Université Libre de Bruxelles

Children's (Grades 1 to 5) implicit learning of French orthographic regularities was investigated through nonword judgment (Experiments 1 and 2) and completion (Experiments 3a and 3b) tasks. Children were increasingly sensitive to (a) the frequency of double consonants (Experiments 1, 2, and 3a), (b) the fact that vowels can never be doubled (Experiment 2), and (c) the legal position of double consonants (Experiments 2 and 3b). The latter effect transferred to never doubled consonants but with a decrement in performance. Moreover, this decrement persisted without any trend toward fading, even after the massive amounts of experience provided by years of practice. This result runs against the idea that transfer to novel material is indicative of abstract rule-based knowledge and suggests instead the action of mechanisms sensitive to the statistical properties of the material. A connectionist model is proposed as an instantiation of such mechanisms.

Most definitions of implicit learning focus on the idea that people's behavior becomes sensitive to the structural features of a situation in such a way that the adaptation does not result from people's intentional use of their explicit knowledge of these features (Perruchet & Gallego, 1997). Although the literature on implicit learning (for reviews, see Berry & Dienes, 1993; Cleeremans, Destrebecqz, & Boyer, 1998) is grounded in laboratory studies, many introductory texts claim that implicit learning is a fundamental process involved in domains as diverse as first and second language learning, category elaboration, reading and writing acquisition, and acquisition of knowledge about the physical world or social skills (e.g., Reber, 1993). There is, however, a sharp contrast between these widespread conceptions and the

content of actual research. On the one hand, implicit learning researchers tend to be evasive when attempting to generalize their conclusions to real-life behavioral abilities. On the other hand, experts on learning-related topics that have direct relevance out of the laboratory (such as cognitive development in general) tend to ignore potentially important results from the implicit learning literature.

This lack of reciprocity is made clear by the conspicuous paucity of studies specifically designed to explore the links between implicit learning and real-life learning. One example is Perruchet and Vinter's (1998a) exploration of the links between implicit learning and developmental theories. Another attempt to ground implicit learning research in a more naturalistic context is provided by studies on second language acquisition (Michas & Berry, 1994; Schmidt, 1994). However, despite the fact that language acquisition constitutes a prototypical implicit learning situation, relevant literature about implicit learning was not even mentioned in a recent annual review on language acquisition (McDonald, 1997).

Our first goal in writing this article is to contribute to bridging the gap between laboratory research and natural situations by investigating whether the conclusions derived from the former generalize to the latter. Our second and main goal is to exploit the specific properties of natural situations—such as the very extended time scale over which real-world learning takes place—to shed light on one of the most controversial issues in the implicit learning literature, namely the question of whether knowledge acquired in implicit learning situations can be described as abstract. In what follows, we first outline the main arguments involved in this debate. Because most studies designed to explore abstraction have used the artificial grammar paradigm, we focus on this paradigm. We then introduce the natural learning situation used in our research and detail how the characteristics of this situation make it an attractive vehicle through which to explore the development of abstract knowledge in real-life learning situations.

Sébastien Pacton and Pierre Perruchet, Laboratoire d'Etude de l'Apprentissage et du Développement (LEAD/CNRS), Université de Bourgogne, Dijon, France; Michel Fayol, Laboratoire de Psychologie Sociale de la Cognition et de Psychologie Cognitive (LAPSCO/CNRS), Université Blaise Pascal, Clermont Ferrand, France; Axel Cleeremans, Séminaire de Recherche en Sciences Cognitives, Université Libre de Bruxelles, Brussels, Belgium.

This research was supported by the Université de Bourgogne, the Centre National de la Recherche Scientifique (Unité Mixte de Recherche [UMR] 5022), the Région de Bourgogne (Acquisition Apprentissage Formation Education program), and the Université Libre de Bruxelles (in support of Inter-University Poles of Attraction Program P/4-19). Axel Cleeremans is a Research Associate of the National Fund for Scientific Research (Belgium). We thank Catherine Transler and Pierre Barrouillet for discussing the ideas presented in this article.

Correspondence concerning this article should be addressed to Sébastien Pacton, who is now at the Laboratoire Cognition et Développement, UMR 8605, CNRS-Paris V, 71, Avenue Edouard Vaillant, 92774, Boulogne-Billancourt Cedex, France. Electronic mail may be sent to Sebastien.Pacton@psycho.univ-paris5.fr.

Artificial Grammars and the Abstraction Issue

In the standard artificial grammar learning paradigm (Reber, 1967, 1993), participants first study a set of letter strings generated from a finite-state grammar that defines legal letters and permissible transitions between them. Typical instructions do not mention the existence of a grammar and are framed so as to discourage participants from engaging in explicit, intentional analysis of the material. Participants are then subsequently informed about the rule-governed nature of the strings and asked to categorize new grammatical and nongrammatical letter strings. Participants are typically able to perform this task with better-than-chance accuracy, while remaining unable to articulate the rules used to generate the material.

This empirical outcome, first demonstrated in the pioneering work of Arthur Reber (1967), has been unambiguously confirmed by a vast number of subsequent experimental studies involving many variants of the situation (for recent edited books, see Berry, 1997; Stadler & Frensch, 1998). Its interpretation, however, remains controversial. Reber's (1967) original proposal was to assume that participants have internalized the constraints embodied by the generation rules over training. Rule abstraction is assumed to occur during the study phase, when participants are exposed to a sample of letter strings generated from the grammar. During the test phase, participants are assumed to use the acquired knowledge, stored in an abstract format, to judge the grammaticality of new items. Clear illustrations of this reasoning can be found in subsequent studies concerning sequence learning (e.g., Lewicki, Czyżewska, & Hoffman, 1987; Lewicki, Hill, & Bizot, 1988), number recognition (McGeorge & Burton, 1990) and prediction (Kushner, Cleeremans, & Reber, 1991).

The original results turned out to be empirically robust, but subsequent studies demonstrated that performance improvement could be explained without assuming rule abstraction. Most of this counterliterature was directed against the evidence stemming from artificial grammar learning settings. The first attack was made by Brooks (1978), who showed that correct grammaticality judgments for new items could be dependent on the perceived surface similarity between these new items and the stored traces of study items. According to Brooks, no abstraction occurs during the familiarization phase: The strings are simply memorized as whole exemplars. Then, during the test phase, the grammaticality of each novel string is evaluated as a function of its overall similarity with the most similar memorized item. It is important to note that similarity is computed in a nonanalytical way, that is, simply by counting the number of common features without assigning separate weights to each. Brooks's account was successful because typical grammatical strings are, on average, more similar to the items of the study phase than nongrammatical items are. As a consequence, in most cases, grammaticality and nonanalytical similarity overlap to a large extent and hence predict similar classification performance. In a subsequent study, however, Vokey and Brooks (1992) were able to independently manipulate similarity to specific training items and grammaticality within their stimulus set, and found that the similarity hypothesis (Brooks, 1978) failed to account for all aspects of participants' classification performance. Vokey and Brooks thus concluded that grammaticality and similarity both contributed about equally to overall classification performance—a suggestion that has enjoyed widespread acceptance by many other

authors (e.g., Knowlton & Squire, 1996; Meulemans & Van Der Linden, 1997; Reber, 1993).

According to other authors, the claim that classification performance is based both on unconscious rule abstraction and similarity to exemplars lacks parsimony. These authors have therefore explored the possibility that participants' knowledge consists neither of abstract rules nor of whole memorized exemplars, but of small memorized chunks of information such as bigrams or trigrams (e.g., Dulany, Carlson, & Dewey, 1984; Perruchet, 1994a; Perruchet & Pacteau, 1990, 1991; Servan-Schreiber & Anderson, 1990). The assumption of so-called fragmentary accounts of implicit learning is simply that task demands naturally induce participants to segment the material in such small units, in the same way that we spontaneously chunk telephone numbers. Classification performance in the context of artificial grammar learning tasks would then be based on the extent to which the novel items can be processed using the memorized chunks. A novel string will tend to be judged as grammatical to the extent that its chunks appear familiar to participants. Several studies have now suggested that such fragment-based accounts are capable of accounting for performance in a variety of implicit learning paradigms. For instance, Perruchet (1994a) showed how such a model could account for all aspects of performance in the Vokey and Brooks (1992) situation. Likewise, the sequence-learning results obtained by Lewicki et al. (1988) were reinterpreted by Perruchet, Gallego, and Savy (1990), those of McGeorge and Burton (1990) by Wright and Burton (1995), and those of Kushner et al. (1991) by Cleeremans (1994) and Perruchet (1994b).

Fragment-based approaches have much in common with connectionist models in that they do not rely on the existence of separate learning systems. Indeed, both of them assume that learning involves the development of a sensitivity to the distributional or statistical features of the training material. Many connectionist models have been developed to account for performance in implicit learning situations. Most are based on the simple recurrent network (SRN) architecture first developed by Elman (1990; see also Cleeremans, 1993). While it is still unclear at this point the extent to which fragment-based and connectionist accounts overlap in their predictions, we nevertheless refer to both approaches as "statistical" or "distributional" accounts in the following, so as to highlight the central property they share and that distinguishes them from other alternatives.

Although statistical approaches have proven successful in accounting for most empirical findings in the implicit learning literature, results regarding the possibility of transfer to novel material appeared to be problematic for those approaches. For instance, in the so-called letter-transfer paradigm, people are asked for grammaticality judgments about strings generated by a grammar identical to the one used in the familiarization phase but instantiated with a novel set of letters. In this situation, the training and test stimuli thus share the same underlying abstract structure but involve different surface forms. Typical results in this paradigm indicate that participants remain able to perform above chance on the discrimination task, although their level of performance is very low and in all cases lower than in the standard condition, where surface features do not differ between training and test. This pattern of results is known as the *transfer decrement* (see Dienes & Altmann, 1997; Gomez, 1997; Mathews et al., 1989; Shanks, Johnstone, & Staggs, 1997; Whittlesea & Wright, 1997). The same

pattern of results has also been observed for transfer in other tasks sharing the same fundamental principles (e.g., Altmann, Dienes, & Goode, 1995; Reber, 1969). Initially, the observation that transfer to novel surface forms actually occurs was taken as an indication that the underlying knowledge is rule based. For instance, the fact that participants trained on strings such as VVPTN and TTVPN are able to transfer to a new string such as ZZLJH was interpreted as evidence that an abstract rule such as "Legal strings begin with a doubled letter" had been learned (e.g., Knowlton & Squire, 1996; Reber, 1993). This account of transfer has now been challenged, on the basis of the fact that independence from surface features (as observed in the experiments described above) does not necessarily entail that the underlying knowledge is rule based. To quote Redington and Chater (in press), "The key point here is that surface-independence and rule-based knowledge are orthogonal concepts" (p. 4). Indeed, although traditional conceptions of the notion of rule typically involve surface-independent features, and although statistical accounts have likewise typically been expressed in a way that appears to make their mechanisms rely on surface features, other combinations of both dimensions are nevertheless possible: Rules can involve surface features, and, crucially for our concern, memory and statistical processes can apply to features that are abstracted away from their sensory content. Thus, although transfer indeed unquestionably reveals some form of abstractive process, one can question the claim that this form of abstraction, conceived hereafter as any knowledge of the structure of the training stimuli that is independent of their specific instantiation with a particular vocabulary of features (see Redington & Chater, 1996, for other meanings of abstraction),¹ is indicative of rule formation and rule use.

One possibility is that the study strings stored in memory under their surface forms are subsequently recoded during the test phase so that their abstract or relational features map with those of specific test items. For example, when presented with the test item ZZLJH, participants could remember that VVPTN, for instance, was displayed during the study phase, and they could see both items as similar in that both start with a double letter. Note that there is no rule abstraction here; at issue is only the coding of a specific item, whereas rule abstraction would imply the abstraction of regularities common to a large number of training items (Brooks & Vokey, 1991; Redington & Chater, 1996). Alternatively, it may be assumed that abstract coding operates on-line, during the study phase, and that statistical processes apply to the abstract features in exactly the same way as they may apply to surface features. The end result is that participants will become sensitive to the repetition of the first letter at test and will identify as grammatical any item that shares this feature, even when it is instantiated with other letters. Perruchet and Vinter (in press) have shown that the available data evidencing transfer can be accounted for by the action of statistical principles acting on very simple abstract or relational primitives, such as they can be provided by low-level perceptual processes. Importantly, these primitives can be extracted by simple connectionist networks.

The possibility of accounting for transfer performance within a connectionist framework has been used as a powerful argument to show that statistical approaches are consistent with the available demonstrations of transfer (Dienes, Altmann, & Gao, 1999; McClelland & Plaut, 1999; Seidenberg & Elman, 1999), which runs against the claim of Marcus (1998) that connectionist models such

as SRNs are not capable of "outside" generalization.² Because we rely on a connectionist network to account for our own transfer data later in this article, the details of how such networks abstract are postponed for the *Discussion* section. It is nevertheless important to mention from now on that considering abstract or relational similarity rather than surface similarity does not change anything in the way transfer occurs. In particular, statistical approaches account naturally for the pervasive phenomenon of transfer decrement. Indeed, in any cases, transfer relies on the amount of similarity between study and test material, and similarity, whether computed from surface or deep features, varies along a graded dimension.

By design, rule-based theories have no difficulty in accounting for abstraction. However, the fact that transfer performance on novel items remains lower than performance on familiar material is problematic for such models because genuine rule-based abstract knowledge is supposed to be completely independent from surface features. For instance, Anderson (1993) noted that "abstraction refers to the generality of production rules. Production rules do not require that a specific stimulus be present; the rules will apply in any stimulus condition that satisfies the pattern specification of the condition" (p. 35; see also Smith, Langston, & Nisbett, 1992). In the context of artificial grammar learning studies, Whittlesea and Dorken (1997) likewise posited that "a subject who learned a useful rule would have equal success in transfer on stimuli presented either in the same or different features, because the rule is applicable regardless of the features in which items are presented" (p. 66). Manza and Reber (1997) acknowledged this implication of their own abstractionist view: "A system that rep-

¹ Redington and Chater (1996) proposed to consider three different notions or degrees of abstraction. In a first, trivial sense, abstract knowledge is simply knowledge that goes beyond the raw stimulus, in the sense that any stimulus will always be interpreted to some extent. For instance, patterns of light and dark disposed in specific arrangements on a page are interpreted as strings of letters when perceived. In a second sense, abstraction refers to abstraction over the surface properties of individual training items. For instance, in artificial grammar learning, bigrams and trigrams can be used to recode the training strings into more compact representations. These representations are abstract in the sense that they capture some aspects of the structure of the ensemble of training strings, for instance their redundancy. Finally, a third meaning of abstraction in this context refers to knowledge that abstracts away from the specific vocabulary used in the training set. Thus, both strings of colored patches and strings of letters could be recoded in similar abstract representations that capture the deep internal structure of the stimuli in a way that does not depend on their specific instantiation with a particular vocabulary of features. It is the nature of the mechanisms that subtend this third notion of abstraction that constitutes the focus of interest in the implicit learning literature.

² Marcus (1998) suggested that there is an important distinction between two types of generalization, which he called "within-training-space" generalization and "outside-training-space" generalization (p. 258). According to Marcus, simple recurrent networks are capable of "within" generalization but not "outside" generalization. Marcus argued that it is the outside-training-space generalization that provides the strongest evidence for rules. However, it must be stressed that the application of this distinction to any specific situation is problematic, due to the difficulty involved in defining participants' training space. Indeed, the training space is entirely dependent on the nature of the coding of the study material and particularly the extent to which the coding involves abstract dimensions.

resented knowledge in this pure abstract form would result in transfer judgment accuracy being equal to control judgment accuracy which, of course, was not what we typically found" (p. 102).

Thus, an essential prediction of any system using abstract rules to represent its knowledge about some domain is that its transfer performance on novel items should be just as good as its performance on familiar items. One way of reconciling the ubiquitous phenomenon of transfer decrement with the assumption that knowledge is rule-based is to assume that abstract knowledge requires extensive practice to fully develop and that practice is generally insufficient in existing studies. This explanation is consistent with the position advocated by Manza and Reber (1997), who described artificial grammar learning as a gradual phenomenon that starts from a state where "each individual letter would have to have its sensory components understood" (p. 101) and that continues up to a point where "an 'abstractor' would come into play, gradually removing irrelevant surface elements and leaving only structural elements in the representation" (p. 101). According to this perspective, cases where transfer decrement is observed would then correspond to a point in training where this latter process is not yet fully completed.

