

Do learners need semantics to spell syntactic markers? Plural spellings in real vs. pseudowords in a French L2 setting

Lisa Klasen $^{1} @\cdot$ Sonja Ugen $^{2} \cdot$ Carole Dording $^{1} \cdot$ Michel Fayol $^{3} \cdot$ Constanze Weth 1

Accepted: 30 January 2023 © The Author(s) 2023

Abstract

Inaudible syntactic markers are especially difficult to spell. This paper examines how 455 fourth graders spell silent French plural markers in a dictation with real and pseudowords after one year of formal French instruction (L2). The Generalized Linear Mixed Model analysis shows first that noun plural spelling (real and pseudo) is a strong predictor for verb and adjective plural spelling. Second, the performance on real verb plural is higher than the performance on real adjective plural. In contrast, the performance on pseudoadjective plural is higher than on pseudoverb plural. Our findings indicate the strong influence of semantics and frequency in instruction input on plural spelling: noun plural is semantically grounded, and nouns are most frequent in the curriculum. Verbs and verb plural are also frequent, and inflection is mostly taught by means of memorizing the verb inflection paradigm. Adjectives are taught least frequently. The findings are discussed in the context of French L2 instruction, as the extremely low results on adjectives and pseudoverbs seem to be a consequence of instruction methods.

Keywords French written plural \cdot Plural spelling acquisition \cdot Word categories \cdot Real and pseudowords \cdot L2 French instruction

Orthographic rules are related to the written (orthographic) word and its representation in the lexicon, in general (Bahr et al., 2012; Coulmas, 2003). Many writing

Lisa Klasen Lisa.klasen@uni.lu

¹ Department of Humanities, University of Luxembourg, Campus Belval, Maison Des Sciences Humaines, 11, Porte Des Sciences, 4366 Esch-Sur-Alzette, Luxembourg

² Luxembourg Centre for Educational Testing (LUCET), University of Luxembourg, Esch-Sur-Alzette, Luxembourg

³ University of Clermont Auvergne, Clermont-Ferrand, France

systems contain inflection morphemes, that are written at word level but encode (morpho-)syntactic information (hereafter referred to only as "syntactic") on phrase or clause level. Those written morphemes that are highly systematic in the writing system but inconsistent or not present in phonology have been observed to be more difficult to spell than morphemes that are present in phonology (Bourdin et al., 2011; Morin et al., 2018; Weth, 2020). Such markers exist in French, but also in other written languages such as English (Nunes et al., 1997a, 1997b) and Dutch (Sandra & Van Abbenyen, 2009). The spelling of inflection morphemes that are not present or inconsistent in phonology follow seemingly simple rules. The rules are prominent in spelling instruction, such as ", if plural, add < s > on nouns and adjectives " in French (Brissaud & Fayol, 2018) or , if past tense, add < ed > at the end of regular verbs" in English (Nunes et al., 1997b). These apparently simple orthographic rules require complex syntactic reasoning (Nunes, 1997b; Weth, 2020). Although a syntactic marker like plural spelling in French, for instance, always refers to the underlying syntactic structure of number agreement, a writer does not need to apply syntactic reasoning in all cases but can also rely on semantics or frequency (Fayol et al., 2006; Sandra, 2011).

This paper studies syntactic markers in written French. In French spelling, every word of a plural noun phrase, and the verb in agreement with a plural subject must be marked by a plural morpheme, except for some rare invariant wordforms (e.g., *nez* 'nose'). There are two different regular plural morphemes: <s> for constituents (determiners, nouns, and adjectives), within the noun phrase and < nt > for inflected verbs (3rd person plural). Except for the determiner, plural morphemes are, in general, not represented in phonology (for the exceptional case of liaison, where the plural marker is audible, see Simoës-Perlant et al., 2013). The determiner at the beginning of noun phrases thus distinguishes number morphology audibly and systematically, e.g., *Le chien noir court* [la fjɛ̃ nwaß kuß] 'The black dog runs' versus *Les chiens noirs courent* [le fjɛ̃ nwaß kuß] 'The black dogs run'. The homophony but heterography between singular and plural in nouns, verbs and adjectives is a major reason for spelling difficulties (Largy & Fayol, 2001).

