British Journal of Educational Psychology (2004), 74, 533–550

www.bps.org.uk

Orthographic revision: A developmental study of how revisers check verbal agreements in written texts

Pierre Largy[1]*, Alexandra Dédéyan[1] and Michel Hupet[2]
[1] PSY.CO (E.A. 1780), Université de Rouen, France
[2] Université Catholique de Louvain, Belgium

Background. Writing is a complex activity involving various cognitive processes in the planning, the transcription and the revision of written texts. The present study focused on the revision of written texts within a developmental approach.

Aims. The study aimed to examine whether children and adults use different procedures to detect and revise erroneous grammatical agreements. It was predicted that children would use a slow algorithmic procedure while adults would use a fast automatized procedure.

Sample. One hundred and twenty participants from 5th grade to undergraduate levels (24 per level) participated in the study.

Method. The participants were asked to decide as quickly as possible whether a visually presented sentence had any agreement error. The French experimental sentences were of the type 'The N1 of the N2 + Verb', in which N2 was either a plausible subject of the following verb (e.g., The guard of the prisoners watches) or an implausible subject (e.g., The guard of the safes watches). Correctness and latency of the responses were recorded.

Results. The main results showed that only the younger participants were affected by the subject-role plausibility of N2, and that there was no difference in response latency between their correct and incorrect responses. These observations support the hypothesis that the younger participants systematically apply a time-consuming algorithmic procedure to verify the agreement; since one step of this procedure consists in searching for the subject of the verb, these participants were frequently misled by the subject-role plausibility of N2. On the contrary, the older participants were not affected by the plausibility of N2, but were frequently misled by erroneous agreements between N2 and the verb. These observations support the view that these older participants use a fast decision strategy based on the co-occurrence of formal

* *Correspondence should be addressed to Pierre Largy, Université de Toulouse le Mirail, U.F.R. de Psychologie, F – 31058 Toulouse cedex 9, France (e-mail: largy@univ-tlse2.fr).*

indices. Their correct answers, however, were slower than their incorrect ones; this suggests that they also sometimes use a time-consuming controlled procedure.

Conclusion. The study shows that along with the acquisition of writing expertise, the revising activity itself is progressively facilitated and gradually automatized by substituting a fast direct decision strategy for a slow and laborious use of revision rules.

Writing is generally viewed as a complex and resource-consuming activity involving various cognitive processes from conceptual planning to the graphic transcription of a written text. A consequence of this complexity is that a writer's attention is regularly shared between the different aspects of the production and in such conditions writers make mistakes of different types (Hayes, 1996; Largy, Fayol, & Lemaire, 1996). Revising therefore constitutes an activity which is necessary not only for children learning to write, but also for the adults who have a daily practice of writing. It is thus no surprise that revising is present in every model of text production (Alamargot & Chanquoy, 2001).

As such, revising constitutes a multi-faceted activity that is particularly difficult to study due to the numerous possible interferences of many types of information: semantic, syntactic, phonological, orthographic, etc. It is however generally considered that revisions relative to surface characteristics (punctuation, spelling, grammar) are to be distinguished from revisions relative to the meaning or the coherence of the text. It is also frequently assumed that if writers revise the surface more than the meaning of the text, it is mainly because the surface revision is easier to make and less resource demanding, particularly for inexperienced writers (see Hacker, Plumb, Butterfield, Quathamer, & Heineken, 1994; Yagelski, 1995).

In some languages, however, even the revising of surface forms can be resource demanding. In written French for example, as in many other languages, the grammatical rule of subject-verb agreement is very simple: a singular subject requires a singular verb inflection whereas a plural subject requires a plural verb inflection. The difficulty in written French arises from the fact that for most nouns and verbs and for the third person pronouns, the morphology of plurality is silent: the particular written marks of plurality (-**s** for nouns, adjectives and third person pronouns, -**nt** for the third person of the plural of the verbs) simply have no corresponding pronunciation, and therefore are not easy to learn. For example, 'Il chante' (He sings) and 'Ils chantent' (They sing) are written differently but are pronounced in exactly the same way. This explains why persistent difficulties throughout primary education are associated with the plural number morphology (Fayol & Largy, 1992): the French-speaking children have to learn many written marks and their grammatical function without referring to oral language marking. Analogous difficulties exist for English-speaking children; erroneous homophonic substitutions for example ('week' instead of 'weak', 'to' instead of 'two', or 'seen' instead of 'scene'), due to the fact that a particular phoneme can be represented by several graphemes, clearly indicate that phonological information influences graphic patterns (see Ellis, 1979; Hotopf, 1983). Similarly, Beers and Beers (1992) showed how the plural mark /-s/ for English nouns varies in the written productions of young children as a function of its oral pronunciation, i.e., /s/, /z/ or /iz/.

It remains to be established how a young inexperienced writer deals with such a silent morphology of the plural in French, both in production and in revision. For production, Totereau, Thévenin, and Fayol (1997), and Fayol, Hupet, and Largy (1999) described the acquisition of this ability for the subject-verb agreement as a progressive

automatization of the agreement procedure, in reference to the ACT model of Anderson (1993). According to these authors, the automatization emerges as the final step of the following evolution:

At first, the children begin to discover the oppositions between linguistic marks of number and to associate them with oppositions between the notions of singularity and plurality; at this stage, they must encode the morphology of number for nouns and verbs under a declarative form. In a second phase, the children become able to interpret a particular number mark in the absence of the other term of the opposition; they progressively understand what the absence or presence of these marks means, but cannot use them in spontaneous written productions. In a third phase, the children are able to apply appropriately algorithms or condition-action rules of the following type: *If a word is a noun, and if this noun is plural, then it must be ended with an -s; or, if a word is a verb, and if the subject of that verb is plural, then it must be ended with an -nt*. Controlled and laborious use of such production rules is rules is characteristic of this third phase; the child must access declarative knowledge in memory, test the relevance of the conditions and of the actions and edit the current word. The final phase corresponds to an automatization of the agreement procedure. As children read and write more and more sentences, they encounter multiple instances of 'Article + Noun + Verb (+ Complement)', which is the most frequent sequence of words in written French (Dubois, 1965). The frequency and consistency of these multiple exposures lead the children to switch from controlled algorithmic computations to direct, fast and effortless memory retrieval (Logan, 1988, 1992). In other words, when older children and adults have to write a sequence of the type 'Plural article + Plural noun + Verb', the plural verbal inflection is likely to be automatically triggered by the plural noun in preverbal position.