This analysis shows that rule-based and distributional approaches offer contrasted predictions about the relationship between performance on familiar and novel material after different amounts of practice. Specifically, rule-based approaches would predict a progressive convergence of performance on familiar and novel material. This prediction follows directly from the perspective that learning involves the development of knowledge that is first sensitive to the low-level surface features of the stimulus material and that subsequently becomes increasingly independent from these features up to the point where it reflects only the abstract, deep features of the material (e.g., Manza & Reber, 1997). In contrast, distributional approaches would predict continued lower levels of performance on novel material, even after extensive training, because the similarity between familiar and novel situations, which determines transfer in any statistical approach, remains unchanged across time. Consistently, statistical approaches would thus further predict continued parallelism between performance on familiar and novel material over practice.

Distinguishing between these diverging predictions might prove to be impossible using standard laboratory conditions. Indeed, training in such circumstances is necessarily restricted, both in duration and in the number of stimuli experienced by participants. As mentioned above, from an abstractionist standpoint (e.g., Manza & Reber, 1997), the development of fully abstract, rule-based knowledge requires considerable time as well as exposure to a very large number of training instances. As a result, one reason why transfer decrement is regularly observed in laboratory conditions might simply be that training has never been sufficiently extended or intensive enough to actually allow for the emergence of pure abstract knowledge. This observation constitutes the cornerstone of our study's rationale. To address the debate concerning abstraction, we proposed to explore whether and to what extent transfer to novel forms occurs in the context of a natural learning situation that extends over an extremely long duration (i.e., years). To address these issues in a natural learning setting, we explored children's development of knowledge about orthographic regularities in printed language—a rich domain that we briefly describe in the following section.

Implicit Learning Out of the Lab: Processing Orthographic Regularities

First language acquisition undoubtedly involves implicit learning. However, it is somewhat problematic to use this situation as a paradigm through which to study the development of abstraction, for several reasons. First, spoken language is learned during a period in development that is particularly challenging to empirical investigation. Second, we cannot exclude the possibility that language learning relies, at least to some extent, on specific hard-wired mechanisms (e.g., Chomsky, 1965; Pinker, 1994). Third, language acquisition is at least partly grounded in instructed learning, through both parental feedback and schooling.

For all these reasons, we have chosen to focus on a much more restricted skill, namely, the development of sensitivity to certain orthographic regularities based on experience of printed language not explicitly taught at school. Traditional models of spelling acquisition (e.g., Ehri, 1986; Frith, 1985; Henderson, 1985) depict spelling development as a sequence of stages involving the ability to use different types of information. Initially, young spellers are supposed to use only two sources of information: phoneme-grapheme correspondences and letter names. As they progress through the stages, they gradually gain the ability to use orthographic and morphological information. However, more recent studies have shown that even young children appear to be sensitive to lexical orthographic regularities and that this knowledge influences spelling much earlier than stage models of spelling development would predict (e.g., Cassar & Treiman, 1997; Nation, 1997; Pacton, Fayol, & Perruchet, in press; Treiman, 1993). For instance, Cassar and Treiman (1997) conducted a series of experiments aimed to show that even young writers are sensitive to orthographic regularities such as which letters can be doubled and the position at which such doublets can occur. These sorts of orthographic regularities are particularly well suited for the purpose of our study because (a) they are not explicitly taught during instruction, and (b) children become sensitive to them during a period in development over which experimental investigations are unproblematic. Thus, they constitute a genuine example of an acquired sensitivity to a complex feature of printed language that is not driven by the intentional exploitation of explicit knowledge about this feature. Indeed, to quote Cassar and Treiman (1997), "It is unlikely that the children's parents or teachers explicitly taught that words do not begin with double consonants. More likely, the children learned this convention through their experience with print" (p. 637). However, as another illustration of the gap that exists between domain-specific and laboratory research, they never referred to general (e.g., implicit) learning theory in their article.

To assess children's sensitivity to regularities in printed language, Cassar and Treiman (1997) used an *orthographic constraints* test (see Treiman, 1993) in which children are shown a set of pairs of nonwords such that one nonword in each pair conforms to a constraint of the English writing system whereas the other nonword does not. For instance, to test children's knowledge about which letters can or cannot be doubled, the material included pairs such as *noss-novv* and *yill-yihh*. Children were asked to choose which item within each pair looked more like a real word. First and second graders (although not kindergartners) tended to prefer nonwords that contained permissible doublets over nonwords that contained doublets that never occur (e.g., children judged that *noss*

and *yill* look more like real words than *novv* and *yihh*). On the basis of these results, Cassar and Treiman concluded that these children were sensitive to frequency of doubling. However, an alternative interpretation of these results is that preference for a nonword such as *yill* over a nonword such as *yihh* may reflect not children's sensitivity to the consonants' frequency of doubling but rather their sensitivity to the frequency of single letters. Indeed, as *l* is more frequent than *h*, it is unclear whether children state that *yill* would be a better word than *yihh* because *l* is doubled more often than *h* or simply because *l* is a more frequent letter than *h*.

The phenomenon of developing sensitivity to the identity of the letters that may or may not be doubled is ill suited to explore abstraction because there is no single high-level rule that could provide a concise definition of which letters may or may not be doubled in English. This property instead constitutes an idiosyncratic property of each letter. Other orthographic regularities such as the position at which doublets may occur within words can be described with a rule. Again, using the orthographic constraints test described above, Cassar and Treiman (1997) presented pairs of nonwords such as *nuss-nnus* to test children's sensitivity to the legal position of doublets. They found that even kindergartners judged *nuss* as more wordlike than *nnus* and concluded that children knew the position at which consonants can be doubled in English. However, children's preference for nonwords such as *nuss* over other nonwords such as *nnus* may be due to the learning of letter-specific features. For instance, participants may consider that *nuss* looks more like a word than *nnus* because they have learned that *ss* and *nn* occur frequently in the end of words but not at the beginning.

A better way to explore the abstract character of children's orthographic knowledge about the legal position of doublets consists of testing children's knowledge about the legal position of doublets with consonants that are *never* doubled. Some nonword pairs in Cassar and Treiman's (1997) study were aimed to test whether children's knowledge of the legal position of doublets is general. According to Cassar and Treiman, it was indeed general because children chose *juss* as more wordlike than *jjus* as often as they chose *nuss* as more wordlike than *nnus*, even though *j* is never doubled in English. They wrote, "whether or not a consonant normally occurs as a doublet in English had little influence on the results. . . . [T]he children had generalized the knowledge that words do not begin with double consonants to include consonants they had never seen as doublets, such as *jj*" (Cassar & Treiman, 1997, p. 637). However, children's preference for *juss* over *jjus* fails to demonstrate that children's knowledge about the legal position of doublets is general. Indeed, as *j* is never doubled but *s* is frequently doubled in English, children's preference for *juss* over *jjus* may simply reflect their knowledge of the identity of allowable doublets. One way to get around this difficulty would consist, for instance, of asking children which one of two nonwords such as *jukk* and *jjuk* looks most like a word. In this latter case, the training and test stimuli have different surface forms but the same underlying abstract structure, as in the letter-transfer paradigm commonly involved in artificial grammar learning. Thus, this characteristic of the use of double letters (their position) allows us to transpose both a problematic (the abstraction issue) and a paradigm (letter transfer) commonly used in laboratory contexts to a naturalistic context.

Experimental Design

We used nonwords rather than words because using words would not allow us to determine whether children are sensitive to particular orthographic regularities or whether they simply know the spellings of specific words. The orthographic regularities explored in this study have been obtained using a computerized database for written and spoken French known as Brulex (Content, Mousty, & Radeau, 1990). Although this database may be unrepresentative of the words children are exposed to, we have assumed that the infra-lexical orthographic regularities present in this database do not differ from those present in children's books.

Some comments are necessary concerning the French orthographic regularities that we have exploited. In French, vowels are never doubled. Certain consonants can be doubled whereas others cannot. Double consonants occur only in medial position (i.e., never in initial or final positions). The format (single vs. double) of a consonant does not generally lead to a modification in the phonological form of French words. For instance, *om* and *omm* are similarly pronounced /om/ in words such as *pomme* (apple) and *moment* (moment), and *ul* and *ull* are similarly pronounced /yl/ in words such as *bulle* (bubble) and *formule* (formula). In fact, the consonant *s* constitutes the only case where the format of a consonant changes its pronunciation in French: *s* is pronounced /z/ (e.g., *rose*, meaning pink) whereas *ss* is pronounced /s/ (e.g., *grosse*, meaning big) between two vowels. Experiments reported in this article did not include nonword pairs in which one nonword contains *s* and the other *ss* so that phonological and orthographic factors are not confounded. This ensures that sensitivity to the orthographic regularities relevant for our experiments resulted only from exposure to printed texts. Concerning the frequency of double consonants, some consonants are frequently doubled (e.g., *m* or *l*), whereas others are infrequently or never doubled (e.g., *k* or *x*). Taken as a whole, there is a link between the frequency of a given consonant presented in isolation and in a doublet. For instance, *m* is both more frequent and more often doubled than *k*. More generally, the Pearson product-moment correlation between the overall number of occurrences of a consonant in a single format and this consonant in a double format is .59. However, it is possible to select consonants occurring frequently in both single and double formats (e.g., *m* or *l*) and consonants occurring frequently only in single format (e.g., *c* and *d*). This latter characteristic is interesting because it enables us to test whether children are really sensitive to the frequency of double consonants or only to the frequency of single consonants.

Experiment 1 was devoted to examining whether and when children are sensitive to the consonants' frequency of doubling, through a judgment task of nonwords by contrasting consonants that differ only in their frequency of doubling in French (e.g., *l* or *m* vs. *c* or *d*).

Experiment 2 was aimed at investigating, through a nonword judgment task, children's sensitivity to the fact that vowels are never doubled and to the legal position of double consonants (two regularities that can be described with a rule), in addition to their sensitivity to the identity of the consonants that can versus cannot be doubled (a regularity that cannot be described with a rule). The main goal of Experiment 2 was to exploit the fact that double consonants are permissible only in the medial position of words to investigate whether children acquire genuine knowledge of this

rule. Experiment 2 applied the strategy involved in the letter-transfer paradigm in artificial grammar context to orthographic regularities in order to determine whether sensitivity to the legal position of double consonants extends (and in what proportions) to consonants that are never doubled. As previously mentioned, evidence for an effect of this variable with frequently doubled consonants is not conclusive because it may be due to the learning of letter-specific features. For instance, if participants consider that *lammi* looks more like a French word than *llami*, this may be due to the fact that they have learned that *mm* and *ll* occur frequently in the middle of words but not at the beginning or at the end. However, if participants select nonwords such as *xevvu* more often than nonwords such as *xxevu*, because *xx* and *vv* never occur in French this preference cannot be attributed to the learning of letter-specific features. As suggested by laboratory studies involving the letter-transfer paradigm, we expected that the sensitivity to legal position learned from a set of doubled letters would transfer to another, unseen set of doubled letters with a diminution of amplitude. However, the crucial question relates to the persistence of transfer decrement when exposure to the study material extends over many years. Children may also become sensitive to the fact that vowels cannot be doubled in French (a characteristic that can be described with a rule) because they learned this property for each vowel, in the very same way that they learned that some consonants are never doubled. To test whether abstraction occurs, in Experiment 2 we explored whether children select nonwords containing double vowels (e.g., *tuuke*) more often than nonwords containing a doublet formed with consonants that are never doubled in French (e.g., *tukke*) when they are paired together. As neither *uu* nor *kk* has been seen in the past, a higher rate of rejection for *uu* than for *kk* would suggest that some form of abstraction has occurred.

Experiments 3a and 3b were aimed at exploring whether the preference judgment effects observed in Experiments 1 and 2 generalize to a production task. Treiman (1993) provides some indications suggesting that such a generalization is valid. She examined the writings of first-grade children who were encouraged to write without any corrections from teachers. She reported that letters that frequently occur as doublets were more likely to occur as doublets in children's writings than letters that are infrequently or never doubled. She also observed that children tended not to make errors such as *bbal* for *ball*. In Experiments 3a and 3b, we explored children's sensitivity to regularities of the use of double consonants in a more constrained completion task.

Experiment 1

Experiment 1 explored whether and when children are sensitive to the frequency of double consonants when the potential confound of the component (single) consonants is removed. In French, despite the fact that there is a correlation between the number of occurrences of a single consonant and this consonant in a double format (.59), it is possible to tease apart consonants frequent in both single and double format and consonants frequent in single format but infrequently or never doubled. Let us consider the six consonants *c*, *d*, *l*, *m*, *s*, and *v*. The mean percentage of occurrences of *l*, *m*, and *s* on the one hand and of *c*, *d*, and *v* on the other hand differ only slightly (55.3% and 44.7%, respectively). However, the mean percentage of occurrences of *ll*, *mm*, and *ss* is much higher

than that of *cc*, *dd*, and *vv* (95.4% and 4.6%, respectively). This difference, according to whether consonants are doubled or not, allows us to test whether children are really sensitive to the frequency of double consonants. If children are sensitive to the frequency of single consonants alone, the selection rate of nonwords containing *l*, *m*, and *s* on the one hand versus *c*, *d*, and *v* on the other hand should not differ significantly according to whether these consonants occur in a single or a double format. By contrast, if children are sensitive to the frequency of double consonants, nonwords including *l*, *m*, and *s* (rather than *c*, *d*, and *v*) should be selected more often when consonants occur in a double format than in a single format.

Method

Materials. There were 18 pairs of six-letter nonwords labeled "double." Each nonword of these pairs contained a doublet in the medial position. One nonword contained *ll*, *mm*, or *ss* (three consonants the frequency of which is high in both single and double formats), whereas the other contained *cc*, *dd*, or *vv* (three consonants the frequency of which is high in single format but low in double format). The other letters making up the members of a pair were the same. Examples of these pairs are *ommera-ovvera* and *obacci-oballi*.

There were also 18 pairs of five-letter nonwords called "single." They were like the double ones with the exception that the six consonants *l*, *m*, *s*, *c*, *d*, and *v* occurred in a single format. Thus, one nonword included *l*, *m*, or *s*, whereas the other included *c*, *d*, or *v*. Examples of these pairs are *imose-idose* and *oryse-oryve*.

Twelve pairs of nonwords labeled "filler" were constructed. These pairs, which were designed only to make the characteristic that constitutes the focus of interest of Experiment 1 (i.e., the frequency of double consonants) less salient, were not analyzed. In six of the filler pairs, a nonword contained one of the six doublets *cc*, *dd*, *ll*, *mm*, *ss*, and *vv*, whereas the other contained a nonexistent doublet in French. Apart from the doublets, which were always in the medial position, the nonwords used the same letters in their spelling. Examples of these pairs are *illoma-ikkoma* and *oxubi-ollubi*. In the six other filler pairs, one nonword contained one of the six single consonants *c*, *d*, *l*, *m*, *s*, and *v*, whereas the other included a single consonant never doubled in French. Examples of these pairs are *oseru-okeru* and *axile-adile*.

Three nonword pairs were created as practice items. Only one item of these pairs could be a French word (*vilo-xtyu*; *rsmt-baso*; *natu-gnhd*). Complete lists of the stimuli are presented in Appendix A.

For each type of pair (i.e., single, double, and filler) the six consonants *c*, *d*, *l*, *m*, *s*, and *v* (in both single and double formats) were on the right in half of the trials and on the left in the other half. The pairs were written on sheets of paper (21 cm × 3 cm) using lowercase letters in 14-point, New York type. They were presented in a random order and stapled together just behind the three practice nonword pairs in order to make a booklet. Each page of the booklet was numbered (at the top right). We presented only one nonword pair on each page to prevent the participants from going back over and correcting their earlier responses.

Procedure. Participants were told that the experimenter had made up new words that no one had ever seen or heard before and that their help was needed in deciding how a made-up word is like a known word. The children were told that their task was to look at each pair of nonwords and to circle the nonword they thought looked most like a word. They were then asked to look at the first practice pair and to circle that item on their booklet. This procedure was repeated with the remaining two practice items and the actual experiment began. The nonword pair numbers were called out to avoid missing items and to promote the careful consideration of each pair.

Participants. Participants included 20 first graders (mean age: 6 years, 8 months), 20 second graders (mean age: 7 years, 11 months), 20

third graders (mean age: 8 years, 9 months), and 20 fourth graders (mean age: 9 years, 11 months). The experiment was conducted at the end of March (i.e., 7 months after the beginning of the school year).