In the past three decades, many studies have been conducted to analyze the spelling and acquisition of French plural inflection according to different word categories. Overall, the studies have shown two major spelling patterns. For young French L1 learners (Thévenin et al., 1999; Fayol, 2003), nouns are spelled better than adjectives, and adjectives better than verbs. For more experienced French L1 learners (Fayol et al., 2006; Totereau et al., 2013; Van Reybroeck & Hupet, 2009), as well as L2 French learners (Ågren, 2008; Ågren, 2009; Bílici et al., 2018; Weth et al., 2021) the spelling pattern differs: nouns are spelled best, followed by verbs and then by adjectives. What these studies show together is that noun plural leads to the highest results, compared to verb and adjective plural. It seems that, if learners do not mark plural on nouns, they most probably do not mark plural on any other word category (Totereau et al., 2013; Weth et al., 2021). The latter conclude that noun plural is acquired first and might be a precondition to acquire adjective and verb plural by analyzing the spelling results of the different word categories. The few existing

longitudinal studies seem to confirm that noun plural is acquired first (Largy et al., 2004; Ågren, 2008; Ågren, 2009).

Indeed, the spelling proficiency and acquisition order of plural markers across the word categories are influenced by semantics, frequency, position, and lexicality. This paper studies the influence of semantics and frequency on French plural spelling compared to plural spelling requiring syntactic reasoning. More precisely we analyse the influence of the plural spelling performance of nouns onto the plural spelling performance of verbs and adjectives. First, we will present these influencing factors in more detail.

Only noun plural is semantically grounded (Fayol et al., 2006) as a reference to tangible objects can be made, e.g., *le chien* 'the dog', *les chiens* 'the dogs'. The semantic information of number in tangible objects makes it easy for learners to grasp the meaning and functioning of the noun plural morpheme and its importance: if more than one dog is referred to, add < s > . Of course, noun plural is also syntactic as the grapheme < s > represents the information ,plural' of the head of the noun phrase. The plural of verbs and adjectives on the other hand is solely syntactic and only refers to the agreement structures: adjectives agree with the noun, and verbs with the subject-noun-phrase. The solely syntactic plural of verbs and adjectives makes it more challenging to grasp and leads to difficulties not only for primary school children, but also for advanced French L1 learners in secondary school (Le Levier & Brissaud, 2020).

Input frequency has a great impact on spelling, as shown by studies on French (Largy et al., 2004; Lété et al., 2008; Ågren & Van de Weijer, 2013a, 2013b), as well as on other languages like Dutch (Sandra & Van Abbenyen, 2009) or English (Kemp & Bryant, 2003). Concerning French plural markers, two different kinds of input frequency must be distinguished: frequency of the plural markers, and frequency of whole inflected word forms.

Concerning the frequency of the plural markers, the marker <s>occurs more frequently and more regularly in the written input of L1 learners than <nt>(Fayol, 2003; Fayol et al., 2006). First graders therefore tend to overgeneralize < s > as plural marker. Consequently, they often correctly mark plural on nouns and adjectives. The overgeneralization leads to less correct spellings and to more substitution errors (<s>instead of <nt>) on verbs. The frequency of the plural markers differs when looking at the instruction input. Noun plural is taught first in the curriculum, which leads to a significant presence of nouns and the marker $\langle s \rangle$ in the input. Verb inflection also plays an important role. It is taught early and practiced repeatedly, leading to a high number of words ending with <nt>and thus often to correct verb spellings (Granget, 2005). Only few adjectives are included in the instruction input. The early instruction of $\langle nt \rangle$ and the later and less frequent instruction of the plural marker of adjectives < s > generate a change in the frequency pattern compared to early L1 learners. This leads to better spelling results on verbs, than on adjectives for more advanced French L1 learners, as well as for L2 learners from the beginning of instruction.

Concerning the frequency of the inflected word forms, singular word forms are, in general, more frequent than plural word forms. Some words occur almost exclusively in the plural form, e.g., *les parents* 'the parents' (Fayol, 2003). The high

exposure to such inflected words in the input can lead to rote learning, that means to the memorization of whole inflected word forms (Lété et al., 2008; Martinet et al., 2004). Learners would then store, retrieve, and spell inflected words without using any rule information (Kemp & Bryant, 2003; Sandra, 2011). Studies by Cousin et al. (2002) and Largy et al. (2007) have shown that learners mark plural more frequently on words, they only encountered in plural, and more frequently omit plural markers on words, they only encountered in singular. Relying on frequency only without applying a syntactic rule can be a source of errors when spelling plural (Largy et al., 1996; Totereau et al., 1998).

Position of a given word in relation to the audible cue is also a factor that influences the correct spelling due to the working memory. In French, the determiner is the sole audible signal to trigger the plural spelling of the following words in a sentence. When spelling a word that is positioned right after the determiner having signalized ,plural ', the audible signal is still present in the working memory and the chance for plural marking of this word is high. The more words between a word and the determiner, the smaller the influence of the audible signal (Fayol et al., 2006). Hence, words positioned right after the determiner, are more often marked correctly than those positioned further away. Beyond the position effect, these words are usually the default position for nouns (Weth et al., 2021).