Fayol *et al.* (1999) demonstrated this gradual automatization of the subject-verb number agreement by relating it to different types of errors in sentences of the type 'N1 of N2 + V' at different stages of learning. They showed how the participants' performance moved from constant errors in 7-year-old writers (who do not use plural marks at all) to adult-like attraction errors in 11-year-old writers through an intermediate phase characterized by an attention-demanding and easily disruptable computation of the agreement.

In the first phase, errors mainly consist of the absence of any plural inflections. Since the plural marks of nouns and verbs are undetectable by ear, the children do not use them when transcribing sentences. In the second phase, plural marks are available, and there are fewer agreement errors. However, this only occurs in unconstrained conditions of writing; under time constraints or in a dual task condition which are likely to disrupt the application of a controlled algorithm, agreement errors substantially increase. These errors mainly consist of the absence of plural marks, as in the first phase of development. In the third phase, children produce a new type of error resulting from a fast agreement procedure based on a 'proximity concordancy' principle (Francis, 1986). The children produce attraction errors of the same type as those produced by adults, i.e., errors that consist in making the verb agree with the immediately preceding noun or pronoun (as in 'La fille des voisins arrivent', lit. 'The daughter of the neighbours arrive', or 'Il les chantent', lit. 'He them sing').

According to Fayol *et al.* (1999), such errors in older children's and adults' written productions are due to a failure of control processes that are supposed to check the verbal agreement that has been automatically elicited. Yet, little is revealed about this

control procedure, except that it is likely to fail particularly in situations of temporary working memory overload.

To better understand such control operations, it is important to gain more information about how young or older writers actually revise the spelling of grammatical agreements. Largy (2001) recently made a first attempt to bring together the production of number marks and their revision by French-speaking children. He showed that the children's performance is better when asked to revise the spelling of grammatical agreements than when asked to produce them. Appropriate declarative knowledge about nominal and verbal agreements is available to these children, but they still use an attention-demanding procedure to produce such agreements. Therefore, their declarative knowledge may not be evident in spontaneous writing for the simple reason that the graphic execution itself may require a large amount of the available cognitive resources (Bourdin & Fayol, 1994). When they only have to revise written sentences, on the contrary, young children are able to use their declarative knowledge to simply add a missing inflection or suppress an inappropriate one.

More recently, Largy and Dédéyan (2002) suggested that children and adults also use different procedures to check grammatical agreements. An inexperienced young writer would apply analogous algorithms both in the production and in the revision of number marks. In the revision, the algorithm could be as follows: *If a word is a noun, if this word is plural, and if it does not end with an -s, this mark has to be added to the word; if a word is a verb, if the subject of this verb is plural, and if the verb does not end with an -nt, this mark has to be added to the verb*. Along with more practice of reading and writing, however, the revising itself is likely to be progressively automatized: the slow and laborious use of revision rules would be gradually replaced by an application of the proximity concordancy principle which gives rise to the above-mentioned attraction errors. An automatic use of this principle would lead to correct decisions for sentences like 'Noun-sing + Verb-**e**' (La fill**e** dans**e**, The girl dances) or 'Noun-plur + Verb-**nt**' (Les fille**s** danse**nt**, The girls dance), but would lead to erroneous decisions for sentences like 'Noun1-sing + Noun2-plur + Verb-**nt**' (La fill**e** des voisin**s** danse**nt**, The girl of the neighbours dance). This does not exclude the possibility for an adult to use a slow, attention-demanding algorithm of revision, but such a revision could only be occasional, limited for instance to infrequent or complex sentential configurations and to conditions in which the reviser allocates all the necessary attentional resources. Such a view is supported in English language by a recent study by Pearlmutter, Garnsey, and Bock (1999). They investigated adult comprehenders' sensitivity to agreement violations using a self-paced reading and eyetracking methodologies. They showed for instance that readers were sensitive to a locally distracting number-marked Noun2 giving rise to what these authors call a 'seeming agreement violation'; it is particularly worth noting that the readers had more difficulties (reflected in longer reading times and in higher regressive saccade probabilities) with a sentence like 'The key to the cabinets was rusty' than with 'The key to the cabinets were rusty'. Such processing disruptions, particularly for the former sentence type, indicate that the readers spent more time to check an agreement that seemed to violate the proximity concord principle. That is, they were less likely to check the agreement for the latter sentence type in which the proximity concord principle is respected.

It was the aim of the present study to confirm and extend Largy and Dédéyan's (2002) preliminary observations in order to get a deeper insight into the developmental pattern of the revising activity. Young and older participants were required to check

the spelling of grammatical agreement in sentences of the type 'N1 of N2 + Verb' in which N2 was either a plausible subject of the following verb (e.g., 'Le gardien des prisonniers surveille', lit. 'The guard of the prisoners watches') or an implausible subject (e.g., 'Le gardien des coffres surveille', lit. 'The guard of the safes watches'). The rationale for this manipulation is as follows.

If a reviser applies an algorithmic procedure to verify a verbal agreement, a step of which consists in identifying the subject of the verb, it can be hypothesized that his/her performance will be affected by a variable that is likely to interfere with this step. On the contrary, if a reviser uses a decision strategy simply based on the presence versus absence of co-occurring number marks, his/her performance should not be affected by such a variable. In this regard, the manipulation of the subject-role plausibility of the preverbal noun is particularly interesting, and is inspired by the following observations gained with adult writers in a study by Hupet, Fayol, and Schelstraete (1998). They showed that, in a single task condition (transcribe orally presented sentences), there were fewer agreement errors for sentences like 'Dans l'herbe chantent les grillons' (lit. In the grass sing the crickets), in which the noun N1 that immediately precedes the verb is not a semantically plausible subject for that verb, than for sentences like 'Avec le coq chantent les poules' (lit. With the cock sing the hens) in which N1 constitutes a semantically plausible subject of the verb. In both types of sentences, the preverbal position of N1, its initial position in the sentence and its definiteness make it a good candidate for the grammatical function of subject of the sentence (McDonald & Heilenman, 1991; MacWhinney, 1987). These characteristics constitute misleading cues capable of inducing erroneous role assignments and therefore agreement failures. However, for sentences of the first type, the subject-role implausibility of N1 is in conflict with the other cues, and this may be used by the system to ascertain the grammatical role of N1 before starting the agreement procedure; this is no problem in a single task condition in which the memory load is rather low and resources available both to detect and resolve the conflict between different cues. For sentences of the second type, on the contrary, since the subject-role plausibility of N1 is in accordance with the other cues, the execution of the assignment procedure is likely to result in unnoticed misassignments and therefore erroneous agreements, whatever the available resources may be.