Results

Figure 1 shows the selection percentage of nonwords including *l*, *m*, and *s* (henceforth “*lms* nonwords”), rather than *c*, *d*, and *v* (henceforth “*cdv* nonwords”) as a function of the type of nonword pairs and grade level. Analyses using *t* tests showed that at each grade level, children selected *lms* nonwords (high frequency in both single and double formats) more often than *cdv* nonwords (high frequency only in single format) for both single and double nonword pairs, all $t(19) > 3.0, p < .007$.

Analyses of variance (ANOVAs) with the variables grade level (1 to 4) and type of nonword pairs (single vs. double) were conducted. The dependent variable was the percentage of selection of *lms* nonwords (rather than *cdv* nonwords). The data were analyzed both by subject (F_1) and by item (F_2). There was a main effect of the type of nonword pairs: $F_1(1, 76) = 83.8, MSE = 5.6, p < .001; F_2(1, 34) = 59.9, MSE = 8.8, p < .001$. The *lms* nonwords were selected more often for double nonword pairs (84%) than for single nonword pairs (66%). This effect was significant as early as in first grade, $t(19) = 2.5, p = .02$. The effect of grade level was significant: $F_1(3, 76) = 6.8, MSE = 9.5, p < .01; F_2(3, 102) = 16.7, MSE = 4.3, p < .01$. The *lms* nonwords were selected less in Grade 1 (65%) than in Grades 2, 3, and 4 (75%, 78%, and 82%, respectively): $F_1(1, 76) = 17.5, MSE = 9.5, p < .001; F_2(1, 34) = 32.6, MSE = 5.7, p < .001$. The selection rate of *lms* nonwords did not differ significantly from Grade 2 to Grade 4 ($F_1 < 1; F_2 < 1$). The Grade Level \times Type of Nonword Pairs interaction was significant: $F_1(3, 76) = 3.4, MSE = 5.6, p = .02; F_2(3, 102) = 4.9, MSE = 4.3, p = .003$. This interaction resulted from a higher amplitude of the effect of the type of nonword pairs in Grades 3 and 4 (26% and 25%, respec-

tively) than in Grades 1 and 2 (13% and 12%, respectively): $F_1(1, 76) = 8.4, MSE = 5.6, p = .004; F_2(1, 34) = 11.3, MSE = 4.6, p = .002$.

Discussion

Experiment 1 was conducted in order to examine children acquiring sensitivity to the frequency of double consonants presented in legal position. Because there is a link between the frequency of a given consonant presented in isolation and in a doublet, we contrasted nonwords including consonants with a high frequency of occurrence in both single and double formats with nonwords including consonants with a high frequency in single format only. For both single and double nonword pairs, children preferred *lms* nonwords rather than *cdv* nonwords. This difference might reflect children’s sensitivity to the consonants’ overall (single plus double) frequency (the overall number of occurrences of *l*, *m*, and *s* being 26% higher than that of *c*, *d*, and *v*). However, children’s sensitivity to the consonants’ overall frequency, or to the frequency of single consonants failed to account for the effect of the type of nonword pairs, that is, to account for why children selected *lms* nonwords (rather than *cdv* nonwords) more often for double nonwords than for single nonwords. The fact that the effect of the type of nonword pairs was observed as early as in Grade 1 shows that even first graders were sensitive to the frequency of double consonants. Moreover, this sensitivity appeared to increase from Grade 2 to Grade 3.

Experiment 2

In addition to children’s sensitivity to the identity of the consonants that can (or cannot) be doubled (as in Experiment 1), Experiment 2 exploited two other characteristics of the use of double letters in French: (a) the fact that vowels cannot be doubled and (b) the legal position of double consonants.

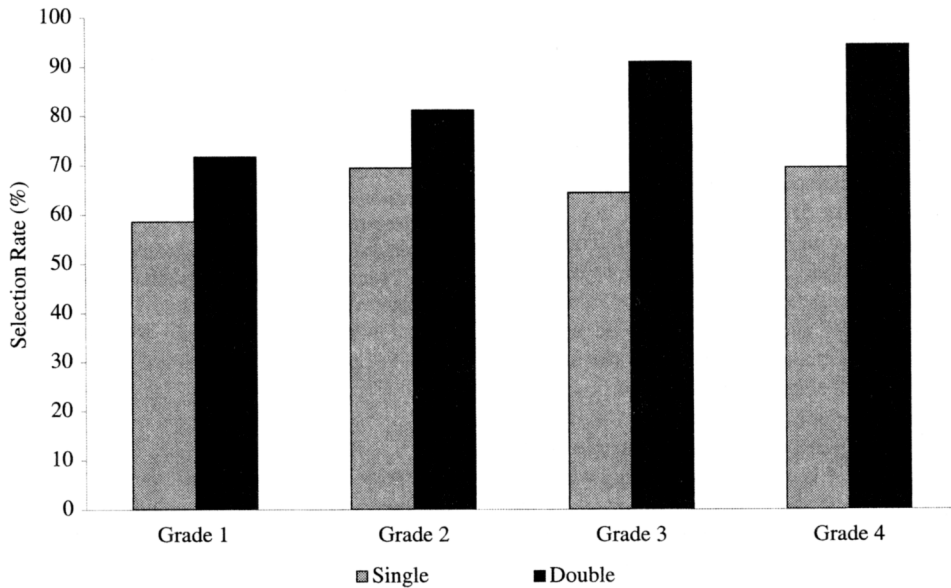


Figure 1. Selection percentage of nonwords including *l*, *m*, and *s* (*lms* nonwords), rather than *c*, *d*, and *v* (*cdv* nonwords) as a function of the type of nonword pairs (single vs. double) and grade level (Experiment 1).

The fact that vowels cannot be doubled in French can be construed as a rule. However, children may become sensitive to this aspect because they learned this property for each vowel, in the very same way that they learned that some consonants are never doubled. The occurrence of abstraction can nevertheless be evaluated by assessing whether children select in different proportions nonwords containing a doublet formed with consonants that are never doubled in French (e.g., *tukke*) and nonwords containing double vowels (e.g., *tuuke*) when they are paired together. As neither *uu* nor *kk* have been seen in the past, a higher rate of rejection for *uu* than for *kk* would suggest that some form of abstraction has occurred. Indeed, the fact that *u* is never doubled not only is an idiosyncratic property of this letter, as it is for *k*, but may be derived from the rule that vowels cannot be doubled.

The fact that double consonants are not permissible at the beginning and at the end of words in French can also be construed as a rule. To explore whether children are sensitive to this rule, we presented them with pairs of nonwords, with one nonword including a double consonant located in medial (legal) position and the other nonword a double consonant located in the initial or final (illegal) position. As expected, the obtained results were positive. However, these results do not indicate genuine abstraction. They might be a by-product of learning the idiosyncratic properties of each consonant. To test this hypothesis, we used a transfer situation, in some ways analogous to the conventional letter-transfer paradigm in the artificial grammar learning literature (see the introduction). Children's choices were compared depending on whether the doublet was formed with consonants frequently or never doubled in French. For example, participants were asked which one of the two nonwords *lammi* or *llami*, on the one hand, and *xevvu* or *xevvu*, on the other, looked more like a French word (*m* and *l* are frequently doubled; *x* and *v* are never doubled). A preference for a nonword such as *lammi* over *llami* but not for *xevvu* over *xevvu* would mean that no transfer occurred and would be indicative of a failure of abstraction. On the other hand, if participants were to also express a preference for *xevvu* over *xevvu*, this would indicate that transfer did occur.

Results from the letter-set change paradigm in artificial grammar learning (e.g., Dienes & Altmann, 1997; Gomez, 1997; Manza & Reber, 1997; Mathews et al., 1989; Reber, 1969; Shanks, Johnstone, & Staggs, 1997; Whittlesea & Wright, 1997) suggest that some transfer will occur but that performance with the studied material will remain better than performance with the unseen material. However, this transfer decrement may be due to the limited amount of practice given in laboratory context. Transposing the abstraction issue in natural context is thus a consideration of special interest because it allows us to study whether this difference evolves across grade level. In keeping with a widespread view, let us assume that empirical regularities are initially learned through sensitivity to the idiosyncratic properties of the material and that abstract knowledge progressively evolves from this initial sensitivity. We should observe an initial stage in which no transfer occurs, although there is evidence of learning when the test material is identical to the studied material, followed by the progressive appearance of some transfer. This evolution should presumably end with a rule-governed pattern of behavior, in which performance is no longer influenced by whether the material has been seen before or not (Manza & Reber, 1997). Any discrepancy with regard to this anticipated pattern would suggest that the

abstractionist view, at least the one relying on a stringent definition of rule (Anderson, 1993; Smith, Langston, & Nisbett, 1992; Whittlesea & Dorken, 1997) needs revision.

Method

Materials. Twelve nonword pairs were constructed to test participants' sensitivity to the identity of consonants that are doubled in French. Each nonword of a pair included a double consonant in the medial position. In one nonword it was a frequent double consonant (e.g., *ff* or *ll*), and in the other it was a doublet formed with consonants that are never doubled in French (e.g., *hh* or *jj*). Apart from the doublets, the constituent letters of the two nonwords of a pair were the same. Examples of these pairs are *ullate-ujate* and *bihhot-billot*. Hereafter, these pairs are noted *CCf* and *CCn* (with *f* denoting frequent and *n* denoting never).

We used 24 pairs of six-letter nonwords to test participants' sensitivity to the fact that vowels cannot be doubled in French. The constituent letters of the two nonwords of a pair were the same. The difference was that one nonword included a double consonant and the other a double vowel. The double consonant and the double vowel were both situated in the medial position. For 12 pairs noted *VV/CCf*, the double consonant was a double consonant frequently doubled in French (e.g., *ll* or *mm*). Examples of these pairs are *tillos-tillos* and *buumer-buumer*. For 12 other pairs, noted *VV/CCn*, the double consonant was formed with consonants that are never doubled in French (e.g., *hh* or *jj*). Examples of these pairs are *ruujer-rujjer* and *ahhire-ahhire*.

We used 24 pairs of six-letter nonwords to test participants' sensitivity to the position in which consonants can be doubled in French. The constituent letters of the two nonwords of a pair were the same. The difference was that one nonword included a double consonant placed in the medial (legal) position and the other a double consonant situated in the initial or final (illegal) position. As far as the illegal position was concerned, the doublet was at the beginning of the nonword for half of the items and at the end for the other half. Twelve pairs noted *Position CCf* included doublets formed using consonants frequently doubled in French (e.g., *ll*, *nn*, or *rr*). Examples of these pairs are *nnulor-nullor* and *golirr-gollir*. Twelve pairs noted *Position CCn* included doublets formed using consonants that are never doubled in French (e.g., *hh* or *xx*). Examples of these pairs are *xihhel-xxihel* and *rixohh-rixohh*.

For each type of pair, legal nonwords were on the right in half of the trials and on the left in the other half. All the nonword pairs described above were written on sheets of paper, placed in a random order and stapled together behind three nonword practice pairs (those used in Experiment 1) to form a booklet, as in Experiment 1. Each page of the booklet was numbered (at the top right). Complete stimulus lists are presented in Appendix A.

Procedure. The procedure used in Experiment 2 was exactly the same as in Experiment 1.

Participants. Participants included 20 first graders (mean age: 6 years, 11 months), 20 second graders (mean age: 7 years, 10 months), 20 third graders (mean age: 8 years, 6 months), 20 fourth graders (mean age: 10 years, 0 month), and 20 fifth graders (mean age: 10 years, 10 months). The experiment was conducted at the end of the school year.

Results

Identity of the consonants that are doubled in French. The selection rate of nonwords including a frequent double consonant (*CCf*) rather than a doublet formed with consonants that are never doubled in French (*CCn*) are shown in Figure 2. Analyses using *t* tests showed that, at each grade level, participants expressed a preference for nonwords containing *CCf* over nonwords containing *CCn*, all $t(19) > 7.3$, $p < .001$. ANOVAs with the variable

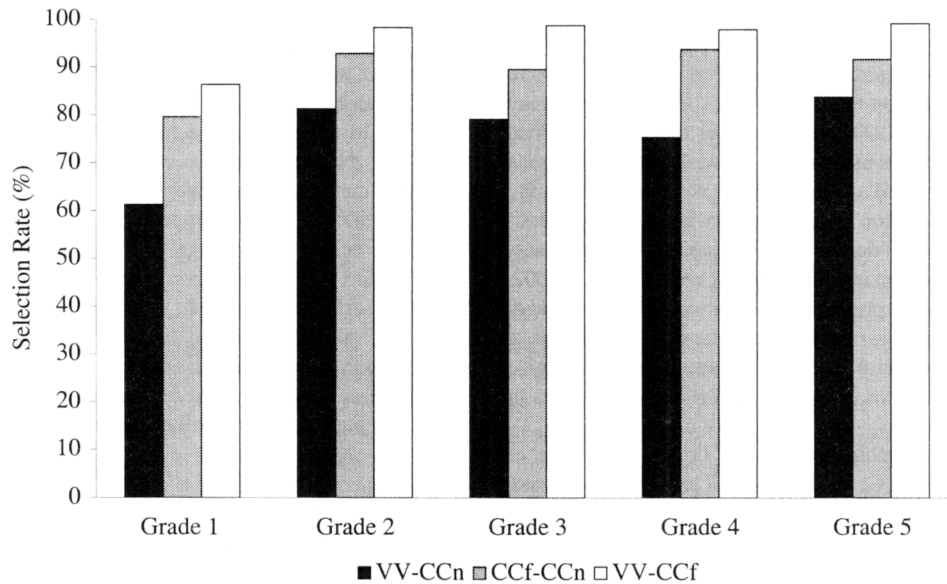


Figure 2. Selection rate of nonwords including a doublet of never doubled consonants, rather than a double vowel (VV-CCn); a frequent double consonant, rather than a doublet of never doubled consonants (CCf-CCn); a frequent double consonant, rather than a double vowel (VV-CCf) (Experiment 2).

grade level (1 to 5) were conducted. The dependent variable was the selection rate of nonwords including CCf. There was a main effect of grade level: $F_1(4, 95) = 4.4, MSE = 2.2, p < .002; F_2(4, 44) = 8.3, MSE = 1.9, p < .001$. The selection rate of nonwords including CCf (rather than CCn) increased significantly from Grade 1 (80%) to Grade 2 (93%), $F_1(1, 95) = 11.9, MSE = 2.2, p = .001; F_2(1, 11) = 16.0, MSE = 2.7, p = .002$, but did not vary significantly from Grade 2 to Grade 5 ($F_1 < 1; F_2 < 1$). Thus, sensitivity to the identity of the consonants that can be doubled was observed as early as in Grade 1 and increased from Grade 1 to Grade 2.

Only consonants (not vowels) can be doubled in French. Figure 2 shows the selection rate of nonwords including (a) a frequent double consonant rather than a double vowel (VV/CCf pairs) and (b) a doublet formed with consonants that are never doubled in French rather than a double vowel (VV/CCn pairs). Analyses using *t* tests showed that at each grade level, participants expressed a preference for nonwords containing CCf over nonwords containing VV, all $t_s(19) > 9.8, p < .001$. More interesting with regard to the abstraction issue is the case of VV/CCn nonword pairs. At each grade level, nonwords containing CCn were more often chosen than nonwords containing VV. This effect was nevertheless significant only from Grade 2 and above: $t(19) = 1.7, p = .10$, in Grade 1; all $t_s(19) > 8.0, p < .001$, in Grade 2 and above. Thus, children preferred nonwords such as *tukke* over *tuuke* even though they had never seen *kk* nor *uu*.