As the previous paragraphs show, the plural spelling of real words can be influenced by frequency, semantics, and position. To exclude the influence of rote learning and frequency on spelling, several studies introduced the use of pseudowords (Kemp & Bryant, 2003; Martinet et al., 2004; Mussar et al., 2020; Nunes et al., 1997b). As learners cannot rely on memorized word forms when spelling pseudowords, they seem to be the ideal test material to measure whether learners apply spelling rules, or not (Sandra, 2011). Nunes et al. (1997b) used pseudowords to assess the spelling of < ed > endings in English. According to the authors, their results showed that the learners made a distinction between regular and irregular pseudoverbs, even though they were completely unfamiliar. Kemp and Bryant (2003) used pseudowords to test the spelling of the English regular plural ending < s > . The performance on pseudowords was lower than on real words. The authors interpreted their results as an effect of frequency-based rote learning.

The plural spelling of nouns in real words versus pseudowords provides insights to disentangle semantics from the syntactic processes assumed in syntactic spelling. On the syntactic level, pseudonouns are, like real nouns, the head of the noun phrase and determine the number of all words within the noun phrase. But on a semantic level, pseudonouns differ from real nouns as no reference to tangible and intangible objects can be made. Thus, learners cannot rely on semantics, but have to imply plural rules to correctly spell plural on pseudonouns.

This study

This study analyzes the influence of semantics and frequency on the spelling of plural across the word categories, based on two hypotheses. As seen in the literature, the plural spelling performance of nouns is highest compared to adjectives or verbs. Therefore, we seek to understand if noun plural is the precondition for the spelling of verb and adjective plural. The first tested hypothesis (H1) is, that the spelling of noun plural predicts the spelling of verb and adjective plural, regardless of lexicality (real or pseudowords). H1 is in line with findings observed in other studies in fifth grade and secondary school (Totereau et al., 2013; Weth et al., 2021) and it is tested whether these results hold for fourth graders who learned French as a foreign language (L2) for one year. Additionally, H1 adds whether these findings also hold for pseudonouns. Including pseudonouns in the analyses allowed us to test our hypothesis without the influence of noun plural semantics. In pseudonouns, the semantic aspect of noun plural is lacking as no reference to tangible objects can be made. Furthermore, pseudonouns are not part of the orthographic lexicon and rote learning is therefore excluded. In other words, we test the application of rules.

Noun plural is the starting point for all our analyses. In hypothesis 2, we looked at the spelling of verbs compared to adjectives under the condition of high noun proficiency. The second hypothesis is divided into two parts, testing real verbs and real adjectives (H2a), and pseudoverbs and pseudoadjectives (H2b).

Verb inflection is a big part of French instruction. This leads to a high frequency of the <nt> marker and whole inflected verb forms in instruction input. Adjectives on the other hand are introduced later and are therefore less frequent. Hence, the spelling performance on verb plural is higher than on adjective plural. This is in line with other studies on L1 French (Bosse et al., 2021; Van Reybroeck & Hupet, 2009) and L2 French (Ågren, 2009; Bîlici et al., 2018; Weth et al., 2021). This study seeks to add knowledge if the performance difference between verb and adjective plural already shows after only one year of French as L2 instruction. In H2a, we assume that the spelling performance on real verb plural is higher than on real adjective plural, under the condition of high noun proficiency.

Additionally, in H2b we tested whether the performance difference between verbs and adjectives holds for pseudowords. When spelling pseudowords the pupils cannot rely on memorized inflected word forms but have to rely on the syntactic structure of the sentence in order to apply the orthographic plural rules correctly on the pseudowords. In H2b we assume that the spelling performance on pseudoverbs is higher than on pseudoadjectives, under the condition of high noun proficiency.

Method

Participants

The current study takes place in Luxembourg and examines the performance of Luxembourgish fourth graders on French plural spelling. In total, 572 Luxembourgish fourth graders from 37 different classes of 19 different schools participated in this study. For all pupils, French is a foreign language. At the time of the study, the pupils were in their second year of formal French instruction. They have followed the same educational curriculum since first grade: They learned to read and

write in German in grade one, learned oral French as a foreign language from second grade on and to read and write in French from third grade on (MENJE, 2011). The schools were situated in the north and the south of Luxembourg and had a mean estimated Highest International Socio-Economic Index (HISEI) of 42.76 (MENJE, 2015). The schools all followed the same curriculum. As we focus on French as L2, pupils with French as L1 were excluded from the analyses (n=29). To ensure the same curricula input, pupils who were enrolled later than grade 1 in a Luxembourgish school were excluded from the analyses (n=6). Additionally, in order to ensure that the pupils were able to fulfill the French syntactic spelling test, we tested general French orthography (adapted version of BELEC, Bodé et al., 2009; Mousty et al., 1994) and French syntactic comprehension (adapted version of *É.co.s.se*, Lecocq, 2013) in a preliminary test. Pupils not reaching a score of at least 50% correct in these preliminary tests were excluded from the analyses (n=35). In addition, 47 pupils did not participate in the syntactic spelling test. In total, 455 from 572 participants were included in the analyses of the syntactic spelling test.