In contrast, when the memory load was higher (as it was the case when the participants had to transcribe a sentence and simultaneously mentally add a series of digits), erroneous agreements increased for both types of sentences. This increase, however, was less important for sentences in which N1 was an implausible subject, but the participants' performance at the concurrent task (number of correct additions) was also lower for these sentences in comparison to the other ones. This last observation is particularly interesting for it clearly shows that controlling the agreement is a resource-consuming process.

The aim of the present study is to assess the extent to which the subject role plausibility may affect the revising performance, and the extent to which such an influence may vary with the participants' age. Participants from 10 to 22 years old were asked to decide as quickly as possible whether the correct form of the verb was used in sentences of the type 'N1 of N2 + Verb', and in which N2 was either a plausible or an implausible subject of the verb. Firstly, it is hypothesized that the younger participants are more likely to be affected by this variable than the adults. If the younger participants apply an algorithmic procedure to verify the agreement, the risk exists –particularly under time pressure– that they will incorrectly assign the grammatical subject function

to N2, which will result in incorrect responses both for sentences like 'N1-sing of N2-plur + V-sing', and sentences like 'N1-sing of N2-plur + V-plur'. On the other hand, if the younger participants apply an algorithmic procedure in every case, there should not be any latency difference between their correct and incorrect responses; their response latency indeed should essentially depend on the time needed for taking the various steps to achieve the whole agreement procedure, and should not depend on whether a particular step gives rise to a correct or an incorrect decision. In contrast, if the older participants do not apply such an algorithmic procedure, they should not be affected by the subject-role plausibility. If their response is simply based on the detection of co-occurring formal indices (i.e., on whether N2 and the Verb have similar number marks), they should be frequently misled by a number mismatch between N2 and the verb, but not more frequently for sentences with a plausible N2 than for sentences with an implausible N2. Considering however that older participants remain capable of applying a controlled algorithmic verification procedure, it can be further hypothesized that their correct responses (which are presumed to be based on a time-consuming algorithmic procedure) would be slower than their incorrect ones (based on a fast decision strategy).

Method

Participants

One hundred and twenty participants spanning five age categories (from 5th graders with a mean age of 10,1 to undergraduates with a mean age of 21,7) volunteered to participate in the study (see Table 1 for details). All participants were native speakers of French and none of them had ever repeated a year at school. The 5th to 11th graders were all described by their teachers both as reading normally and as presenting no particular writing problems. A preliminary phase aiming at familiarizing the participants with the procedure (which simply consisted in reading sentences on a screen) also identified participants whose reading times were much slower than the mean reading time for each age group and who could be excluded from the study.

Table 1. Distribution of participants according to their school level and age in years and months

School level	Youngest	Mean age	Oldest
5th grade	9.9	10.1	10.9
7th grade	12.1	12.7	13.3
9th grade	13.7	14.4	14.8
11th grade	16.2	16.7	17.4
Undergraduate	19.3	21.7	22.1

Materials

Sixty-four experimental sentences of the type 'N1 of N2 + Verb' were designed in such a way that N1 was always singular and N2 always plural, which is the sentential configuration that has been showed to elicit more errors than any other one (see Bock

& Miller, 1991; Fayol, Largy, & Lemaire, 1994). In half of these sentences, N2 could have been a semantically plausible subject for the verb; in the other half, N2 could not have been such a plausible subject. Furthermore, in half of the experimental sentences, the incorrect form of the verb was used (i.e., the verb erroneously ended with the plural verbal flexion -nt), while in the other half, the verb was correctly agreed.

The subject role plausibility of N1 and N2 had been controlled in a preliminary study in which 60 students (different from the participants) were asked to assess the plausibility of simple sentences of the type 'The noun + verb'. For 576 sentences of this type, the students were asked to assess on a 7-point scale the extent to which the noun was a highly plausible (=7) or implausible (=1) subject of the verb. For the 64 experimental sentences to be used in the present study, we selected as N1 the half of the nouns with a high plausibility score (>5.5); whereas the other half of the nouns with a high plausibility score (>5.5) was used in sentences in which N2 had to be a plausible subject of the verb. Nouns with a low plausibility score (<1.5) were used in sentences in which N2 had to be an implausible subject of the verb.

The experimental material was thus constituted as follows:

- sixteen sentences N1 + N2 plausible + V correct (e.g., 'Le directeur des secrétaires décide', lit. 'The director of the secretaries decides');

- sixteen sentences N1 + N2 plausible + V erroneous (e.g., 'La cliente des vendeuses bavardent', lit. 'The client of the salesgirls chat');

- sixteen sentences N1 + N2 implausible + V correct (e.g., 'La fille des rues danse', lit. 'The girl of the streets sings');

- and sixteen sentences N1 + N2 implausible + V erroneous (e.g., 'Le commandant des casernes crient', lit. 'The commandant of the barracks shout').

Filler sentences were also used to prevent the participants from identifying the specific characteristics of the experimental sentences; to this aim, the fillers were designed to allow variations in the grammatical structure of the sentences (N1 of N2 + V as for the experimental sentences, but also N + V, and N1 + V + N2), in the type of agreement error (incorrect verb agreement as in the experimental sentences but also incorrect noun agreement) as well as in the localisation of the errors (at the end or not of a sentence). A first group of 48 fillers consisted of sentences of the type 'N1 of N2 + V' in which N1 and N2 were respectively singular-singular, plural-plural and plural-singular; the verbal agreement was correct in half of these fillers, and erroneous in the other half. Another group of 64 fillers consisted of simple sentences of the type 'N + V'; in half of these sentences, there was an incorrect number agreement either on the noun N ('Le garçons chante' lit. 'The (singular) boys sings') or on the verb V (e.g., 'Les réveils sonne', lit. 'The (plural) clocks rings'). A last group of 22 fillers consisted in sentences of the type 'N1 + V + N2' aiming at presenting the participants with sentences where the incorrect agreement was not at the end of the sentence (e.g., 'Le meurtrier sèment la terreur', lit. 'The (singular) murderer sow the terror').