ANOVAs with the variables grade level (1 to 5) and type of nonword pairs (VV/CCf vs. VV/CCn) were conducted. The dependent variable was the selection rate of nonwords including a double consonant. There was a main effect of grade level: $F_1(4, 95) = 5.9, MSE = 4.9, p = .002; F_2(4, 88) = 22.5, MSE = 2.2, p < .001$. The selection rate of nonwords including CC increased significantly from Grade 1 (74%) to Grade 2 (90%), $F_1(1, 95) = 15.1, MSE = 4.9, p < .001; F_2(1, 22) = 43.5, MSE = 2.8,$

$p < .001$, and did not vary significantly from Grade 2 to Grade 5 ($F_1 < 1; F_2 < 1$). There was a main effect of the type of nonword pairs: $F_1(1, 95) = 121.6, MSE = 2.3, p < .001; F_2(1, 22) = 102.2, MSE = 4.7, p < .001$. Nonwords including double consonants were selected more often for VV/CCf (96%) than for VV/CCn pairs (76%). The Grade Level \times Type of Nonword Pairs interaction was nonsignificant ($F_1 < 1; F_2 < 1$). This lack of interaction means that the difference between performance involving abstraction processes (VV/CCn) and performance that can be accounted for without any form of abstraction (VV/CCf) did not vary significantly as a function of grade level. To examine whether this parallelism over grade level extended to between-subjects differences, we computed the correlation between the selection rates of double consonants for VV/CCf pairs and for VV/CCn pairs. When computed over the entire sample, the Pearson correlation was .58 ($p < .001$). Because this value may be due to the effect of grade level, the overall correlation was partially redundant with the information provided by the ANOVA. However, when grade level was partialled out, the between-subjects correlation was still a substantial .55 ($p < .001$). Thus, both ANOVAs and correlational analyses indicate a parallelism on performance that involves abstraction (i.e., VV/CCn pairs), and performance that can be accounted for without any form of abstraction (i.e., VV/CCf pairs).

Note also that nonwords including CCf were selected more often when the other member of the pair was a VV (96%) rather than a CCn (90%): $F_1(1, 95) = 32.5, MSE = 0.96, p < .001; F_2(1, 22) = 16.2, MSE = 3.2, p < .001$. This difference was stable across grade level as attested by the nonsignificant Grade Level \times Type of Nonword Pairs interaction ($F_1 < 1; F_2 < 1$). Thus, participants always performed better on pairs for which selection could additionally depend on an abstract property (higher selection of CCf for VV/CCf pairs than for CCn/CCf pairs) regardless of

grade level, and the difference between these two types of pairs persisted across grade level.

Position in which a double consonant can occur in French. Figure 3 shows the selection rate of nonwords including a doublet in the medial (legal) position rather than in the initial or final (illegal) positions. Analyses using *t* tests showed that at each grade level, participants expressed a preference for nonwords containing a doublet in a legal position (rather than in an illegal position) regardless of whether the doublet was formed with consonants frequently or never doubled in French, all $t_s(19) > 3.7, p < .002$. Thus, children expressed a preference for nonwords such as *jukker* over nonwords such as *jjuker* even though they had never seen *jj* nor *kk*. This result provides a clear indication that children's sensitivity to the legal position of double consonants transfers to unseen double consonants.

ANOVAs with the variables grade level (1 to 5) and type of nonword pairs (Position CCf vs. Position CCn) were conducted. The dependent variable was the selection rate of nonwords including a doublet in the medial position. There was a main effect of grade level: $F_1(4, 95) = 3.1, MSE = 4.2, p = .018$; $F_2(4, 88) = 8.3, MSE = 2.6, p < .001$. The selection rate of nonwords containing a doublet in a legal position increased significantly from Grade 1 (74%) to Grade 2 (85%), $F_1(1, 95) = 7.8, MSE = 4.2, p = .006$; $F_2(1, 22) = 16.5, MSE = 3.3, p = .001$, but did not vary significantly from Grade 2 to Grade 5 ($F_1 < 1$; $F_2 < 1$). There was a main effect of the type of nonword pairs (Position CCf vs. Position CCn): $F_1(1, 95) = 168.8, MSE = 1.2, p < .001$; $F_2(1, 22) = 13.6, MSE = 23.8, p = .001$. Nonwords including a doublet in legal position were chosen more often when the constituent letters of the doublets were frequently doubled consonants (91%) rather than never doubled consonants (74%). This difference, which reflects transfer decrement, was constant across grade level as attested to by the nonsignificant Grade Level \times Type of Nonword Pairs interaction ($F_1 < 1$; $F_2 < 1$). This absence of

interaction suggests that the difference between performance on new material, for which abstractive processes are involved (Position CCn), and performance that can be accounted for without any form of abstraction (Position CCf) did not vary significantly as a function of grade level. As for the property explored previously (i.e., the fact that vowels cannot be doubled), we examined whether the parallelism over grade level extended to between-subjects differences by computing correlation between the selection rates of doublets in legal position for Position CCf pairs and for Position CCn pairs. When computed on the entire sample, the Pearson correlation was .64 ($p < .001$). When grade level, which may be responsible for part of the correlation, was partialled out, the between-subjects correlation was still a sizable .63 ($p < .001$). Thus, these analyses indicate a between-grades and between-subjects performance parallelism on transfer and nontransfer material.

Discussion

Experiment 2 confirmed Cassar and Treiman's (1997) findings that children possess some knowledge about orthographic regularities much earlier than traditional stage models of spelling development would predict (Ehri, 1986; Henderson, 1985). For instance, as early as Grade 1, children were sensitive to the identity of consonants that can or cannot be doubled, to the fact that vowels can never be doubled, and to the legal position of double consonants.

Experiment 2's results further suggest that children's orthographic knowledge is in some sense abstract. This conclusion emerges from (a) the fact that children preferred nonwords with a doublet formed by consonants that are never doubled over nonwords including a doublet formed with vowels and (b) the fact that children's sensitivity to the legal position of double consonants extended to consonants that are never doubled. With respect to the

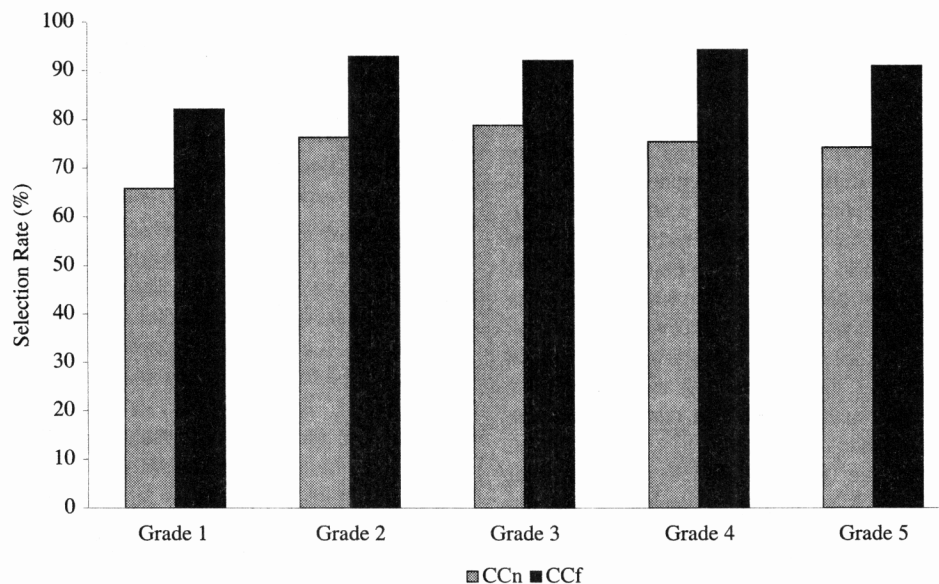


Figure 3. Selection rate of nonwords including a doublet in the medial (legal) position, rather than in the initial or final (illegal) positions for both never doubled consonants (CCn) and frequently doubled consonants (CCf) (Experiment 2).

legal position of double consonants, the high score observed in the transfer condition contrasts with the low (but better-than-chance) performance typically observed in laboratory contexts. This discrepancy illustrates the value of transposing laboratory research to realistic situations, where the effects of very extensive training can be assessed. However, whatever the rule we consider, performance was always better with frequently doubled consonants (rather than doublets of never doubled consonants), and there was a parallelism between performance on seen and unseen material. This transfer decrement phenomenon, which was stable across grade level, suggests that children's orthographic behavior cannot be readily described as rule based. The theoretical implications of these results are postponed until the General Discussion.

Experiments 3a and 3b: Overview

Experiments 3a and 3b were conducted to explore whether the results of Experiments 1 and 2 extend to production (completion) tasks. Experiment 3a was aimed at generalizing the results relative to children's sensitivity to the frequency of double consonants, an orthographic regularity that *cannot* be described with an abstract rule. Experiment 3b was aimed at generalizing the results of Experiment 2 about the legal position of double consonants, an orthographic regularity that *can* be described with an abstract rule.

Experiment 3a

Method

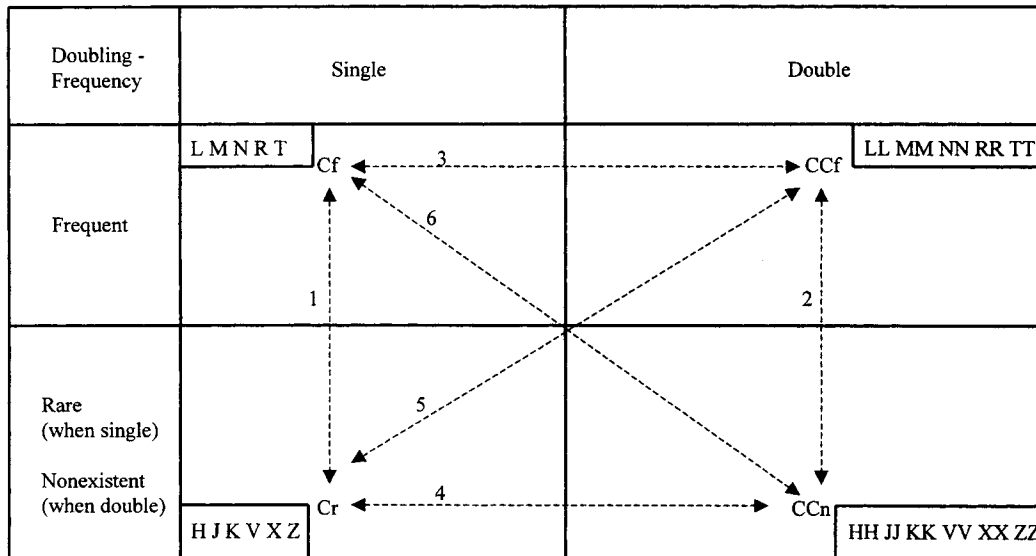
Materials. The stimuli consisted of trisyllabic nonwords with a blank in the medial position (e.g., *tuba_ir*; *u_otir*) and two patches that were placed in the middle of the bottom line. Each patch consisted of

a single consonant or a double consonant. Combining the format (single vs. double) and frequency (frequent vs. never for the double format; frequent vs. rare for the single format) factors produced four types of patches: frequent single consonant (Cf); rare single consonant (Cr); frequent double consonant (CCf); and never double consonant (CCn). The pairwise combination of these four types of patches resulted in the creation of six types of pairs as illustrated in Figure 4 (Cf/Cr, Cf/CCf, Cr/CCf, CCf/CCn, Cf/CCn, Cr/CCn). Twelve pairs of each type were created.

For each type of item, one type of patch was on the right for half of the trials and on the left for the other half. The 72 items were written on sheets of paper, placed in random order and stapled together just next to three practice items to form a booklet as in Experiments 1 and 2. The practice items involved no ambiguity in filling the blank since only one alternative was valid in French (e.g., *rato_er*, with *a* and *t* as alternatives). Each page of the booklet was numbered (at the top right). Complete stimulus lists are presented in Appendix A.

Procedure. Children were told that the experimenter had made up new words that no one had ever seen or heard. Afterward the experimenter specified that these new words contained a blank. At this point the children were asked to look at the first practice nonword and to fill the blank with one of the two proposals located below this new word so that it most resembled a word. This procedure was repeated with the remaining two practice items before the actual experiment began. As in Experiments 1 and 2, the experimenter called out item numbers.

Participants. Participants were 20 first graders (mean age: 6 years, 4 months), 20 second graders (mean age: 7 years, 9 months), 20 third graders (mean age: 8 years, 2 months), 20 fourth graders (mean age: 9 years, 9 months), 20 fifth graders (mean age: 10 years, 5 months), and 20 sixth graders (mean age: 11 years, 4 months). Testing took place at the beginning of January (i.e., 4 months after the beginning of the school year).



- (1) Cf/Cr = u_otir (k/n)
- (2) CCf/CCn = tuba_ir (kk/ll)
- (3) Cf/CCf = i_oter (l/mm)
- (4) Cr/CCn = i_aler (z/kk)
- (5) Cr/CCf = tago_er (h/nn)
- (6) Cf/CCn = juba_er (m/xx)

Figure 4. Material used in Experiment 3a. Cf = frequent single consonant; Cr = rare single consonant; CCf = frequent double consonant; CCn = never double consonant.

Results

Table 1 shows the completion rate for each type of items mentioned in Figure 4. The study of the Cf/Cr items allowed us to assess children's sensitivity to the frequency of single consonants. At each grade level, children used Cf more often than Cr, all $t(19) > 3.7$, $p < .001$. The main effect of grade level was significant: $F_1(5, 114) = 11.0$, $MSE = 2.4$, $p < .0001$; $F_2(5, 55) = 13.0$, $MSE = 3.4$, $p < .0001$. The use of Cf rather than Cr increased significantly from Grade 1 (66%) to Grade 2 (82%): $F_1(1, 114) = 14.9$, $MSE = 2.4$, $p = .0002$; $F_2(1, 11) = 9.6$, $MSE = 6.3$, $p = .01$.

The study of participants' behavior for items for which one of the two patches was a doublet of never doubled consonants allowed us to explore children's knowledge of the identity of the consonants that can (or cannot) be doubled. At each grade level, and for each of the three types of items (Cf/CCn; Cr/CCn; CCf/CCn), CCn items were used less frequently than the alternatives, all $t(19) \geq 3.2$, $p \leq .004$. Note, however, that a wider use of Cf for Cf/CCn items and of CCf for CCf/CCn items could simply reflect children's sensitivity to the frequency of single consonants (see criticisms with regard to Cassar & Treiman, 1997).

Sensitivity to the frequency of single consonants alone does not prove a preference for Cr over CCn for Cr/CCn items (e.g., *ra_olit-ra_olit; xkk*). Because the critical consonants (e.g., *x, k*) are infrequent in single format but are never doubled, the fact that at each grade level, children used Cr more often than CCn (78% on average), all $t(19) > 5.7$, $p < .001$, would reflect children's sensitivity to the consonants' frequency of doubling. Alternatively, this result could be due to a general trend of using single consonants more often than double consonants, irrespective of the identity of the consonants. In this case, the completion rate with single (rather than double) consonants should not differ for Cr/CCn and

Cf/CCf items. By contrast, if the consonants' frequency of doubling influences participants' completions, the completion rate with Cr for Cr/CCn items should be higher than the completion rate with Cf for Cf/CCf items. ANOVAs with the variables grade level (1 to 6) and type of items (Cr/CCn vs. Cf/CCf) were conducted. The dependent variable was the completion rate with a single, rather than a double, consonant. These analyses indicate that single consonants were used more frequently for Cr/CCn items (78%) than for Cf/CCf items (54%): $F_1(1, 114) = 158.5$, $MSE = 3.1$, $p < .001$; $F_2(1, 22) = 14.4$, $MSE = 55.9$, $p = .001$. The main effect of grade as well as the Grade \times Type of Items interaction were significant: $F_1(5, 114) = 2.2$, $MSE = 5.7$, $p = .058$; $F_2(5, 110) = 6.0$, $MSE = 3.5$, $p < .001$; and $F_1(5, 114) = 4.4$, $MSE = 3.1$, $p = .001$; $F_2(5, 110) = 6.4$, $MSE = 3.5$, $p < .0001$. This interaction can be explained as follows: Whereas completions with Cr for Cr/CCn items were more frequent than completions with Cf for Cf/CCf items from Grade 2 to Grade 6 (mean difference = 27%), all $t(19) \geq 3.8$, $p \leq .0012$, this difference was not significant in Grade 1 (difference = 6%), $t(19) = 1.5$, $p = .14$. Indeed, whereas first graders more frequently used single (rather than double) consonants for both Cr/CCn and Cf/CCf items, older children more frequently used single consonants only for Cr/CCn items. First graders' general preference for single consonants is also evidenced by the fact that first graders did not use Cr and CCf differently for Cr/CCf items (48%), $t(19) = 0.4$, $p = .73$, whereas older children used CCf more often than Cr for those items (79% on average), all $t(19) > 6.8$, $p < .0001$. Thus, the results obtained for the Cr/CCn items demonstrate that from the second grade onward, children are really sensitive to the consonants' frequency of doubling.