Due to the multilingual situation of Luxembourg, the participants had various language backgrounds. Language background (L1) was determined by means of a

Table 1BackgroundInformation (L1, Gender,Age, Estimated HISEI, andPreliminary Tests: SyntacticFrench Comprehension and	Background information					
	Ll	N (%)				
General French Orthography)	Luxembourgish	150 (32.97)				
for the total sample ($N = 455$)	Portuguese	133 (29.23)				
	Other	61 (13.41)				
	Missing	111 (24.40)				
	Gender					
		N (%)				
	Female	222 (48.79)				
	Male	225 (49.45)				
	Missing	8 (1.76)				
	Age					
		<i>M</i> (SD)				
	Months	119.09 (8.07)				
	HISEI					
		<i>M</i> (SD)				
	Estimated	42.76 (7.61)				
	Preliminary tests					
		<i>M</i> (SD)				
	Syntactic French Comprehension (max. score = 100%)	78.03 (12.78)				
	General French Orthography (max. score = 100%)	80.65 (9.89)				

parent questionnaire or teacher's information. As 24.40% of the information was missing (due to questionnaires not filled out by parents), it was only possible to give a tendency of the home languages for our sample. Even though the pupils in our sample have different home languages all the pupils followed the same curriculum. Therefore, we did not include L1 in our analyses, as we assumed that their verbal input of French is mostly related to classroom and thus similar for all the children. For an overview on all background information concerning the participants, see Table 1.

Procedure and material

All tests took place between October 2019 and December 2019. Test administration was standardized as a group test in the classroom. Instructions were given by a member of the research group or a trained student assistant. Test administrators followed a script, and the dictation tests were played by means of an audio CD.

The syntactic spelling test consisted of two gap dictation tests, one with real words and one with pseudowords, constructed identically. We tested three different (pseudo)word categories: nouns, verbs, and adjectives. The plural markers for all word categories were regular plural morphemes with < s > for nouns and adjectives and < nt > for verbs (3rd person plural). We excluded *liaison*, audible word endings and irregular verbs with audible plural inflection.

To ensure the comparability of the different test items we controlled the sentence structures and the positions of the test items within the sentences. The test items (underlined, see Table 2) were tested within six different sentence structures, depending on the word category: DN, DNV, DANV, DNAV, DAN, DNA (D=determiner, A=adjective, N=noun, V=verb). Additionally, the items were equally tested in initial and final position within the sentence, e.g., DANVDN and DNVDAN. We excluded pronouns and proper names.

	Word category					
	Noun	Verb	Adjective			
Number of target items (per real and pseudowords)	6	9	12			
Example sentences						
Real words	<i>Mes copains mangent</i> <i>une pizza.</i> 'My friends eat a pizza.'	<i>Les crocodiles ouvrent</i> <i>la bouche.</i> 'The crocodiles open their mouth.'	Mes grands frères regar- dent un film. 'My big brothers watch a film.'			
Pseudowords	Mes joupalles chan- tent une chanson. 'My joupalles sing a song.'	Les machines pras - sent un robot. 'The machines prassent a robot.'	Les bames femmes dansent dans la salle. 'The bames women dance in the room.'			

 Table 2
 Overview of the Test Design with the Word Categories, Number of Target Items and Example

 Sentences for Real and Pseudowords of the Syntactic Spelling Test with Target Items in bold

There were 27 test (plural) items (6 nouns, 9 verbs, 12 adjectives) and 12 (singular) filler items (3 nouns, 3 verbs, 6 adjectives) for each test (real and pseudowords). The real target words were chosen from the official schoolbooks used in third and fourth grade French instruction in Luxembourg. The number of syllables of the test items was controlled: all test items had no more than two syllables. Additionally, we controlled for frequency with the French lexical data base MANULEX (Lété et al., 2004). Only words with a Standard Frequency Index (SFI) larger than 45 were chosen for the test. A SFI of 40 is the value of a word form that occurs once in a million words (Lété et al., 2004). Further, we ensured that the singular form and the plural form of an item were comparably frequent. The vocabulary used to form the sentence contexts was adapted to fourth graders' L2 French skills. The pseudowords were created with the pseudoword generators *Wuggy* (Keuleers et al., 2010) and *WordGen* (Duyck et al., 2004) and followed French graphotactic regularities.