The words used both in experimental and filler sentences were all taken from the Novlex database (Lambert & Chesnet, 2001) that gives the inventory of the words most frequently used in primary school books and best-sellers for young children; this ensured that all the words chosen would be known by the children and that familiarity of lexical items was unlikely to affect performance.

Procedure

The participants were tested individually. Each participant was successively presented with two lists of 104 sentences (32 experimental ones, and 72 fillers), in counterbalanced order and with a short break between the two lists. For each sentence that appeared on the screen of a microcomputer, the participants were simply asked to decide as quickly as possible whether an agreement error occurred. In each age group, half of the participants were asked to press a left button whenever they thought the sentence had no agreement error, and a right button if they thought the sentence had an agreement error; for the other half, the buttons were reversed. Before starting the first list, each participant was presented with 10 sentences of various types with the aim of getting the participants accustomed to the task; it also established that every participant understood the instructions, and more particularly understood the notion of 'agreement'. Every 26 sentences (i.e., four times for each list) the participant was reminded that he had to answer as quickly as possible.

The whole experimental session lasted from 45 minutes for the youngest participants to 30 minutes for the older ones.

Results

To account for the participants' performance, two dependent variables were examined. The first one was the number of revising errors which was subdivided into two types: (a) the number of false alarms, i.e., the number of correct agreements that the participants took for wrong ones, and (b) the number of missed errors, i.e., the number of wrong agreements that the participants took for correct ones. The second dependent variable was the participants' response latency which was analysed according to whether the responses were correct or incorrect for sentences with or without an agreement error.

Analysis of errors

The mean percentages of errors (false alarms and missed errors), according to school levels and experimental conditions, are given in Table 2.

False alarms

The mean percentages of false alarms were analysed in a two-factor analysis of variance: 5 (School level: 5^{th}, 7^{th}, 9^{th}, 11^{th}, undergraduates) $\times$ 2 (subject role plausibility of the preverbal noun N2: plausible or implausible) with repeated measures on this second factor. The analysis first revealed a significant main effect of the school level: $F(4,184) = 2.9$, $p = .02$, $MSe = 2041$. Post hoc analyses (Duncan Test) showed that the mean percentage of false alarms first decreases from the 5^{th} (24.7%) to the 9^{th} (10.5%) grade level ($p = .01$), and then increases from 9^{th} grade level to the undergraduate one (24.8%) ($p = .02$). On the other hand, the effect of the subject role plausibility of N2 is marginally significant: $F(1,46) = 3.86$, $p = .55$, $MSe = 2646$. The participants are less likely to make a false alarm when N2 was an implausible subject of the verb (16.5%) than when it was a plausible one (23.1%). There was no significant interaction between the school level and the subject role plausibility of N2: $F(4,184) = 1.07$, *ns*.

Table 2. Mean percentages of false alarms and missed errors according to school level and plausibility of Noun2 (standard deviations in brackets)

False alarms	Noun2	
	Plausible	Implausible
School level		
5th grade	32.3 (29.1)	17.2 (19.1)
7th grade	28.2 (37.2)	19.0 (27.9)
9th grade	16.7 (24.4)	4.4 (09.2)
11th grade	15.1 (18.8)	15.6 (29.5)
Undergraduate	23.4 (27.2)	26.3 (31.3)
Missed errors	Noun2	
	Plausible	Implausible
School level		
5th grade	82.8 (24.1)	51.6 (34.1)
7th grade	47.6 (38.3)	18.0 (29.6)
9th grade	34.5 (38.1)	13.3 (21.8)
11th grade	15.1 (23.1)	17.7 (21.5)
Undergraduate	45.8 (42.1)	40.1 (42.7)

Missed errors

The percentages of missed errors were submitted to the same analysis of variance as the false alarms. The analysis first showed a significant main effect of the school level, $F(4,184) = 18.63$, $p < .001$, $MSe = 18713$; the mean percentage of missed errors first regularly decreases from the 5th grade level (67.2%) to the 11th grade one (16.4%), with intermediate values for the 7th and 9th grade levels (32.8% and 23.9% respectively), and then substantially increases at the undergraduate level (43%). The analysis also showed a significant main effect of the subject role plausibility of N2: $F(1,46) = 13.64$, $p < .001$, $MSe = 17398$; on the whole, participants were more likely to detect an erroneous agreement when N2 was an implausible subject of the verb (missed errors = 28.1%) than when it was a plausible one (missed errors = 45.2%). There is also a significant interaction between the School level and the Subject role plausibility, $F(4,184) = 2.66$, $p = .03$, $MSe = 2673$.

The effect of the subject role plausibility was further tested for each school level (Figure 1). As hypothesized, the effect of the plausibility of N2 is significant only for the younger participants; the observed effect was in the predicted direction: there were more missed errors when N2 was a plausible subject of the verb, 5th grade: $F(1,46) = 13.46$, $p < .001$, $MSe = 17719$; 7th grade: $F(1,46) = 8.99$, $p < .005$, $MSe = 10537$; 9th grade: $F(1,46) = 5.56$, $p = .02$, $MSe = 5359$. For the older participants (both 11th grade and undergraduate levels), on the contrary, the plausibility of N2 had no significant effect, respectively, $F(1,46) = .16$, *ns*; $F(1,46) = .22$, *ns*.