A second way of exploring children's sensitivity to the consonants' frequency of doubling consists of comparing the use of CCf

Table 1
Percentage Completion Rate for Each Type of Items

Type of items	Grade					
	1	2	3	4	5	6
Cf for Cf/Cr items						
<i>M</i>	65.8**	81.7**	87.1**	91.2**	86.7**	91.7**
<i>SD</i>	18.9	12.5	10.9	9.5	11.6	12.1
Cf for Cf/CCn items						
<i>M</i>	80.8**	80.0**	86.7**	92.1**	85.8**	95.8**
<i>SD</i>	15.5	20.7	13.3	14.4	17.7	6.9
Cr for Cr/CCn items						
<i>M</i>	75.4**	77.1**	77.5**	79.6**	73.3**	85.4**
<i>SD</i>	20.0	17.5	19.3	14.9	17.6	15.3
Cf for Cf/CCf items						
<i>M</i>	69.6**	52.9	55.4	53.3	43.3*	51.7
<i>SD</i>	16.9	17.1	21.5	14.9	14.4	17.6
CCf for Cr/CCf items						
<i>M</i>	48.3	74.6**	74.6**	79.2**	84.2**	81.7**
<i>SD</i>	21.2	10.9	16.1	18.6	12.6	13.9
CCf for CCn/CCf items						
<i>M</i>	66.2**	85.0**	87.1**	91.2**	91.7**	95.0**
<i>SD</i>	22.4	13.1	13.1	10.6	9.4	8.3

Note. Cf = frequent single consonant; Cr = rare single consonant; CCn = never double consonant; CCf = frequent double consonant.

* $p < .05$. ** $p < .001$.

for Cr/CCf items versus CCn/CCf items. If children are sensitive only to the frequency of single consonants, CCf should not be used differently for Cr/CCf items versus for CCn/CCf items. By contrast, if children are sensitive to the consonants' frequency of doubling, CCf should be used more often for CCn/CCf items than for Cr/CCf items. Note that a more frequent use of CCf for CCn/CCf items than for Cr/CCf items from the second grade onward could not be explained by a general preference for single consonants since children did not use Cf and CCf differently for Cf/CCf items (see above). ANOVAs with the variables grade level (1 to 6) and type of items (Cr/CCf vs. CCn/CCf) were conducted. The dependent variable was the completion rate with CCf rather than CCn or Cr. The use of CCf differed as a function of grade: $F_1(5, 114) = 19.1, MSE = 4.1, p < .0001$; $F_2(5, 110) = 28.6, MSE = 4.5, p < .001$. The use of CCf was significantly higher in Grades 2 to 6 (84%) than in Grade 1 (57%), $F_1(1, 114) = 86.5, MSE = 4.1, p < .0001$; $F_2(1, 22) = 47.6, MSE = 12.4, p < .001$, and did not differ significantly from Grade 2 to Grade 6 ($F_1 < 1$; $F_2 < 1$). The CCf patches were more often used for CCn/CCf items (86%) than for Cr/CCf items (74%): $F_1(1, 114) = 57.6, MSE = 2.3, p < .001$; $F_2(1, 22) = 7.0, MSE = 30.9, p < .001$. This effect of type of items did not differ as a function of grade level, as attested by the nonsignificant Grade Level \times Type of Items interaction ($F_1 < 1$; $F_2 < 1$). Thus, the fact that second to sixth graders, who do not exhibit a general preference for single (rather than double) consonants, used Cr for Cr/CCf items more often than CCn for CCn/CCf items indicates that they were sensitive to the consonants' frequency of doubling.

Discussion

Experiment 3a was aimed at generalizing the conclusions drawn from Experiment 1 and Experiment 2 concerning children's sensitivity to the frequency of double consonants to a completion task. Experiment 3a revealed an almost systematic preference for single over doubled consonants in Grade 1. This preference was no longer observed in older children. Sensitivity to the frequency of single and double consonants was observed as early as in Grade 1, and increased from Grade 1 to Grade 2. More important, Experiment 3a showed that from Grade 2 onward, children's completions were clearly influenced by the frequency of double consonants even when the potential confound of the component (single) consonants was removed. Thus, Experiment 3a showed that sensitivity to the frequency of double consonants reported in Experiment 1 through a judgment task also influenced children's behavior in production tasks.

Experiment 3b

Experiment 3b explored children's sensitivity to the legal position of double consonants in French by means of a completion task. Nonwords were presented in pairs, with a blank in the medial position in one nonword (e.g., *a_iger* or *gira_er*) and at the beginning or at the end in the other (e.g. *_iraver*, *alogi_*). Participants had to fill the two blanks with two alternatives presented below each nonword pair. One of these two alternatives was a single consonant, the other a double consonant. This somewhat indirect method was needed in order to avoid biasing the probability of correct responses. Indeed, asking children to fill in initial,

medial, or final blanks in nonwords with a single or double consonant would have been problematic because completion with a single consonant is legal in all three positions, whereas completion with a double consonant is legal only in the medial position. As in Experiment 2, the abstract character of children's knowledge about the legal position of double consonants in French was assessed with consonants that are frequently or never doubled in French.

Method

Materials. Thirty-six pairs of trisyllabic nonwords, including one nonword with a blank located in the medial position (legal position for a doublet) were built. The other nonword of a pair included a blank situated at the beginning for half of the items and at the end for the other half (illegal position for a doublet). Below each of these 36 pairs were two alternatives. One consonant was presented in isolation; the other as a doublet.

In 24 pairs, the alternatives proposed to fill the blank in the nonwords were consonants that are frequently doubled in French (*l, m, n, p, r, or t*). Examples of these pairs, labeled Position CCf, are *iméva__i__éva* (*t* and *rr* as alternatives) and *gira_er__iraver* (*ll* and *m* as alternatives). In 12 pairs, the two alternatives proposed to fill the blank in the nonwords were consonants which are never doubled in French (*h, j, k, w, x, z*). Examples of these pairs, labeled Position CCn, are *a__oli__adoli__* (*j* and *ww* as alternatives) and *ro__ébit__orébit* (*z* and *kk* as alternatives). There were also three practice pairs including a blank, with two letters (in single format) below them (e.g., *ralo__ra_on* with *u* and *b*).

For the two types of pairs (Position CCf and Position CCn), nonwords with a blank in the medial position were on the right for half of the trials and on the left for the other half, and vice versa for nonwords with a blank at the beginning or at the end. This distribution also held for the two proposed consonant completions (one in single format, the other in double format). The 36 pairs described above were written on sheets of paper (21 cm \times 3 cm) using lowercase letters in 14-point New York type. They were put in a random order and stapled together just behind the three practice nonword pairs to make a booklet. Each page of the booklet was numbered (at the top right). Complete lists of the stimuli are presented in Appendix A.

Procedure. Children were told that the experimenter had made up new words that no one had ever seen or heard. Next, the experimenter indicated that these new words were presented in pairs and that they each contained a blank. At this point, the children were asked to look at the first practice nonword pair and to complete each nonword with one of the two possibilities located below the new words in such a way as to make each nonword look most like a word. This procedure was repeated with the two remaining practice items before the actual experiment began.

Participants. The same participants took part in Experiment 3b as in Experiment 3a. Experiment 3b was run one week before Experiment 3a.

Results

Figure 5 shows the completion rates with the doublet in the medial position and the single consonant in initial or final positions (henceforth, correct completions). Analyses using *t* tests show that at each grade level, correct completions outnumbered incorrect completions for Position CCf pairs, all $ts(19) > 2.5, p < .02$. This effect was significant only from Grade 2 and above for Position CCn pairs: $t(19) = 2.0, p = .06$, in Grade 2; all $ts(19) > 3.4, p < .003$, from Grade 3 and above. Thus, as early as in the beginning of the first grade, children's sensitivity to the legal position of double consonants influenced their completions when frequently

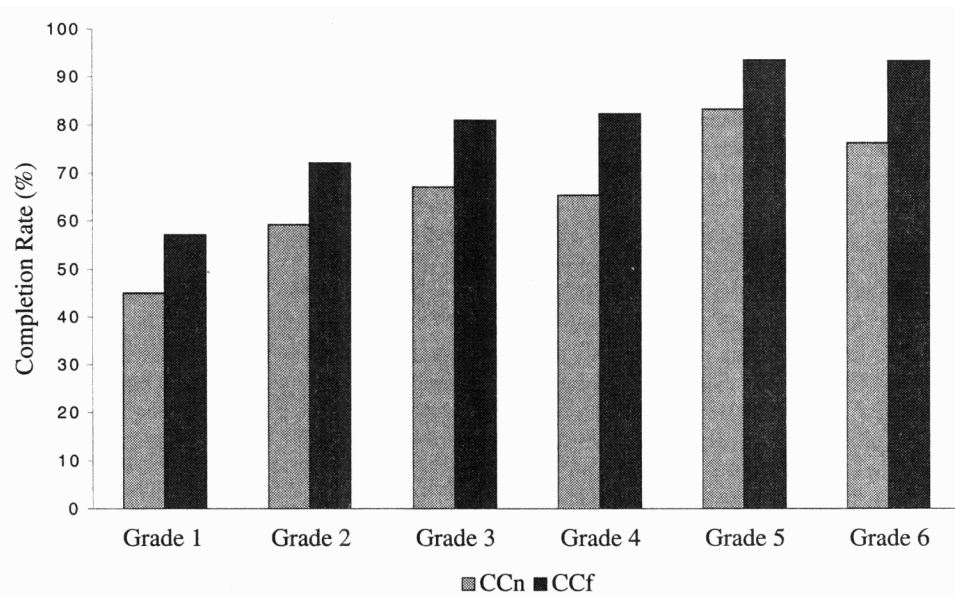


Figure 5. Completion rate with the doublet in the medial position and the single consonant in initial or final positions (Experiment 3b). CCn = never double consonant; CCf = frequent double consonant.

doubled consonants were used, and this result extended to never doubled consonants from Grade 2 onward.

ANOVAs with the variables grade level (1 to 6) and type of items (Position CCf vs. Position CCn) were conducted. The dependent variable was the rate of correct completions. There was a main effect of grade level: $F_1(5, 114) = 15.1$, $MSE = 27.9$, $p < .001$; $F_2(5, 170) = 56.4$, $MSE = 4.1$, $p < .001$. The rate of correct completions increased as a linear trend of grade level: $F_1(1, 114) = 63.5$, $MSE = 27.9$, $p < .001$; $F_2(1, 34) = 151.2$, $MSE = 6.5$, $p < .001$. There was a main effect of the type of items, $F_1(1, 114) = 98.1$, $MSE = 6.8$, $p < .001$; $F_2(1, 34) = 14.6$, $MSE = 25.3$, $p < .001$: Correct completions were more frequent for Position CCf items than for Position CCn items (80% vs. 66%). This difference, which reflects transfer decrement, was stable across grade levels, as suggested by the nonsignificant Grade Level \times Type of Items interaction ($F_1 < 1$; $F_2 < 1$). This absence of interaction means that the difference between performance on new material for which abstractive processes are involved (Position CCn) and performance that can be accounted for without any form of abstraction (Position CCf) did not vary significantly as a function of grade level. As for the judgment task, we examined whether the parallelism over grade level extended to between-subjects differences by computing the correlation between the rate of correct completions for Position CCf items and for Position CCn items. When computed on the entire sample, the correlation was .74 ($p < .001$). As in Experiment 2, the correlation computed on the entire sample was sensitive to age-related differences and hence was partially redundant with the information conveyed by the ANOVA. However, when grade level was partialled out, the between-subjects correlation was still a sizable .64 ($p < .001$). Thus, as with the judgment task, these analyses are indicative of a between-grades and between-subjects parallelism of performance on transfer and nontransfer material.

Discussion

The purpose of Experiment 3b was to generalize the conclusions from Experiment 2 concerning the legal position of double consonants to a completion task. Participants were asked to fill blanks located in the medial position in one nonword of a pair, and in initial or final positions in the other. They were offered two possibilities: a single consonant or a double consonant. As early as in Grade 1, children's sensitivity to the legal position of double consonants influenced their completions when frequently doubled consonants were used, and this result extended to never doubled consonants from Grade 2 onward. The relationship between performance in transfer and nontransfer situations for children tested only 4 months after the beginning of the school year is interesting because it clearly eliminates a possible scenario according to which rules would be present from the beginning and would determine the entire improvement in performance. Indeed, in the case of a developmental primacy of rules, performance levels in both transfer and nontransfer conditions should have improved in concert. Also of particular interest with regard to the abstraction issue is the higher rate of correct completions when frequently doubled consonants (Position CCf items) rather than never doubled consonants (Position CCn items) were used. Furthermore, this difference, which corresponds to the transfer decrement, was stable across grade levels, and extended to between-subject differences.

General Discussion

Laboratory studies of implicit learning have had relatively little impact on the issue of learning in real-world situations so far. By and large, workers in domain-specific topics involving learning issues seemingly ignore the literature on implicit learning, and researchers from the implicit learning domain have not assessed

whether conclusions based on laboratory studies apply to real-world contexts. In this article, we have attempted to contribute to fill the gap between laboratory research and natural learning situations by exploring whether the conclusions issued from the former generalize to the latter.

Our second and main goal was to address a central issue in the implicit learning literature, namely, the abstract character of knowledge acquired during an implicit learning episode. Evidence relevant to this issue comes essentially from the so-called letter-transfer paradigm, in which participants are asked to judge the grammaticality of a set of stimuli that have the same deep structure but are instantiated with completely novel surface features. Typical results indicate (a) that participants are still able to perform above chance on the transfer discrimination task, although (b) their level of performance is very low and, in all cases, lower than in standard transfer conditions, where the items' surface features remain unchanged (Dienes & Altmann, 1997; Gomez, 1997; Manza & Reber, 1997; Mathews et al., 1989; Reber, 1969; Shanks et al., 1997; Whittlesea & Wright, 1997). The underlying problematic issue that we have attempted to address in our study is whether the observation of successful transfer to novel surface forms necessarily indicates, as some authors claim, that abstract rule-based knowledge has been acquired (e.g., Knowlton & Squire, 1996; Reber, 1993). Indeed, not only are statistical approaches capable of accounting for at least some aspects of the data, but the fact that performance on exemplars instantiated with novel features tends to be systematically lower than performance on exemplars instantiated with familiar features is at odds with the assumptions of rule-based, abstractionist approaches. This reasoning only holds, however, if one can rule out that participants have had the opportunity to develop fully abstract knowledge in the first place, that is, if one can assume that they had sufficient experience with the stimulus material.

To overcome the fact that the time course of performance over very long duration is difficult to explore in laboratory settings, we transposed the problematic of implicit learning to a real-world context. Accordingly, we explored whether participants' behavior tended to become rule directed after massive amounts of practice impossible to obtain in laboratory settings. Acquired sensitivity to certain orthographic regularities based on experience of printed language was chosen as a real-world learning situation because it fulfilled several desirable properties for well-controlled experiments.

In the following discussion, we summarize our main empirical results and assess their implications about the underlying theoretical issues. We also describe simulation work meant to explore the extent to which the sorts of learning mechanisms typical of statistical approaches are capable of accounting for our data.

Summary of the Main Results

The experiments reported in this article explored three characteristics of the use of double letters in French: the identity of consonants that can be doubled, the legal position of double consonants, and the fact that vowels cannot be doubled. Cassar and Treiman (1997) have suggested that children were sensitive to the frequency of double consonants because first and second graders judged that *noss* and *yill* look more like real words than *novv* and *yihh*. In contrast, we argued that such a preference may not reflect

children's sensitivity to the frequency of double consonants because it is possible that children express a preference for *yill* over *yihh* simply because *l* is more frequent than *h*, rather than because *l* is more frequently doubled than *h*. To get around this methodological difficulty, Experiment 1 contrasted consonants that are frequent in both single and double formats (e.g., *l* or *m*) versus consonants that are frequent in single format but not in double format (e.g., *c* or *d*). With this potential confound removed, our results indicated that children were sensitive to the frequency of double consonants as early as in Grade 1 and that this sensitivity increased with grade level. Through a judgment task, Experiment 2 showed that as early as in Grade 1, children were sensitive to the identity of the consonants that can or cannot be doubled with a material similar to the one used in Cassar and Treiman's study. Finally, Experiment 3a showed that sensitivity to the frequency of double consonants (Experiment 1) and to the identity of consonants that can be doubled (Experiment 2) also influenced children's behavior in production (completion) tasks. The fact that similar results were obtained in judgment and completion tasks is congruent with Treiman's (1993) studies, in which orthographic regularities influenced both children's writings (naturalistic study) and judgments of nonwords (experimental studies).