The sentences were recorded by a native speaker. Every sentence was repeated twice. The time between the repetitions was three seconds, and five seconds between the different sentences. Real and pseudowords were presented in separate booklets. There were additional explanations for pseudowords. Each booklet contained 39 sentences. Every sentence contained one gap. The pupils were asked to complete the missing word.

Only the plural endings were considered and coded as correct if $\langle s \rangle$ for noun and adjective plural and $\langle nt \rangle$ for verb plural was marked. Errors were categorized into omissions (no plural marker) or substitutions (wrong plural marker). As only 2.79% of all answers were substitution errors, they were not analyzed separately. Cronbach's alpha was 0.889 for the real word test and 0.907 for the pseudoword test.

Results

We conducted a Generalized Linear Mixed Model (GLMM) using RStudio (version 2021.09.0 Build 351; RStudio Team, 2020) with the package glmmTMB (version 1.1.3; Brooks et al., 2017) to predict the spelling of adjective and verb plural

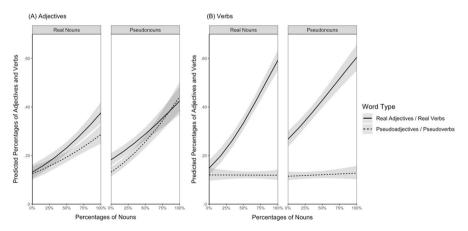


Fig. 1 llustration of the Predicted Percentages on the Word Categories (Adjectives, Verbs) according to the Word Type (Real Adjectives, Pseudoadjectives, Real Verbs, Pseudoverbs) and to the Percentages of Correct Real Nouns and Correct Pseudonouns

(*perc_adj_verb*) with the spelling of noun plural (*perc_nouns*). For the visualization, we used the packages ggeffects (version 1.1.2.; Lüdecke, 2018) and ggplot2 (version 3.3.6., Wickham, 2016). Pure estimates given by the glmmTMB summaries (as indicated in the Tables 4 and 5) may be difficult to interpret. Therefore, we used predicted percentages of adjectives and verbs in Fig. 1 to ease the understanding of the results. As percentages have a lower and an upper boundary in our study, we used the distribution/family *beta* with the link function *logit* (Schmettow, 2021). The estimation method was the maximum likelihood. We tested H1 and H2 with a fitted model that takes the following R syntax form:

perc_adj_verb ~ perc_nouns * wordcategory * wordtype * nountype + (1|class/child_in_class)

The variable *perc_adj_verb* is a dependent variable containing all the percentages of correct plural adjectives and verbs (pseudo and real) and the variable *perc_nouns* is the independent variable containing the percentages of correct plural nouns (real and pseudo). To analyze the hypotheses H2a and H2b, we focus on the lower (0%) and the upper boundary (100%) of the variable *perc_nouns*. The independent variables *wordcategory* (levels: adjective, verb), *wordtype* (levels: real adjective and pseudoadjective, real verb and pseudoverb) and *nountype* (levels: real noun, pseudonoun) are categorical. With the variable *perc_nouns* we predict the influence of noun spelling on *perc_adj_verb*, the spelling of verbs and adjectives (see slope in Fig. 1). With the variable *wordcategory* we analyze the difference between verb and real adjectives, and pseudoverbs and pseudoadjectives. With the variable *nountype* we distinguish between the influence of real noun spelling and pseudonoun spelling. Please note that real noun spelling and pseudonoun spelling are independent predictors in the GLMM.

The variables *class* and *child_in_class* are added as random factors into the model to control for variability between the classes and the children nested within classes. The variance explained by the children within classes (*child_in_class*) is equal to 0.39. The variance between the different classes (*class*) is 0.03. The model has a BIC value of -5943.4 and an AIC value of -6061.2. Before reporting the detailed results of the GLMM in relation to the hypotheses, we briefly present the descriptives.