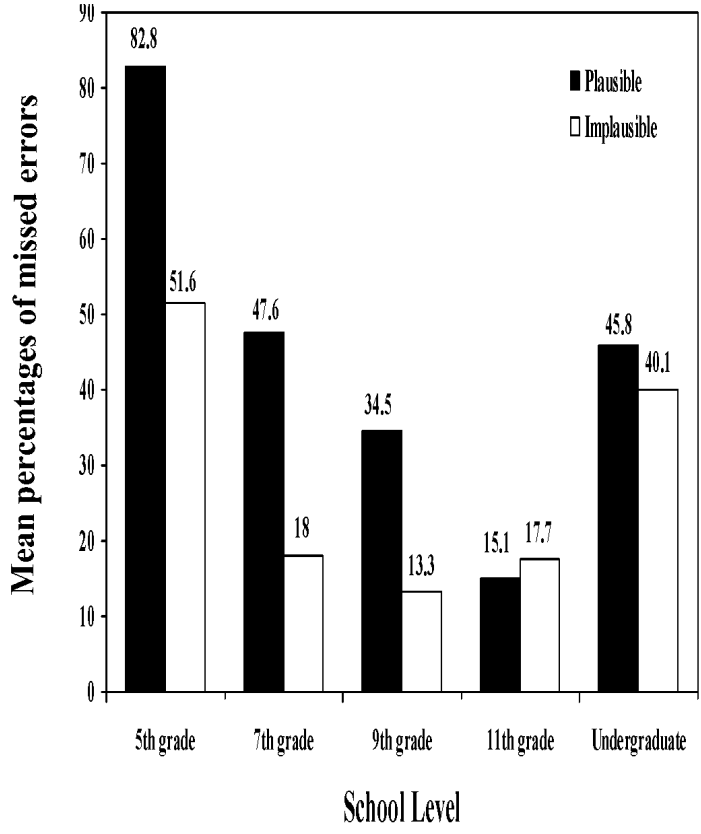

Figure 1. Mean percentages of missed errors in each age group as a function of the subject-role plausibility of N2

Analysis of response latency

The results of response time, measured in milliseconds, according to school levels, accuracy of answers and experimental conditions, are presented in Table 3.

Sentences without agreement error

For sentences in which the verb was correctly agreed, two types of responses will be considered: (a) the correct responses, when the participants correctly decide that the sentence has no agreement error, and (b) the incorrect responses, when the participants wrongly decide that the sentence has an agreement error.

The response latencies were analysed in a three-factor analysis of variance: 5 (School level: 5^{th}, 7^{th}, 9^{th}, 11^{th}, undergraduates) $\times$ 2 (subject role plausibility of the preverbal noun N2: plausible or implausible) $\times$ 2 (Response type: correct or incorrect), with repeated measures for the last two factors. The analysis first showed a significant main effect of the School level, $F(4,368) = 44.83$, $p < .001$, $MSe = 104105$; on the whole, the response latency regularly decreases from the 5^{th} grade level (5466 msec) to the undergraduate level (2799 msec), with intermediate values for the intermediate levels (3675 msec for the 7^{th} graders, 3145 msec for the 9^{th} graders and 3392 for the 11^{th} graders). The analysis also showed a significant main effect of the Response type, $F(1,92) = 5.6$, $p = .01$, $MSe = 126656$; on the whole, correct responses were slower

Table 3. Mean response time (in msec) for correct and incorrect responses as a function of sentence type, plausibility of N2 and of school level (standard deviations in brackets)

Exact sentences	Noun2			
	Plausible		Implausible	
Response type	Correct	Incorrect	Correct	Incorrect
School level				
5th grade	5415 (2365)	5122 (1691)	6169 (2407)	5160 (1875)
7th grade	3265 (0764)	3613 (1633)	4282 (1542)	3534 (1834)
9th grade	3278 (0708)	3090 (0897)	2810 (0794)	3403 (1198)
11th grade	3934 (2049)	3805 (1883)	2796 (0338)	3034 (0874)
Undergraduate	3826 (2117)	2372 (1155)	2842 (0410)	2152 (0120)
Erroneous sentences	Noun2			
	Plausible		Implausible	
Response type	Correct	Incorrect	Correct	Incorrect
School level				
5th grade	5340 (2374)	4800 (1363)	4858 (1980)	5264 (1964)
7th grade	3981 (2019)	4250 (1075)	3553 (0853)	3944 (1739)
9th grade	3336 (0892)	3119 (0860)	3005 (0894)	3758 (1251)
11th grade	3542 (0964)	3368 (1279)	3619 (1010)	3417 (1290)
Undergraduate	3542 (0964)	2780 (1253)	3506 (1154)	2298 (0965)

(3862 msec) than the incorrect ones (3529 msec). There was no significant main effect of the Subject role plausibility of N2, $F(1,92) = 1.26$, *ns*.

Since the School level interacted significantly with both the Subject role plausibility of N2, $F(4,368) = 3.89$, $p < .005$, $MSe = 8922206$ (see Figure 2), and the Response type, $F(4,368) = 2.84$, $p = .02$, $MSe = 6881877$ (see Figure 3), further analyses were made for each school level.

As hypothesized, the effect of the Response type is significant only for the oldest participants. The undergraduates indeed were the only participants whose correct responses were slower (3334 msec) than their incorrect ones (2262 msec), $F(1,92) = 14.14$, $p < .001$, $MSe = 275761$.

The effect of the Subject role plausibility of N2 is also only significant for the older participants: 11th graders: $F(1,92) = 10.3$, $p < .005$, $MSe = 218619$; undergraduates: $F(1,92) = 4.04$, $p = .05$, $MSe = 868686$. The 11th graders' responses were slower when N2 was a plausible subject of the verb (3869 msec) than when N2 was an implausible subject (2915 msec); the same was true for the undergraduates' responses (3100 msec *vs* 2497 msec). In contrast, the subject role plausibility of N2 has no significant effect on the younger participants' responses: $F(1,92) = .84$, *ns* for the 5th graders, $F(1,92) = .16$, *ns* for the 9th graders, and $F(1,92) = 2.31$, *ns* for the 7th graders.

No other significant interaction was observed. In particular, it is worth noting that the interaction between the Response type and the Subject role plausibility is not significant, $F(1,92) = 1$, *ns*, nor is the interaction between Response type, Subject role plausibility and School level, $F(4,368) = 2.1$, *ns*.