The two other orthographic regularities that were considered can be described by a rule. Experiment 2 showed that as early as in Grade 1, children expressed a preference for nonwords such as *tumme* over nonwords such as *tuume*, suggesting that they were sensitive to the fact that vowels can never be doubled in French. Moreover, from Grade 2 onward, children expressed a preference for nonwords that included a doublet of never doubled consonants (e.g., *kk*) over nonwords that included a double vowel (e.g., *uu*), despite the fact that neither *kk* nor *uu* had ever been seen previously. However, at each grade level, nonwords including a double vowel tended to be rejected more often when the other nonword of the pair included a frequently doubled consonant rather than a doublet of never doubled consonants, and this difference remained very stable across grade levels. It may be hypothesized that children's preference for doublets of never doubled consonants results from the fact that they are easier to pronounce than double vowels. The simulations reported below provide partial support for this interpretation but also show that some important aspects of the results can be accounted for without appealing to phonological information.

The second orthographic regularity that can be described with an abstract rule, namely the legal position of double consonants, is more interesting in that the format (single vs. double) of the consonants used in Experiments 2 and 3b does not lead to a modification in the phonological form of the word. In Experiment 2, as early as in Grade 1, children were sensitive to the legal position of double consonants. For instance, they judged that a nonword such as *bummor* was more wordlike than a nonword such as *bumorr*. This result is congruent with those reported by Cassar and Treiman (1997; e.g., the preference for *nuss* over *nnus*). However, Cassar and Treiman provided only limited evidence that children's knowledge of the legal position of double consonants was general (i.e., independent of the identity of the consonants that constitute the doublets). In fact, children's preference for *juss* over *jjus* in Cassar and Treiman's study may just reflect their knowledge of the identity of allowable doublets. In order to get around this difficulty, we used nonwords such as *bukcox* and *bukoxx*, in

which doublets in both legal and illegal positions were formed with consonants that can never be doubled in French. Thus the training and test stimuli have different surface forms but the same underlying abstract structure, as in the letter-transfer paradigm commonly used in artificial grammar learning. Our results showed that as early as Grade 1, children's sensitivity to the legal position of double consonants extended to never doubled consonants. However, performance was always better when frequently doubled consonants (rather than doublets of never doubled consonants) were used, and there was a parallelism between performance on seen and unseen material. Finally, Experiment 3b showed that the sensitivity to the legal position of double consonants reported in Experiment 2 through a judgment task also influenced children's behavior in production (completion) tasks. The similarity between the results obtained in judgment and completion tasks is congruent with Treiman (1993), who showed that children did not often make errors such as *bbal* for *ball*. It was from Grade 2 onward that children's sensitivity to the legal position of double consonants extended to consonants that can never be doubled in French. The fact that children's sensitivity to the legal position of double consonants tested only 4 months after the beginning of the school year did not extend to consonants that can never be doubled in French suggests that rules are not present from the onset of exposure, and that they cannot be construed as determining all aspects of performance. More important, as in Experiment 2, performance was always better on frequently doubled consonants (rather than doublets of never doubled consonants), and there was a parallelism between performance on seen and unseen material.

Overall, the results reported in these experiments confirm and extend the results of recent studies indicating that orthographic regularities affect children's spellings much earlier than traditional stage models would predict. First, we showed that children were influenced by the frequency of double consonants when the potential confound of the component (single) consonants is removed in both judgment and completion task. Secondly, we showed that children's knowledge about the legal position of double consonants extends to never doubled consonants but that performance was always lower when doublets formed with never doubled consonants were used in both legal and illegal positions. Moreover, there was a between-grades and between-subjects parallelism of performance with frequently and never doubled consonants—an important result with regard to the abstraction issue, which we discuss in the next section.

On the Nature of Abstraction

There is consensus on the idea that transferring some piece of knowledge based on a subset of letters (e.g., the legal position of doubling for frequently doubled letters) to another set of letters in which this piece of information is not available (e.g., the legal position of doubling for never doubled letters) amounts to abstraction. However, interpretation of the phenomenon is controversial. Many authors have claimed that between-letters transfer is indicative of rule-based behavior (e.g., Manza & Reber, 1997). According to such approaches, when provided with sufficient exposure to the material, participants would end up acquiring an abstract rule such as "The initial letter of a word is never doubled." This kind of rule would be qualitatively different from idiosyncratic knowledge about specific letters or words, because the latter can be

directly captured in the data whereas the former needs inferential processing. Because the role of statistical learning mechanisms in the early phases of learning is unquestionable, such approaches thus result in a dualistic view in which abstract rules are based on statistical knowledge and subsequently applied to all relevant material, whether familiar or novel. This dualistic perspective is close to the theoretical position of many authors in the implicit learning literature (e.g., Meulemans & Van Der Linden, 1997).

However, as we pointed out in the introduction, such rule-based approaches are challenged by the transfer decrement phenomenon and by the possibility of accounting for transfer through statistical learning mechanisms that never involve the development of abstract, rule-based knowledge. We discuss both issues in the following paragraphs.

In an authoritative discussion on the use of abstract rules, Smith et al. (1992) posited as the first of their eight criteria for rule use that "performance on rule-governed items is as accurate with unfamiliar as with familiar material" (p. 7; see also Anderson, 1993). Likewise, Whittlesea and Dorken (1997) suggested that someone who learned a useful rule would have equal success in transfer and nontransfer situations. Laboratory studies of implicit learning, however, have consistently indicated that participants exhibit a transfer decrement effect—a result that appears inconsistent with the criterion listed above for the involvement of rule-based knowledge. In the introduction, we hypothesized that a possible argument to help reconcile such results with abstractionist approaches would be to assume that the development of abstract knowledge takes considerably more time than available in typical laboratory settings. Transfer decrement, from this perspective, would then merely be a reflection of incomplete rule-based abstraction. Our experiments provide evidence against this possibility: Transfer decrement persisted even after very extensive training. There was no trend toward a reduction of the magnitude of transfer decrement over the 5 years of training we examined. Our results therefore suggest that even after considerable familiarity with printed material, children's orthographic behavior cannot be readily qualified as rule based in the sense defined above.

Is there still a way to reconcile our data with the idea that performance is rule based? It may be argued either (a) that rules are specific, idiosyncratic, instead of general or (b) that rules have probabilistic instead of deterministic conditions of application.

According to the first interpretation (i.e., specific rules), rules would have limited scope and apply or not as a function of factors such as context and familiarity with the situation. Such an interpretation could account for the fact that performance differs according to the type of consonants (i.e., frequent vs. never doubled) involved in our experiments. However, this interpretation severely undermines what made rule-based knowledge so attractive to early cognitive theorists such as Chomsky (1965), namely the fact that rules should be sufficiently general and abstract so as to support successful application to an increasingly large number of new situations as they develop.

According to the second assumption, rules would be abstract and general but would have probabilistic conditions of firing. Formally, this interpretation retains the advantages characteristic of any rule system, namely that they can easily generalize to new instances. However, one may question the internal consistency of the notion of probabilistic rules, insofar as the involvement of a rule in performance is conventionally inferred on the basis of the

observation that the same actions are systematically performed in a given context. A further problem is that theories using probabilistic rules are not falsifiable because findings that do not follow the rules can always be interpreted as cases where the rules do not apply. Note that when taken to the limit, such theories also make it possible to imagine that we are endowed with an infinite number of rules, most of which are associated with a probability of application of zero—an assumption that is clearly unparsimonious.

In both Experiments 2 and 3b, we observed strong correlations (partialing out grade level) between performance levels on familiar and novel material. This close parallelism suggests that a single processing system underpins performance on both familiar and novel forms. The hypothesis that a single processing system can be responsible for both concrete and abstract forms of knowledge has received strong support from connectionist modeling over the last two decades. Numerous simulations have suggested (a) that the observation of sensitivity to an abstract dimension of the training material need not involve the acquisition of corresponding abstract rules and (b) that a single network is quite capable of processing both rule-abiding items and exceptions based on the same processing and representational resources. A very well known example is Rumelhart and McClelland's (1986) model of the acquisition of the past tense morphology. In the model, not only are regular verbs processed in just the same way as exceptions, but neither are learned through anything like processes of rule acquisition. A somewhat more recent example is provided by Cleeremans, Servan-Schreiber, and McClelland (1989), who showed that an SRN trained to process sequences generated from a finite-state grammar was subsequently able to act as a perfect finite-state recognizer for *any* expression constructed with the same tokens. In other words, the network, on the basis of necessarily limited experience with grammatical expressions, had developed representations of the training material that were sufficiently abstract to enable it to exhibit generalization to an infinite number of novel expressions (but instantiated with familiar features). The network's performance was thus indistinguishable from the performance of the corresponding symbolic, rule-based finite-state recognizer, in spite of the fact that the network's internal representations did not consist of abstract, symbolic rules. Of interest, Cleeremans et al. (1989) showed that, depending on factors such as task demands, the structure of the training set, and the number of hidden units, the network's internal representations may vary on a continuum that extends from exemplar-specific knowledge (e.g., many local associations) to highly synthetic representations corresponding to the states of the finite-state grammar used to generate the training material. These findings thus suggest that rulelike performance need not necessarily depend on rule-based knowledge and that systems based on the processing of exemplars can often end up being just as capable as their symbolic counterparts to process novel items.

We now address the question of whether a connectionist model is capable of accounting for the findings described in this article in the case of the written language.

Simulating Our Results With an SRN

To test whether children's performance on seen and unseen material in our experiments may in fact be accounted for by the operation of a single system, we explored whether an SRN was

capable of providing a qualitative account of our results. The SRN, initially proposed by Elman (1990; see also Cleeremans et al., 1989) is one of the most influential connectionist models in the implicit learning and psycholinguistic literatures. SRNs are typically trained to predict the next element of sequences presented one element at a time to the network and are therefore particularly appropriate to explore tasks involving sensitivity to sequential structure. The SRN is trained to *predict* the successive elements of sequences presented to it one element at a time. To perform this prediction task, the network is presented, on each time step, with element *t* of a sequence, and with a copy of its own internal state (i.e., the vector of hidden units activations) at time step *t* - 1. On the basis of these inputs, the network is to predict element *t* + 1 of the sequence. During training, the network's prediction responses are compared with the actual successor of the sequence, and the resulting error signal is then used to modify its connection weights using the back-propagation algorithm.

To assess the SRN's ability to capture our data, we trained an SRN on the most frequent 1,000 words of a computerized Brulex database for written and spoken French (Content, Mousty, & Radeau, 1990). The network used local representations on both its input and its output pools of units, and words were represented as a sequence of letters presented one a time to the network (see Appendix B for simulation details). During each epoch of training, a random number of words were selected for presentation to the network in a way that respected the frequency of occurrence of the words in the corpus. On average, about 10 words were presented during each epoch. The network was trained for a total of 10,000 epochs (i.e., approximately one million words). At various points during training, we froze the network's connections and tested the network on the set of nonwords used in our experiments by recording the network's prediction responses for each letter of the test items (i.e., the activation of the output unit corresponding to letter *t* + 1 of each nonword upon presentation of letter *t*). To control for random differences in the specific words experienced by the network during training, and for variance that results from the configuration of initial weights used by the network, we conducted 10 replications of this simulation and averaged the networks' responses.

To assess network performance, we first considered the activation of the output units corresponding to doubled letters in the nonwords of the test set. For instance, we examined the activation of the output unit corresponding to the letter *M* in the context of the word *fommir* after the network had been presented with *fom-*. The network's responses on all such cases were averaged separately for (a) consonants that are frequently doubled in French, (b) consonants that are never doubled in French (*k* and *w* were removed from the analysis, because these consonants never occur in the database)³ and (c) vowels. We further distinguished between con-

³ Because the letters are the input units of the network, and because *k* and *w* never occur in the database, the network cannot learn anything about *k* and *w*. Actual people, by contrast, are able to generalize to unseen letters because they have access to other features of the material, such as their form or their size. To make the network able to generalize to unseen letters, the material should thus have been coded using more elementary, lower level features, the combination of which would allow the encoding of both old and new material, as actual people presumably do.

sonants occurring at the beginning of nonwords or in their middle. Consonants occurring in the final position were removed from the analysis because information is provided sequentially to an SRN, and hence the network is not informed about whether the currently predicted letter ends the nonword.

As described below, the network quickly developed sensitivity to the frequency with which consonants can be doubled in the corpus. Thus, for instance, the network's prediction that frequently doubled consonants such as *l* or *m* may follow themselves (as in *ll* or *mm*) is stronger than its prediction that never doubled consonants such as *h* or *x* may follow themselves. In addition, after 10,000 epochs, the activation of a doubled consonant is higher in medial position than in initial position for both kinds of consonants. Furthermore, the activation of double letters increased in medial position and, conversely, tended to decrease in the initial position for both frequently and never doubled consonants. It is important to note that the higher activation of double letters in medial position (vs. in initial position) cannot be solely explained by the fact that the corresponding single letters are more frequent in medial than in initial position. Indeed, despite the fact that certain consonants are more frequent in initial position than in medial position (e.g., *f* occurs as the initial letter of 40 words and in the middle portion of 30 words in our database) whereas other consonants are more frequent in medial position than in initial position (e.g., *m* occurs as the initial letter of 63 words and in the medial position of 124 words in our database), the prediction responses of the network with respect to doubling were similar for these two consonants.

To simulate children's performance in our task, however, we did not compare raw output activations but instead preferred to compare relative activation levels. Indeed, while comparing raw activation levels would be adequate if the task were one of production, our task was a forced-choice test. Thus, children—who would

probably never spontaneously produce doublets such as *xx* (Treiman, 1993)—were forced to choose between nonwords including, for instance, *xx* in the initial position or *xx* in the medial position. Likewise, in our simulations, the overall activation of this doublet is very low (irrespective of its position) because the SRN has never encountered doublets such as *xx*. Thus, computing a ratio (activation in medial position / activation in initial + medial positions) more closely simulates the task proposed to children, in which the only relevant factor is the *relative* activation of the two items.

To assess the network's differential predictions concerning doubled consonants versus doubled vowels, we computed an index meant to reflect the constraints human participants were requested to take into account when faced with the comparison task. Let us call *CCf* the average activation of output units corresponding to the network's prediction of the second instance of frequently doubled consonants and *VV* the average activation of output units corresponding to the network's prediction of the second instance of doubled vowels. The ratio $CCf / (CCf + VV)$ then provides an index of the extent to which the network predicts the second instance of frequently doubled consonants while controlling for its overall tendency to predict doubling. Early during training, this ratio was around .30, indicating a lower activation for frequently doubled consonants than for double vowels (see Figure 6). This results from the network's initial sensitivity to the higher frequency of vowels in the selected database. After more extensive training, however, the ratio increased up to .95, suggesting that the network had become more and more sensitive to the fact that vowels are never doubled in French. Thus, as with human participants, the network became increasingly sensitive to the frequency of doubled letters, and not merely to the frequency of single letters (for if this were the case, the activation of double vowels should not decrease with practice). This result was not just a by-product of the net-

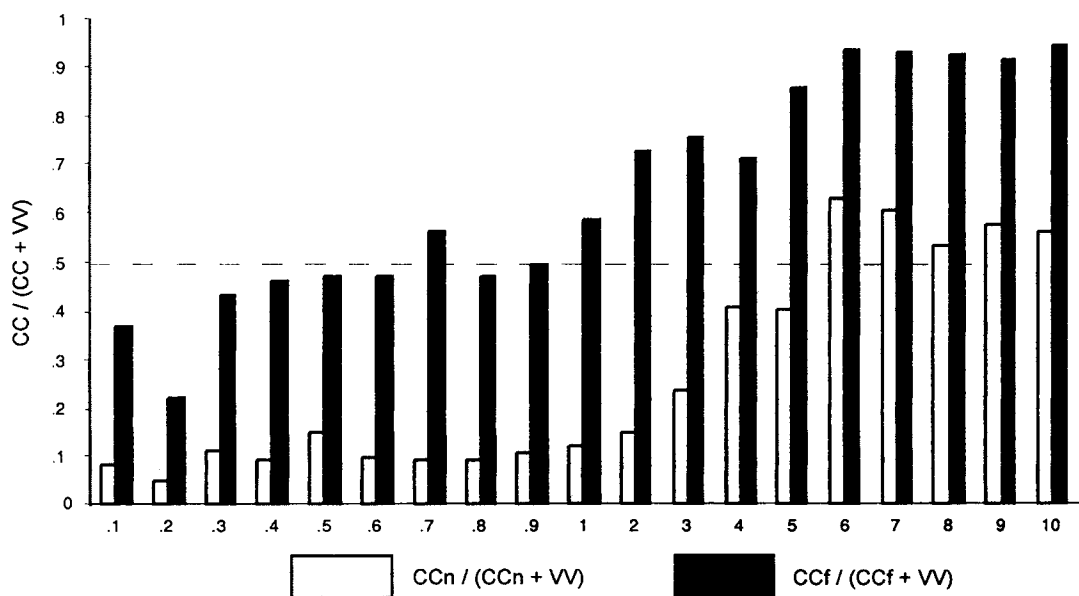


Figure 6. Network prediction that never doubled consonants (CCn) or frequently doubled consonants (CCf) can be doubled, relative to overall network tendency to predict doubling (see text for details). CC = double consonant; VV = double vowel.

work's sensitivity to the fact that certain consonants are frequently doubled, because the ratio also increased when the level of activation in the output layer for a doublet of never doubled consonants was considered. Indeed, in another analysis, we computed the ratio $CCn/(CCn + VV)$, in which CCn indicates the level of activation in the output layer for doublets of never doubled consonants. This ratio increased with training from .08 to .56. Thus, the network became sensitive to the consonant-vowel distinction, as shown by the fact that the network discriminated between the doubling of consonants and vowels, even though neither the specific consonants involved in the comparison nor the vowels are ever doubled in French. Importantly, as with children, there was a close parallelism between performance elicited by VV and CCf items on the one hand and by VV and CCn items on the other hand. This parallelism resulted in a strong correlation across epochs between the two measures ($r = .92$).