The descriptives in Table 3 show that for real words, the spelling performance is highest for noun plural followed by verb and then by adjective plural. For

Table 3Mean (M) and StandardDeviation (SD) in % for RealWords and Pseudowords acrossWord Categories		M (SD)		
		Real words	Pseudowords	
	Nouns	59.93 (30.83)	36.52 (33.52)	
	Adjectives	25.55 (24.54)	21.47 (24.36)	
	Verbs	41.47 (36.34)	8.57 (18.60)	

Table 4Estimates of theGLMM (B), Standard Errors(SE) and z-values (z) of theIntercept and perc_nounsfor Real Noun Spelling andPseudonoun Spelling on RealAdjectives, Pseudoadjectives,Real Verbs, and Pseudoverbs		Real noun spelling			Pseudonoun spelling		
		B	SE	z	В	SE	z
	Real adjectives						
	Intercept	-1.886	.128	-14.778	- 1.499	.088	-17.023
	Perc_nouns	.014***	.002	7.631	.012***	.002	7.335
	Pseudoadjectives						
	Intercept	-1.942	.128	-15.232	-1.889	.088	-21.374
	Perc_nouns	.010***	.002	5.684	.017***	.002	10.004
	Real verbs						
	Intercept	-1.748	.121	-14.435	-1.007	.084	-12.036
	Perc_nouns	.021***	.002	12.486	.014***	.002	9.172
	Pseudoverbs						
	Intercept	- 1.989	.125	-15.863	-2.041	.088	-23.140
	Perc_nouns	000	.002	058	.001	.002	.727

****p*<.001; ***p*<.01; **p*<.05

Table 5 Estimates of the GLMM (*B*), Standard Errors (*SE*) and z-values (z) of the Intercept and Word Category (Adjectives vs. Verbs) for Real (H2a) and Pseudowords (H2b), for High and Low Real Noun Proficiency, and for High and Low Pseudonoun Proficiency

	H2a real words			H2b pseudowords			
	B	SE	z	B	SE	z	
Real noun proficiency							
high							
Intercept	506	.095	-5.347	908	.096	-9.434	
Word category (adjectives vs. verbs)	.885**	* .104	8.533	-1.092***	.110	-9.894	
low							
Intercept	-1.886	.128	-14.778	-1.942	.128	-15.232	
Word category (adjectives vs. verbs)	.138	.149	.929	047	.152	309	
Pseudonoun proficiency							
high							
Intercept	302	.120	-2.508	232	.124	-1.879	
Word category (adjectives vs. verbs)	.729**	** .135	5.390	-1.688***	.144	-11.693	
low							
Intercept	- 1.499	.088	-17.023	- 1.889	.088	-21.374	
Word category (adjectives vs. verbs)	.491**	** .098	5.035	152	.102	-1.485	

***p < .001; **p < .01; *p < .05

pseudowords, the spelling performance is also highest for noun plural, but then followed by adjective and lastly by verb plural. Generally, performance is higher for real than for pseudowords. In H1, we assume that the spelling of noun plural (real and pseudo) is a significant predictor for the spelling of adjective and verb plural (real and pseudo). That means, higher scores on noun spelling (real and pseudo) predict higher scores on verb and adjective spelling (real and pseudo). The results of the GLMM show that this is valid for real words of each word category (adjectives and verbs) and for pseudoadjectives (Table 4). Real noun plural and pseudonoun plural are both significant predictors for the spelling of verb and adjective plural. However, noun spelling, whether real or pseudo, is no significant predictor for pseudoverb spelling. Considering our first hypothesis, these findings mainly support that the performance of noun spelling is a significant predictor for adjective and verb spelling, except for pseudoverbs.

In H2a, we assume that the spelling performance on real verb plural is higher than on real adjective plural, under the condition of high noun proficiency. Our findings show that the plural spelling performance on verbs is indeed higher than on adjectives when real noun proficiency is high (Table 5). The same results can be observed when pseudonoun proficiency is high.

For completeness, we also report the results for low noun proficiency. There is a performance difference between verbs and adjectives, when pseudonoun proficiency is low, but not when real noun proficiency is low.

Thus, the hypothesis that the performance on real verb spelling is higher than the performance on real adjective spelling can be confirmed.

In H2b, we assume that the spelling performance on pseudoverb plural is higher than on pseudoadjective plural, under the condition of high noun proficiency. Our results show that the spelling scores of pseudoadjectives are higher than the spelling scores of pseudoverbs when real noun proficiency is high, and when pseudonoun proficiency is high (Table 5).

For completeness, we also report the results for low noun proficiency. There is no significant difference between pseudoverbs and pseudoadjectives when real and pseudonoun proficiency is low.

Thus, the hypothesis that the performance on pseudoverb spelling is higher than the performance on pseudadjective spelling cannot be confirmed. On the contrary, scores on pseudoadjectives are significantly higher than on pseudoverbs, when noun proficiency is high.

Discussion

The current study investigated the influence of semantics and frequency on the performance on French plural spelling of Luxembourgish fourth graders after one year of French as foreign language instruction.