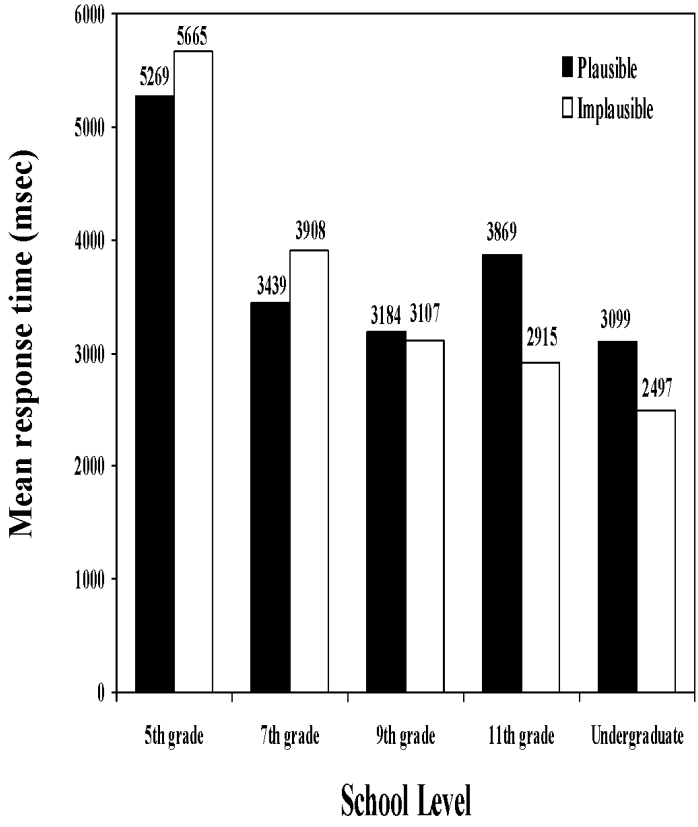

Figure 2. Mean response time (in msec) in each age group as a function of the subject role plausibility of N2 (plausible vs. implausible), for sentences without agreement error

Sentences with an erroneous agreement

For the sentences in which the verb was incorrectly agreed, two types of responses were also considered: (a) the correct responses, when the participants correctly decided that the sentence had an agreement error, and (b) the incorrect responses, when the participants wrongly decided that the sentence had no agreement error. The mean response latencies for those sentences were submitted to the same analysis of variance as above. A significant main effect of School level, $F(4,368) = 32.07$, $p < .001$, $MSe = 609633$, was detected; the mean response latency decreases from the 5th grade level (5063 msec) to the undergraduate level (3022 msec), with intermediate values for the intermediate levels (3932 msec for the 7th graders, 3305 msec for the 9th graders and 3486 for the 11th graders). There was no other significant effect. Considering the responses of all the participants, the variance analysis showed that neither the Subject role plausibility of N2 nor the Response type significantly affected the response latency, respectively $F(1,92) = .45$, *ns*, and $F(1,92) = 1.03$, *ns*.

On the other hand, the School level × Response type interaction was significant, $F(4,368) = 3.5$, $p = .01$, $MSe = 6645742$ (see Figure 4), but none of the other interactions was significant.

The Response type effect was therefore further examined for each school level. Response type was significant only for the undergraduates, $F(1,92) = 19.56$, $p < .001$, $MSe = 232785$, the only participants whose correct responses were slower (3524 msec)

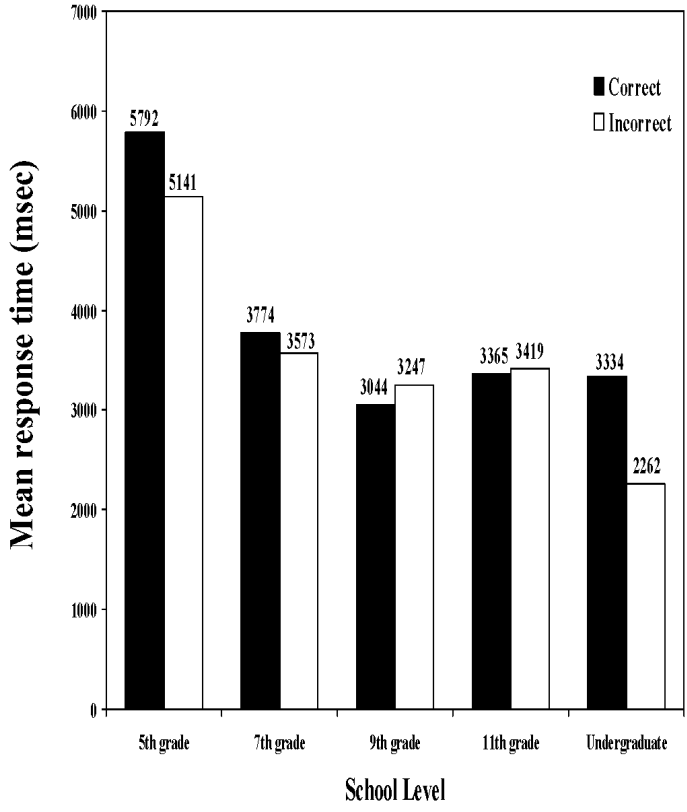

Figure 3. Mean response time (in msec) in each age group as a function of the response type (correct vs. incorrect), for sentences without agreement error

than their incorrect ones (2539 msec). In contrast, the response type had no significant effect on the younger participants' response latency, $F(1,92) = .03$, *ns* for the 5th graders, $F(1,92) = 1.17$, *ns* for the 7th graders, $F(1,92) = 1.69$, *ns* for the 9th graders, and $F(1,92) = .66$, *ns* for the 11th graders.

Discussion

The aim of the present study was to understand how young writers and older ones revise the spelling of grammatical agreements. The question is particularly important in French because of its silent plural morphology where, in general, the written marks of plurality have no corresponding oral forms. Previous research concerned with how writing expertise develops (Fayol *et al.*, 1999; Totereau *et al.*, 1997) showed how French-native children and teenagers progressively acquire the ability to correctly agree the verb of a sentence with its grammatical subject. Focusing on this particular grammatical agreement, the present study aimed to examine whether the acquisition of revising expertise parallels the acquisition of writing expertise. It was hypothesized that, as it has been shown for the production of verb number agreement, young revisers would use a slow and laborious algorithmic procedure to revise the spelling of verb