Note that some discrepancies were observed between participants' and SRN performance. Children's selection of nonwords including CCn was, on average, higher than the simulated ratio, even after 10,000 epochs on the SRN (76% vs. 56%). This difference presumably reflects the impact of phonological information on participants' performance—information with which the network was not provided.

The most striking form of abstract knowledge revealed in our studies pertained to the legal position at which doubling can occur for letters that are nevertheless never doubled in French. It is important to assess whether a standard SRN is able to reproduce

children's performance in this respect because this dimension was our study's correlate of the laboratory letter-transfer paradigm. To examine whether the SRN was able to reproduce this finding, we conducted separate analyses for each of the four cells resulting from the crossing of two dichotomous criteria: the frequency of double consonants and their position. Let us call m the level of activation of the output units corresponding to a double consonant occurring in the medial position and i the level of activation of the output units corresponding to a double consonant occurring in initial position. We computed the ratio $m/(m + i)$. A ratio of .5 indicates that the network predicts doubling irrespective of the location of the letter within the nonword. The extent to which the ratio departs from .5 indicates how well the network uses positional information. A ratio of 1 would correspond to the normatively correct response, a doubling being predicted only in the medial position. For consonants frequently doubled in French, this ratio started around .5 and progressively increased up to .90 after 10,000 epochs of training, as shown in Figure 7. This indicates that the positional information that determines the extent to which doubling is legal was correctly extracted for frequently doubled letters. Crucially, this knowledge transferred to letters that are never doubled in French: There was a close parallelism between the performance elicited by frequently and never doubled consonants. In both cases, performance started to differ from chance level after about 1,000 epochs of training. This parallelism induced a strong correlation across epochs between the two measures ($r = .90$). However, as for children's performance, the effect

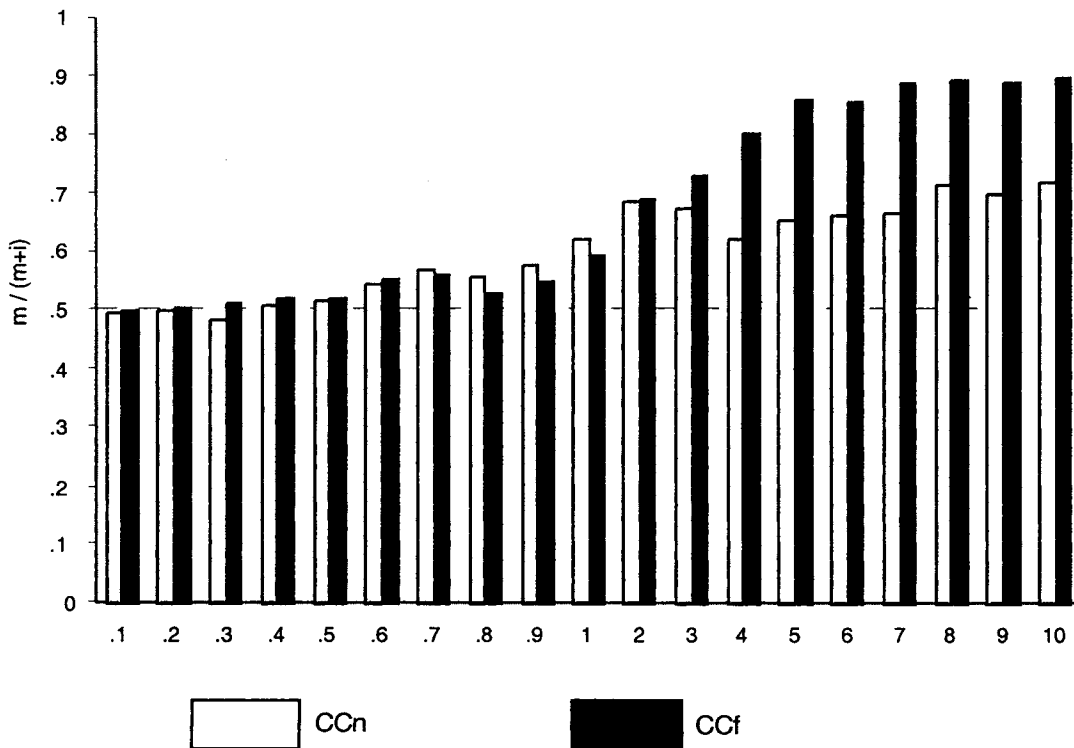


Figure 7. Network prediction that frequently doubled consonants (CCf) or never doubled consonants (CCn) can be doubled in the medial position, relative to overall network tendency to predict consonant doubling independently of position (see text for details). m = activation in medial position; i = activation in initial position.

for unseen material was somewhat reduced in amplitude, with a ratio culminating to .70 at the end of training (Figure 7).

Of course, these findings do not entail rejecting the notion of abstraction, but they cast light on the possible nature of the process. The fact that the form of abstraction revealed in our studies can be successfully simulated by a standard connectionist network that, as most other connectionist models, does not incorporate any mechanism for rule-based abstraction strongly suggests that such mechanisms are simply unnecessary to account for the data. Instead, both the empirical results and our simulation results appear more consistent with a characterization of abstraction as a graded, dynamic process (Cleeremans, 1997; see also Shanks et al., 1997) through which both sensitivity to higher-level features of the material (e.g., a distinction between consonants and vowels) and generalization to novel instances (e.g., the network's performance on never doubled consonants) progressively emerge as a result of exposure to the material. To demonstrate this, it is necessary that we provide a short description of the way learning proceeds in the SRN.

When training starts, the network can be sensitive only to the information provided by the current letter of the word that it is presented with because its internal representations (which also act as context for the next trial) are completely unstructured and merely reflect the initially random connectivity between the network's units. The network will quickly become sensitive to the overall frequency of the various letters, however, and will tend to activate more strongly those output units that represent these frequent letters. This is what explains the network's early sensitivity to vowels in general: Given that the network cannot, at this point, represent and use context information, predicting the most frequent letters irrespective of the context in which they occur is the best the network can do to reduce the error associated with its responses. At this point, the association between each letter and the entire set of its possible successors has become associated with a unique internal representation. This consistency is what enables the network to subsequently become gradually increasingly sensitive to the sequential constraints present in the stimulus material. Indeed, when copied over the context units on the next time step, the stable internal representation associated with each letter can now be used by the network to base its predictions not only on the current letter but also on the context in which this letter occurs. At this point, the network is thus sensitive to first-order sequential constraints. This sensitivity enables it, for instance, to predict that some consonants can be doubled whereas others cannot.

Later in training, the network will further refine its prediction responses by integrating more and more temporal context in the internal representations associated with the different letters. For instance, the network becomes sensitive to positional information because the context that precedes a consonant doublet that occurs in the medial position (i.e., a set of specific letters) is different from the context that precedes the same doublet when it occurs in the initial position (i.e., nothing). Thus, given identical inputs, the network, because it has become sensitive to the context in which these inputs occur, is now able to use this context information in order to produce different prediction responses. At this point, the network is thus sensitive to second and higher order sequential constraints, because its predictions are now based on the contingencies between subsequences of two or more letters and their possible successors. Throughout training, the network's responses

reflect almost exactly the conditional probabilities of occurrence of each letter given an increasingly large temporal context set by previous sequences of letters.

Of importance, the rich set of internal representations that the network develops in response to the prediction task reflects high-level features of the stimulus material in the sense that they encode the functional similarity relationships that exist between different letters. For instance, the fact that consonants in general share the property that they collectively tend to be more often preceded and followed by vowels rather than by other consonants is a high-level feature of the material that the network becomes sensitive to because it is useful to support performance in the prediction task.

Is Generalization to Oral Language Possible?

Our results concerning children's early sensitivity to orthographic regularities echo findings obtained on the acquisition of natural language. A growing number of studies show that infants become sensitive to the prosodic and phonological regularities of their native language quite early during development (see review in Brent & Cartwright, 1996; Christiansen, Allen, & Seidenberg, 1998; Jusczyk, 1997; McDonald, 1997). For instance, Jusczyk, Cutler, and Redanz (1993) reported evidence that 9-month-old English-learning infants preferred spoken words with a strong-weak pattern—a stress pattern typical of English—over words with a weak-strong pattern. Such formal similarities between results obtained with written and spoken languages prompt us to examine whether the conclusions drawn from our studies, particularly those regarding the abstraction issue, can be generalized to oral language.

It could be argued that both forms of language differ according to whether the basic units of analysis are given in the input or must be discovered by the learner. Indeed, in written language, the letters are easily identifiable and are moreover explicitly taught as isolated patterns. This makes it possible to compute frequency information on these units. By contrast, oral language is characterized by a continuous speech flow that requires relevant units to be identified before any frequency counting mechanism can start operating. This identification would require specific, hard-wired processes that lie out of the scope of any frequency-based account.

Although the difference is worth considering, several recent studies have shown that infants are able to learn the words of an artificial language presented orally. For instance, Saffran, Aslin, and Newport (1996) exposed 8-month-old infants to an artificial language composed of four trisyllabic words that were read by a speech synthesizer in immediate succession without pauses or any other prosodic cues. The infants were then tested with the familiarization preference procedure of Jusczyk and Aslin (1995), in which they controlled the exposure duration of the stimuli by their visual fixation on a light. The infants showed longer fixation (and hence listening) times for nonwords than for words, hence suggesting that they had discovered the relevant units of the language. The hypothesis that this ability is subtended by general learning mechanisms was further supported by studies showing that it extends to arbitrary sequences of nonlinguistic stimuli (Saffran, Johnson, Aslin, & Newport, 1999). Finally, Perruchet and Vinter (1998b) demonstrated that the discovery of wordlike units from continuous input such as evidenced in the Saffran and colleagues studies can be accounted for by simple and ubiquitous associative

learning mechanisms (see also Bates & Elman, 1996; Christiansen & Curtin, 1999; Redington & Chater, 1998).

This does not entail that generalization of the conclusions drawn from written language to oral language is fully warranted. One must indeed consider that the acquisition of oral language is phylogenetically more ancient and ontogenetically more precocious than the acquisition of written language, which suggests that oral language relies on specific, innate mechanisms different from those investigated here. In support of such a view, it has been argued that some results in the acquisition of artificial oral language cannot be accounted for by statistical approaches. For instance, Marcus, Vijayan, Bandi Rao, and Vishton (1999) recently showed that 7-month-old infants were able to discriminate between pattern structures such as ABA and ABB, despite the fact that both A and B were instantiated by different syllables during the study phase and during the test phase of the experiment. According to the authors (see also Pinker, 1999), such sensitivity cannot be simulated by connectionist models such as SRNs. However, this claim has been subsequently challenged (Christiansen & Curtin, 1999; McClelland & Plaut, 1999; Negishi, 1999; Seidenberg & Elman, 1999). For instance, McClelland and Plaut (1999) have shown that all the ABB items in Marcus et al.'s (1999) experiment can be recoded as "different-same," whereas none of the other items could be recoded in this manner—an analysis that suggests that infants' performance could in fact be based on an elementary sensitivity to repetition patterns in the stimulus material rather than on the acquisition of abstract rules (but see Marcus, 1999).⁴ Marcus's findings, in this light, would thus illustrate nothing more than the kind of transfer that has been shown to exist in implicit learning studies and should thus be interpreted in the same manner: Transfer can be accounted for by statistical learning mechanisms provided that their input involves some elementary forms of relational coding.

Generalizing our conclusion to oral language also presupposes that transfer decrement, which is consistently observed in the present study and in the implicit learning literature, applies to oral language as well. Against this assumption, Prasada and Pinker (1993) showed that the generalizability of the use of the regular suffix *-ed* to indicate past tense does not appear to depend on similarity to existing regular verbs (whereas generalization from existing irregular verbs to novel ones depends on similarity). This lack of similarity effect can thus be interpreted as indicating that regular inflectional patterns were rule governed. Note, however, that Prasada and Pinker's study involved morphological aspects of the language, whereas experiments such as those of Marcus et al. (1999) or the ones reported in the present article do not. Our conclusions with respect to the learning of orthographic regularities may thus be applicable only to certain aspects of oral language. Further studies are needed to investigate the similarities and divergences between the acquisition of oral and written language. As a case in point, studies should be devoted to exploring whether transfer decrement can be found for the morphological dimensions of written language.

Conclusion

One of the most fundamental issues in cognitive psychology concerns whether novel items are dealt with by applying rules that have been previously abstracted or by relying on their similarity to

previously processed items. This issue is central in domains such as memory (e.g., Hintzman, 1986), categorization (e.g., Barsalou, 1990), language acquisition (e.g., Pinker, 1994; Rumelhart & McClelland, 1986), and implicit learning (e.g., Cleeremans et al., 1998). Debates about this issue have been particularly vivid in the latter domain, in which most of the controversy has been focused on the interpretation of the results obtained in so-called letter-transfer experiments that involve training and test stimuli instantiated over different surface features but that share the same underlying abstract structure. Although the very existence of transfer in these conditions has been initially considered problematic for statistical approaches, it has become recently obvious that these approaches can accommodate such results provided that statistical learning mechanisms apply to elementary relational coding. Rule-based theories account for transfer in natural ways, but these theories have difficulty explaining the transfer decrement phenomenon. One argument is that transfer decrement may be due to an insufficient practice in acquiring the relevant rules. Because this hypothesis was hard to evaluate in laboratory context, because of obvious limits with regard to the duration of experimental sessions, we have used the letter-transfer paradigm in a real-world situation where learning extends over a very large time scale.

Our study confirmed the generalizability of laboratory research to natural contexts and, more specifically, the existence of transfer to unseen material in such natural settings. An important aspect of our study is the observation of very high transfer performance on novel material—a result that has consistently eluded standard laboratory research and that further suggests that the strategy of transposing laboratory paradigms into natural settings whenever possible is a fruitful one. Our main result, however, was that transfer decrement was shown to be an enduring phenomenon even after massive practice (at least five years of exposure to print). Performance on familiar items consistently paralleled performance on unfamiliar items, and parallelism over grade level extended to between-subjects differences. A transfer effect and a pervasive transfer decrement were also obtained in simulation studies using an SRN. Overall, these results suggest that elementary associative learning mechanisms are sufficient to account for the observed pattern of performance. By contrast, the fact that transfer decrement persists over extensive training without any trend toward fading is difficult to reconcile with the idea that participants acquired

⁴ In his reply to McClelland and Plaut, Marcus (1999) argued that their interpretation does not eliminate the need for rules but only displaces it. According to Marcus, there must be "some external device to compute whether any two items are the same," (p. 169) and to implement a rule "that says, for all syllables *x*, *y*, if *x* equals *y* execute condition A, otherwise execute condition B" (p. 169). This argument, however, strikes us as circular (see also Gellatly, 1992). It is clear that the involvement of rule-based knowledge can be invoked whenever any regularity is observed in behavior, but such an interpretation quickly dissolves into absurdity. Should a hungry animal looking for food be assumed to execute a rule such as "when hungry, execute food search"? Likewise, should one assume that sunflowers rely on a rule to determine how to orient their stalks so as to align with the sun's position (Searle, 1992). Infants' sensitivity to the same-different relationship, which is obvious in any habituation experiment, has never been conceptualized as a rule-governed behavior, and need not be.

increasingly abstract, rule-based knowledge about the regularities contained in the material over training.