First, our findings show that noun spelling, whether real or pseudo, is a significant predictor for the spelling of verbs and adjectives (H1). A higher noun spelling performance predicts a higher performance on real verbs and adjectives, and on pseudoadjectives. Noun spelling does not predict the spelling of pseudoverbs. The low results on pseudoverbs show that the influence of noun spelling is limited to real verbs and adjectives, and to pseudoadjectives. The results of H1 confirm the findings of previous studies (Totereau et al., 2013; Weth et al., 2021) that noun spelling influences verb and adjective spelling. Pupils who perform higher in noun spelling, also perform higher in adjective and verb spelling.

Our study expands the results of former studies by including pseudowords. The descriptives show that the performance on real nouns (M=59.93%) is higher than on pseudonouns (M=36.52%). Real noun plural is semantically detectable and triggered by the audible presence of the determiner. For pseudonoun plural, however, the determiner is the only cue to determine the number of a given noun phrase including the pseudo noun. Therefore, the lack of a semantic basis has an influence on the spelling of pseudonoun plural. Still, this lack of plural semantics does not seem to make a difference in predicting the spelling of adjective and verb plural. Indeed, the spelling of noun plural (real or pseudo) seems to be the condition for the spelling of verb and adjective plural, and thus a strong predictor independent of plural semantics.

Concerning the second hypothesis (H2a), our findings show that real verb spelling scores are higher than real adjective spelling scores, when real noun spelling is high. The strong focus on verb inflection in instruction input and the resulting high frequency of verbs in the input lead to a higher performance on verbs compared to adjectives, when spelling real and familiar words. These findings confirm the results on the spelling pattern (V>A) from former studies on advanced L1 learners (Fayol et al., 2006; Totereau et al., 2013; Van Reybroeck & Hupet, 2009), on French L2 learners in primary school (Bîlici et al., 2018; Weth et al., 2021), and on adult L2 learners (Ågren, 2008; Ågren, 2009) and complement the results for young French foreign language learners.

The findings of H2b on the difference between pseudoverbs and pseudoadjectives differ from the results observed on real words. The performance on pseudoadjectives is higher than on pseudoverbs, when real noun and pseudonoun spelling are high. The higher results on pseudoadjectives compared to pseudoverbs suggest that spellers may nonetheless rely on the more frequent marker <s>when they have to spell unknown words. Especially when these words are placed immediately after the determiner, the only audible cue for plurality. The position of the adjectives in relation to the determiner was not part of the hypotheses of this study. However, we tested the adjectives in two different positions: prenominal, which is immediately after the determiner, and postnominal, which is after the noun. When looking at the descriptives, we can see that prenominal pseudoadjectives (M = 26.45%) are indeed spelled better than postnominal pseudoadjectives (M = 16.48%). Looking further at the descriptives, we see that the mean real and pseudoadjective scores are nearly equal and both relatively low (real: M=25.55%; pseudo: M=21.47%). In contrast, the scores for real verbs and pseudoverbs are highly different (real: M = 41.47%; pseudo: M = 8.57%). The extremely low results for pseudoverbs indicate that unknown verbal word forms do not seem to be recognized at all as an inflected word category within the sentence context. Furthermore, the big difference between real and pseudoverb spelling indicates that the high frequency of verbs in instruction does lead to correct spellings on familiar real verbs, but it does not lead to the correct inflection of the category verb in unfamiliar words. The pupils do not

apply the subject-verb agreement, that means, the agreement that goes beyond the nominal phrase. They do not use the correct plural marker (substitution error) or any plural marker at all. We did not analyze substitution errors in our GLMM analysis. However, the descriptive results show that out of all answers given, 8.20% were substitution errors on pseudoverbs, compared to 3.66% on real verbs. Additionally, since some of the verbs are tested in the position usually occupied by postnominal adjectives, misinterpretations seem to occur more frequently here. Even though substitution errors are not very frequent for both, real and pseudoverbs, the stronger tendency to substitution errors for pseudoverbs supports the idea that learners do not recognize pseudoverbs as verbs.

To sum up, our findings show that noun plural seems to be the condition for the spelling of verb and adjective plural (H1). Concerning the performance difference between verbs and adjectives, the results show that real verbs are spelled better than real adjectives (H2a). This spelling pattern is reversed when it comes to pseudowords: the scores on pseudoadjectives are higher than on pseudoverbs (H2b).

Our findings indicate the strong influence of semantics and frequency in instruction input on plural spelling: noun plural is semantically grounded, and nouns are most frequent in the curriculum. Verbs and verb plural are also frequent, and inflection is mostly taught by means of memorizing the verb inflection paradigm. Adjectives are the least frequent word category in the curriculum. This is reflected in the spelling pattern of real words: nouns are spelled best, followed by verbs, and then by adjectives.