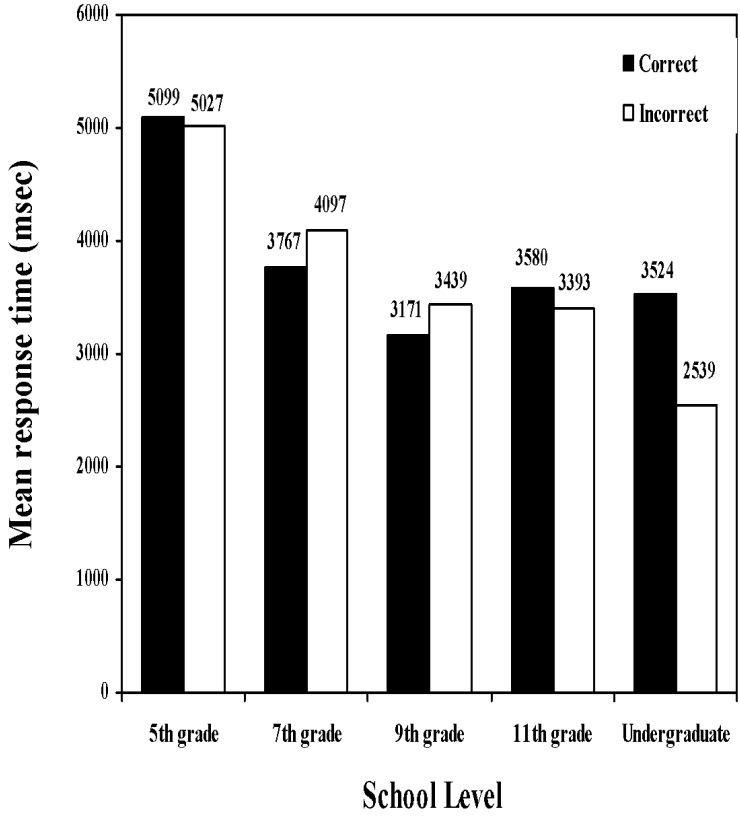

Figure 4. Mean response time (in msec) in each age group as a function of the response type (correct vs. incorrect), for sentences with an erroneous agreement

agreements while adult revisers would use a fast and effortless decision strategy simply based on the detection of matched versus mismatched co-occurring number marks.

Participants from five school levels (from 5th graders to undergraduates) were asked to decide as quickly as possible whether sentences they were visually presented with were correctly or incorrectly agreed. The sentences used as experimental materials were of the type 'Noun1-singular of Noun2-plural + Verb' in which N2 for 50% of the items was a semantically plausible subject for the verb whereas for the remaining items N2 was an implausible subject for the verb. The participants' performance was analysed both in terms of number of revision errors and in terms of response latency.

The data showed that the number of revision errors initially decreases as the younger participants gain more experience of reading and writing but then increases for the oldest participants. Indeed, considering together the false alarms and the missed erroneous agreements, the revision errors decrease from 46% at the 5th grade level to 16% at the 11th grade level, and then increase to 34% at the undergraduate level. The data also showed that the response latency regularly decreases from the 5th grade level (5468 msec) to the undergraduate level (2799 msec). Those observations support the idea that children and adults use different procedures to revise grammatical agreements. Those observations indeed are compatible with the idea that young revisers systematically use an algorithm of verification consisting in applying condition-action rules of the type : *If a word is a verb, if the subject of this verb is singular, and*

if the verb is marked with -nt, the agreement is wrong and the mark -nt must be suppressed. With sufficient practice (around the 9th grade level), the revisers appear to apply a controlled procedure more accurately and more quickly. The adults, on the contrary, appear to use a highly automatized revising procedure relying on proximal co-occurrences of inflections, i.e., on purely formal characteristics of nouns and verbs; in other words, their decision as to whether the verb is correctly or incorrectly agreed largely depends on whether the plurality of the verb and of the immediately preceding noun in the sentence is a match or mismatch. It is therefore predictable that they are frequently misled by the sentential configuration 'N1-singular of N2-plural + Verb-plural', especially in time-constrained conditions. The general pattern of these findings thus confirms that the performance levels corresponding to the two extremities of the U-shaped curve are attributable to two different revising strategies. The youngest participants' frequent errors of revision are due to their imperfect use of a laborious algorithmic verification procedure, whilst the oldest participants' frequent errors are due to a highly risky but fast decision strategy.

The difference between those two procedures is also evident from the differential effect of the subject role plausibility of N2 on younger and older participants' revising performance. When revising relies on an algorithmic procedure that necessarily involves a search for the grammatical subject, some revision errors are likely to result from a misassignment of the subject role, and a misassignment to the preverbal noun N2 is more likely to occur when N2 is a plausible subject. In other words, when relying on a controlled algorithmic procedure, the revision should be better in 'N1 of N2 + Verb' sentences when N2 is an implausible subject. Since the younger revisers are presumed to rely heavily on an algorithmic revision, their performance should be better for sentences with an implausible N2. On the contrary, older revisers who are presumed to simply rely on whether two successive number marks match or mismatch should not be affected by such semantic features of the preverbal items. The reported observations strongly support this view.

The young revisers' performance, however, varies according to whether the participants were presented with correct or incorrect sentences. In the latter case, the data showed that an incorrectly agreed verb was better detected when the preverbal N2 was an implausible subject than when it was a plausible one; as we have just mentioned, this is consistent with the idea that the participants use an algorithmic procedure, a step of which – searching for the grammatical subject – is facilitated when N1 is the only plausible subject. The revision errors in this case may be due to the cognitive overload resulting from both the time constraints and the cognitive costs of the control operations. The idea that the younger participants use an algorithmic procedure is further confirmed by the absence of any difference in response latency between their correct (erroneous agreement is detected) and incorrect (erroneous agreement is not detected) responses. In other words, their revision errors clearly cannot be attributed to faster decisions; whether correct or incorrect, their decisions are more likely to result from the application of a unique procedure. The subject role plausibility of N2, however, had no influence on the younger participants' revision performance when presented with correct sentences. For those sentences, all happens as if the fact that the verb agreed with N1 (and did not agree with N2) facilitated the identification of the subject of the sentence : in those cases, N2 could be immediately disregarded, and consequently its subject role plausibility had no influence on the revision performance.