References

- Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 899–912.
- Anderson, J. R. (1993). *Rules of mind*. Hillsdale, NJ: Erlbaum.
- Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 3, pp. 61–88). Hillsdale, NJ: Erlbaum.
- Bates, E., & Elman, J. (1996, December 13). Learning rediscovered. *Science*, 274, 1849–1850.
- Berry, D. C. (1997). *How implicit is implicit learning?* Oxford, England: Oxford University Press.
- Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hillsdale, NJ: Erlbaum.
- 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: Erlbaum.
- Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1969) and Mathews et al. (1989). *Journal of Experimental Psychology: General*, 120, 316–323.
- Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89, 631–644.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- 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.
- Christiansen, M. H., & Curtin, S. L. (1999). Transfer of learning: Rule acquisition or statistical learning? *Trends in Cognitive Sciences*, 3, 289–290.
- Cleeremans, A. (1993). *Mechanisms of implicit learning: A connectionist model of sequence processing*. Cambridge, MA: MIT Press.
- Cleeremans, A. (1994). The representation of structure in explicit prediction tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 783–809). Cambridge, MA: MIT Press.
- Cleeremans, A. (1997). Principles for implicit learning. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 195–234). Oxford, England: Oxford University Press.
- Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, 2, 406–416.
- Cleeremans, A., Servan-Schreiber, D., & McClelland, J. L. (1989). Finite state automata and simple recurrent networks. *Neural Computation*, 1, 372–381.
- Content, A., Mousty, P., & Radeau, M. (1990). Brulex. Une base de données lexicales informatisée pour le Français écrit et parlé [Brulex: A lexical database for written and spoken French]. *Année Psychologique*, 90, 551–566.
- Dienes, Z., & Altmann, G. T. M. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 107–123). Oxford, England: Oxford University Press.
- Dienes, Z., Altmann, G. T. M., & Gao, S. J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science*, 23, 53–82.
- Dulany, D. E., Carlson, A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, 113, 541–555.
- Ehri, L. C. (1986). Sources of difficulty in learning to spell and read. In M. L. Wolraich & D. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (Vol. 7, pp. 121–195). Greenwich, CT: JAI Press.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179–212.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301–330). London: Routledge & Kegan Paul.
- Gellatly, A. (1992). The misleading concept of cognitive competencies. *Theory & Psychology*, 2, 363–390.
- Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, 33, 154–207.
- Henderson, L. (1985). *Teaching spelling*. Boston: Houghton Mifflin.
- Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence*, 40, 185–234.
- Hintzman, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93, 411–428.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 28, 1–23.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675–687.
- Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 169–181.
- Kushner, M., Cleeremans, A., & Reber, A. (1991). Implicit detection of event interdependencies, and a PDP model of the process. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society* (pp. 215–220). Hillsdale, NJ: Erlbaum.
- Lewicki, P., Czyżewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 523–530.
- Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20, 24–37.
- Luce, D. R. (1986). *Response times*. New York: Oxford University Press.
- Manza, L., & Reber, A. S. (1997). Representing artificial grammar: Transfer across stimulus forms and modalities. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 73–106). Oxford, England: Oxford University Press.
- Marcus, G. F. (1998). Rethinking eliminative connectionism. *Cognitive Psychology*, 37(3), 243–282.
- Marcus, G. F. (1999). Connectionism: With or without rules? Response to J. L. McClelland and D. C. Plaut (1999). *Trends in Cognitive Sciences*, 3, 168–170.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999, January 1). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blancard-Fields, F., Cho, J.-R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1083–1100.
- McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences*, 3, 166–168.
- McDonald, J. L. (1997). Language acquisition: The acquisition of linguistic structure in normal and special populations. *Annual Review of Psychology*, 48, 215–241.

- McGeorge, P., & Burton, A. M. (1990). Semantic processing in an incidental learning task. *Quarterly Journal of Experimental Psychology*, 42A, 597–609.
- Meulemans, T., & Van Der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1007–1028.
- Michas, I., & Berry, D. C. (1994). Implicit and explicit processes in a second language learning task. *European Journal of Cognitive Psychology*, 6, 357–381.
- Nation, K. (1997). Children's sensitivity to rime unit frequency when spelling words and nonwords. *Reading and Writing: An Interdisciplinary Journal*, 9, 321–338.
- Negishi, M. (1999, April 16). A comment on G. F. Marcus, S. Vijayan, S. Bandi Rao, & P. M. Vishton's "Rule learning by seven-month-old infants." *Science*, 283, 77–80.
- Pacton, S., Fayol, M., & Perruchet, P. (in press). The acquisition of untaught orthographic regularities in French. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy*. Dordrecht, The Netherlands: Kluwer.
- Perruchet, P. (1994a). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 223–228.
- Perruchet, P. (1994b). Learning from complex rule-governed environments: On the proper functions of nonconscious and conscious processes. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 811–835). Cambridge, MA: MIT Press.
- 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, England: Oxford University Press.
- Perruchet, P., Gallego, J., & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology*, 22, 493–516.
- 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., & Pacteau, C. (1991). The implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General*, 120, 112–116.
- 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.
- Perruchet, P., & Vinter, A. (1998b). PARSER: A model for word segmentation. *Journal of Memory and Language*, 39, 246–263.
- Perruchet, P., & Vinter, A. (in press). The self-organizing consciousness: A framework for implicit learning. In R. A. French & A. Cleeremans (Eds.), *Implicit learning*. Hove, United Kingdom: Psychology Press.
- Pinker, S. (1994). *The language instinct*. New York: HarperCollins.
- Pinker, S. (1999, January 1). Out of the minds of babes. *Science*, 283, 40–41.
- Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, 81, 115–119.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General*, 125, 123–138.
- Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes*, 13, 129–191.
- Redington, M., & Chater, N. (in press). Knowledge representation and transfer in artificial grammar learning. In R. A. French & A. Cleeremans (Eds.), *Implicit learning*. Hove, United Kingdom: Psychology Press.
- Rumelhart, D. E., Durbin, R., Golden, R., & Chauvin, Y. (1995). Backpropagation: The basic theory. In Y. Chauvin & D. E. Rumelhart (Eds.), *Backpropagation: Theory, architectures, applications* (pp. 1–34). Hillsdale, NJ: Erlbaum.
- Rumelhart, D., & McClelland, J. (1986). On learning the past tense of English verbs: Implicit rules or parallel distributed processing? In McClelland, J. L., & Rumelhart, D. E. (Eds.), *Parallel distributed processing* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–209). New York: Academic Press.
- Searle, J. R. (1992). *The discovery of the mind*. Cambridge, MA: MIT Press.
- Seidenberg, M. S., & Elman, J. L. (1999, April 16). Do infants learn grammar with algebra or statistics? *Science*, 14, 433.
- Servan-Schreiber, D., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 592–608.
- Shanks, D. R., Johnstone, T., & Staggs, L. (1997). Abstraction processes in artificial grammar learning. *The Quarterly Journal of Experimental Psychology*, 50A, 216–252.
- Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science*, 16, 1–40.
- Stadler, M. A., & Frensch, P. A. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.
- 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. W. A., & Dorken, M. D. (1997). Implicit learning: Indirect, not unconscious. *Psychonomic Bulletin and Review*, 4, 63–67.
- Whittlesea, B. W. A., & Wright, R. L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 181–200.
- Wright, R. L., & Burton, A. M. (1995). Implicit learning of an invariant: Just say no. *Quarterly Journal of Experimental Psychology*, 48A, 783–796.

(Appendixes follow)

Appendix A

Stimuli Used in Experiments

Experiment 1

Double Items

assyla–avvyła
essura–eccura
issote–iddote
ommera–ovvera
immatu–iccatu
ummyse–uddyse

illaro–ivvaro
illetu–iccetu
elloba–eddola
abossi–abovvi
urusse–urucce
ogessa–ogedda

ubamme–ubavve
otymma–otycca
avumme–avudde
etullo–etuvvo
oballi–obacci
abolle–abodde

Single Items

osela–ocela
amyra–avyra
yletu–yvetu
ubasi–ubaci
abome–above
ovile–ovive

isara–ivara
emutu–ecutu
olaro–ocaro
oryse–oryve
utuma–utuca
ebuli–ebuci

usate–udate
imose–idose
alebi–adebi
agusa–aguda
evyme–evyde
orile–oride

Filler Items

ammiro–azziro
ovvubi–oxxubi
ilazo–ilamo
ulizo–ulivo

elukke–elusse
utakki–utacci
oseru–okeru
ecevo–ekevo

illoma–ikkoma
elizzo–eliddo
eboxu–ebolu
adile–axile

Experiment 2

Position CCf Items

nullor–nnulor
tiffol–tifol
gollir–golirr
laffum–lafumm

fommir–ffomir
lurret–lluret
rammit–ramitt
lirran–lirann

rannes–rranes
mattir–mmatir
minnal–minall
muttef–muteff

Position CCn Items

xihhel–xxihel
koxxir–kkoxir
lakkex–lakexx
lawwix–lawixx

hojjir–hhojir
jawnor–jjawnor
rixxow–rixoww
tukkoj–tukojj

jukker–jjuker
kijjer–kkijer
mohhek–mohekk
rexxih–rexihh

CCf/CCn Items

billot–bihhot
ommile–oxxile
befful–bekkul
irramé–ijjamé

ullate–ujjate
tannis–tahhis
effiré–exxiré
attoge–akkoge

tummet–tukket
onnavé–ojjave
merral–mehhal
bottit–boxxit

VV/CCf Items

tillos–tiilos
ullars–ulaars
bummer–buumer
ommite–omiite

bannil–baanil
onnare–onaare
reffur–reefur
effire–efiire

cerral–ceeral
irrate–iraate
gattul–gatuul
gottir–gotiir

VV/CCn Items

lihher–liiher
ahhire–ahiire
ohhuré–ohuuré
tajiil–tajiil

rujjer–ruujer
majjir–maajir
rikkar–rikaar
tukkes–tuukes

tekkir–tekiir
raxes–raaxes
loxxar–loxaar
mixxot–miixot

Experiment 3a

Cf/Cr Items

unotir–ukotir
golutir–goluzir
lipéral–lihéral
tiwaler–tidaler

akiler–ariler
ritarer–rixarer
galiter–gahiter
boqalir–bodalar

bézilat–bérilat
uroxir–uropir
ruloté–ruvoté
rinoter–riqoter

Appendix A (continued)

Experiment 3a (continued)

CCf/CCn Items

tuballir–tubakkir
darosser–darozzer
lotannir–lotahhir
bilatter–bilajjer

rébammé–rébakké
garossi–garoxxi
gimarrer–gimahher
gorettir–gorevvir

lirommer–lirozzer
mébanner–mébaxxer
lomerri–lomejji
jubelli–jubevvi

Cf/CCf Items

immoter–iloter
tibarrer–tibamer
jubitter–jubirer
garollue–garotue

garonné–garolé
arrolir–anolir
appagir–aratir
borillé–boripé

géranmir–géramir
attagir–anagir
mérappir–mératir
ammagir–apagir

Cr/CCn Items

izaler–ikkaler
bahité–bazzité
rajulé–rahhulé
tukalir–tujjalir

raxolit–rakkolit
ohiler–oxxiler
rivoler–rihholer
rovvalé–rokalé

motixer–motizzer
tijoler–tixxoler
ravélir–rajjélir
girazer–giravver

Cf/CCn Items

tilagé–tikkagé
jubamer–jubaxxer
bariler–bajjiler
rutabé–ruhhabé

rolubé–rozzubé
bonalot–boxxalot
mudaré–mudavvé
ripagé–ripahhé

rumabir–ruzzabir
ranolir–rajjolir
gabotir–gabovvir
turiper–turikker

Cr/CCf Items

logappe–logake
litarré–litazé
tagonner–tagoher
sarollit–sarowit

attamir–akamir
arrolir–axolir
ammoter–ahoter
allébir–ajélir

azalir–attalir
téganmir–tégaxir
férommir–férowir
appamir–ajamir

Experiment 3b

Position CCf Items

gira__er__iraver (m–ll)
ibaro__e__ibarue (l–mm)
biro__er__iroger (t–nn)
__uméla–a__alir (n–tt)
a__iler__aviler (p–rr)
a__alit__atalit (r–pp)
gilo__e__iloco (p–tt)
gira__e__irale (t–pp)
a__utie__amutie (r–ll)
a__aver__amaver (l–rr)
icéra__e__icérope (m–nn)
a__elir__atelir (n–mm)

ro__ema–rovema__ (t–ll)
a__ivor–anivo__ (l–tt)
a__emir–ravemi__ (n–ll)
péfo__at–péfova__ (l–nn)
ogila__e–molige__ (l–rr)
acimo__e–cimote__ (r–ll)
a__iger–alogi__ (n–tt)
a__olir–avoli__ (t–nn)
umila__e–cumila__ (r–tt)
i__éva–iméva__ (t–rr)
liva__e–livati__ (r–nn)
luba__ie–lubova__ (n–rr)

Position CCn Items

adoli__a__oli (j–ww)
metali__e__alir (w–jj)
i__ober–ilobe__ (w–hh)
o__ager–olage__ (h–ww)
u__alir–umali__ (h–jj)
a__iler–avile__ (j–hh)

ro__ébit__orébit (z–kk)
__ivabot–li__abot (k–zz)
__eliro–le__iro (z–xx)
__uliver–mu__iver (x–zz)
__umolir–ru__olir (x–kk)
__olarie–lo__arie (k–xx)

Note. CCf = frequent double consonant; CCn = never double consonant; VV = double vowel; Cf = frequent single consonant; Cr = rare single consonant.

(Appendixes continue)

Appendix B

Simulation Details

Network Architecture

The network used in the simulations described in this article was a standard simple recurrent network (Elman, 1990) with 15 hidden units and local representations on both the input and the output pools (see below). The network incorporated 1,260 connections among which 1,245 were trainable and 15 were used to copy the hidden unit activations onto the context units on each trial. The network was trained using the back-propagation algorithm and the cross-entropy error measure (Hinton, 1989). Unit activations ranged from 0 to 1. The network used the standard logistic activation function for its hidden units and the so-called softmax, or normalized exponential activation function (see Rumelhart, Durbin, Golden, & Chauvin, 1995) for its output units. In contrast to the standard sigmoid activation function, the softmax function (see Luce, 1986) normalizes the output vector to a sum of 1.0, which makes it particularly appropriate in cases where the target vector involves a single activated element. In such cases, the output of the network can then be interpreted directly as a probability distribution.

Stimulus Representation

Each n -letter word was coded by a series of $n + 1$ pairs of 34-bit vectors presented successively to the network. Within each pair, the first vector represented the input pattern (one of the word's letters) and the second vector represented the target (the next letter of the word). Each vector consisted of a single activated element corresponding to one of 34 possibilities, defined as follows: The 26 letters of the alphabet, 6 accentuated letters (à or â, ç, é or ê or è, ô, î, ù or ú), and two symbols ("[" and "]") to represent the beginning and end of each word. Thus for instance, the word *de* would be coded as the following three successive pairs of input-output patterns: [-d, d-e, and e-].

Training Procedure

On each epoch of training, a number of words were randomly selected from the 1,000-word training set for presentation to the network. Each word was associated with a probability that reflected its frequency of occurrence in the Brulex (Content et al., 1990) corpus. Selection proceeded as follows. First, for each word, a pseudo-random number between 0 and 1 was generated and compared with the word's probability of occurrence as specified in the training set. Second, the word was selected for presentation during the current epoch if its probability was smaller than or equal to the pseudo-random number. This procedure resulted in the selection of approximately 10 words per epoch. The selected words were then presented to the network in random order. For each word, the network was presented with the corresponding sequence of input-output patterns. The weights on its connections were modified using the back-propagation algorithm after each such presentation. The network's context units were reset to .5 on presentation of the "begin" symbol of each word.

Simulation Procedure and Parameters

Ten replications of the simulation were conducted and their results averaged. For each, training was initiated with a different set of random weights selected in the $-.5-.5$ range. All simulations were conducted using a learning rate of .1. Neither momentum nor weight decay were used.

Received June 10, 1998

Revision received December 6, 1999

Accepted June 13, 2000 ■