As Largy and Dédéyan (2002) pointed it out, there is no reason to believe that the

two revision procedures correspond to two strictly distinct developmental steps. On the contrary, the present study provides evidence that the 11th graders' performance shows a pattern which is midway between the youngest participants' pattern and the oldest participants' one; it also reveals that even adult revisers may use both procedures. The 11th graders made relatively few revision errors, which is compatible with the correct use of an efficient algorithmic procedure; they also can come to a decision within a relatively short time, which reflects the development of their expertise, even if it is not based on fully automatized procedures. On the other hand, however, just like the oldest participants, the 11th graders appear to be insensitive to the subject role plausibility of N2; and this is more likely to result from a highly automatic decision procedure based on purely formal aspects of the sentences. These observations are therefore in accordance with the idea that 11th graders still mainly use an algorithmic revision procedure, but also sometimes take their decision on the basis of a direct recognition of co-occurrences of number marks; this second procedure would explain why they make a few revision errors. As an index of their intermediate status, it is also worth noting that the 11th graders, like the younger participants, were faster to reach a decision when N2 was an implausible subject, and were as quick to give a correct response as to give an incorrect one.

The profile of an expert's performance clearly appears in the undergraduates. These participants indeed were clearly affected by the particular syntactic structure of the materials used in the present study : they made the same number of false alarms as the youngest participants (25%), and made more missed errors (43%) than the 7th graders (32%). Yet, their errors seem unlikely to be attributed to an inefficient control procedure: indeed, their decision was both relatively fast and insensitive to the plausibility of N2, which constitute two arguments in favour of an automatized revision strategy. On another hand, the undergraduates' responses were not all incorrect : they sometimes reached the right decision. This suggests that they did not use the same revision procedure in all cases. The observations gained in the present study further support the idea that if their incorrect responses result from the use of a fast decision strategy, their correct ones are more likely to result from the use of a controlled procedure; since this latter procedure is by definition more time-consuming than the former, its application should result in slower responses, which is exactly what was observed with the undergraduates.

In conclusion, the present study confirms the similarities between the developmental patterns that characterize the acquisition of an expertise in the production and the revision of grammatical forms. In both cases, this expertise results from a progressive change in both the writing and revising procedures; resulting from more and more frequent practice, such a change contributes to lower the cognitive costs of processes relative to the management of rather superficial aspects of writing. This view also accounts for why the expertise in revision of grammatical spelling is acquired later than the expertise in production : this delay is likely to be due to the simple fact that revising itself is an activity that only appears some time after the children have learned to produce written sentences.

References

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht, Boston & London: Kluwer Academic.

Anderson, J.R. (1993). *Learning and memory. An integrated approach*. New York: John Wiley & Sons.

Beers, C.S., & Beers, J.W. (1992). Children's spelling of English inflectional morphology. In S. Templeton & D.R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy* (pp.231–251). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bock, J.K., & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45–93.

Bourdin, B., & Fayol, M. (1994). Is written language more difficult than oral language production? A working memory approach. *International Journal of Psychology, 29*, 591–620.

Cameron, C.A., Edmunds, G., Wigmore, B., Hunt, A.K., & Linton, M.J. (1997). Children's revision of textual flaws. *International Journal of Behavioral Development, 20*, 667–680.

Dubois, J. (1965). *Grammaire structurale du français: Nom et pronom*. Paris: Larousse.

Ellis, A.W. (1979). Slips of the pen. *Visible Language, 13*, 265–282.

Fayol, M., & Largy, P. (1992). Une approche cognitive fonctionnelle de l'orthographe grammaticale. Les erreurs d'accord sujet-verbe chez l'enfant et l'adulte. *Langue Française, 95*, 80–98.

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject-verb agreement in written French: From novices' to experts' errors. *Reading and Writing*, *11*, 153–174.

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors. A study in French written language. *Quarterly Journal of Experimental Psychology, 47*, 437–464.

Francis, W.N. (1986). Proximity concord in English. *Journal of English Linguistics, 19*, 309–318.

Hacker, D.J., Plumb, C., Butterfield, E.C., Quathamer, D., & Heineken, E. (1994). Text revision: Detection and correction of errors. *Journal of Educational Psychology, 86*, 65–78.

Hayes, J.R. (1996). A new framework for understanding cognition and affect in writing. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp.1–28). Mahwah, NJ: Lawrence Erlbaum Associates.

Hotopf, W.H.N. (1983). Lexical slips of the pen and tongue: What they tell us about language production. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp.147–199). New York: Academic Press.

Hupet, M., Fayol, M., & Schelstraete, M.A. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology, 89*, 59–75.

Lambert, E., & Chesnet, D. (2001). Novlex: Une base de données lexicales pour les élèves de primaire. *L'Année Psychologique, 101*, 277–288.

Largy, P. (2001). La révision des accords nominal et verbal chez l'enfant. *L'Année Psychologique, 101*, 221–245.

Largy, P., & Dédéyan, A. (2002). Automatisme en détection d'erreurs d'accord sujet-verbe: Etude chez l'enfant et l'adulte. *L'Année Psychologique, 102*, 201–234.

Largy, P., Fayol, M., & Lemaire, P. (1996). The homophone effect in written French: The case of noun-verb inflection errors. *Language and Cognitive Processes, 11*, 217–255.

Logan, G.D. (1988). Toward an instance theory of automatisation. *Psychological Review, 95*, 492–527.

Logan, G.D. (1992). Automaticity and memory. In W.E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays in honor of Bennett Murdock* (pp.234–248). Hillsdale, N.J.: Lawrence Erlbaum Associates.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp.249–308). Cambridge: Cambridge University Press.

McDonald, J.L., & Heilenman, K. (1991). Determinants of cue strength in adult first and second language speakers of French. *Applied Psycholinguistics, 15*, 238–261.

Pearlmutter, N.J., Garnsey, S.M., & Bock, J.K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*, 427–456.

Totereau, C., Thévenin, M.G., & Fayol, M. (1997). Acquisition de la morphologie du nombre en français. In L. Rieben, M. Fayol & C. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp.147–163). Lausanne & Paris: Delachaux et Niestlé.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in

speech: The role of semantic and morphological factors. *Journal of Memory and Language, 34*, 186–215.

Yagelski, R.P. (1995). The role of classroom context in the revision strategies of student writers. *Research in the Teaching of English, 29*, 216–238.

Received 12 February, 2002; revised version received 2 June, 2003