Plural instruction in the curriculum appears to be strongly word-based and does not address plural within the plural agreement chain of the phrase or sentence. This leads to positive results on real nouns, the only words with a semantic plural. The teaching also seems to show first learning effects on the inflection of familiar verbs. However, as soon as the words are unfamiliar, the pupils do not add plural on the word forms anymore. This occurs with adjectives (real and pseudo), the word category hardly present in instruction, and pseudonouns and pseudoverbs. The limits of plural instruction become especially apparent with pseudoverbs: even though much attention is paid to verb conjugation in instruction, pseudoverbs remain almost entirely uninflected. The way verb plural is taught does not simply lead to the memorization of whole inflected word forms. Otherwise, the pupils would probably pluralize all the verbs in our test. The control items in the singular form in our test show that this is not the case. The pupils spelled correctly over 95% of these control words in the singular form. The results on real verbs show that the pupils already seem to have some verb plural knowledge, but they cannot transfer it to unfamiliar words.

Taking the extremely low results on pseudoverbs, and the low results on adjectives in general into consideration, plural spelling instruction should focus on building an understanding of the noun phrase as an agreement unit, as well as of the subject-verb agreement. This would probably lead to better results on the strictly syntactic plural of verbs and adjectives. Cogis (2004) and Brissaud and Cogis (2011) developed syntactic training methods to help learners understand the syntactic relations, such as plural agreement, between the different words in a phrase. Intervention studies on French as majority language (Nadeau & Fisher, 2014), French as minority language (Arseneau & Nadeau, 2018), and French as L2 (Bîlici et al., 2018) have demonstrated that using these syntactic training tools indeed has a significant positive effect on syntactic spelling, compared to control groups following the usual plural instruction. Bîlici et al. (2018) showed that their syntactic training improved the verb spelling proficiency for real and pseudoverbs. These findings underline that plural instruction should highlight the phrase structures and focus on the syntactic agreement within phrases and sentences.

Limitations

One limitation is that we did not explicitly control for the home languages of the pupils in our study. However, the pupils with French as L1 were excluded from our analysis. Therefore, we assumed that the French input was similar for all the children as all the pupils in our sample followed the same national curriculum.

In our syntactic spelling test, the number of test items per word category might appear rather small. However, the test items included an equal number of real words and pseudowords within each word category.

Additionally, the number of test items varied across the word categories: six nouns, nine verbs and twelve adjectives. This is explained by the fact that we distinguished between prenominal adjectives (n=6) and postnominal adjectives (n=6) to test adjectives in two different positions in relation to the determiner. Since the position of the adjective is not the focus of the present paper, we have taken the word forms in both positions into the word category "adjectives ", resulting in a larger number of adjectives compared to verbs and nouns.

Authors contributions LK: Conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, visualization. SU: conceptualization, formal analysis, writing – review & editing. CD: formal analysis, data curation, visualization. MF: conceptualization. CW: conceptualization, resources, data curation, writing – review & editing, supervision, project administration, funding acquisition.

Funding This paper was supported by the GRASP project, funded by the Luxembourg National Research Fund (FNR, CORE program), grand no: C18/SC/12675187/GRASP/WETH.

Declaration

Conflict of interest The authors whose names are listed above certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Ågren, M. (2008). À la recherche de la morphologie silencieuse: Sur le développement du pluriel en français L2 écrit [Searching for silent morphology: On the development of the plural in French as a second language]. Doctoral Thesis (monograph), French Studies, Lund University).
- Ågren, M. (2009). Morphological development in written L2 French: A processability perspective. In J. Keßler & D. Keatinge (Eds.), 2009 (pp. 121–151). Empirical evidence across languages. Cambridge Scholars Publishing.
- Ågren, M., & Van de Weijer, J. (2013a). Number problems in monolingual and bilingual French-speaking children. *Language, Interaction and Acquisition*, 4(1), 25–50. https://doi.org/10.1075/lia.4.1.02agr
- Ågren, M., & Van De Weijer, J. (2013b). Input frequency and the acquisition of subject-verb agreement in number in spoken and written French. *Journal of French Language Studies*, 23(3), 311–333.
- Arseneau, R., & Nadeau, M. (2018). Expérimentation des dictées métacognitives : Quels effets sur l'apprentissage de l'orthographe grammaticale en contexte francophone minoritaire ? [Experimentation of metacognitive dictation: Effects on syntactic spelling learning in French minority context?]. *Canadian Journal of Applied Linguistics*, 21(2), 126–153. https://doi.org/10.7202/1058464ar
- Bahr, R. H., Silliman, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1–9. *Journal of Speech, Language and Hearing Research*, 55(December), 1587–1599. https://doi.org/10.1044/1092-4388(2012/10-0335)
- Bîlici, N., Ugen, S., Fayol, M., & Weth, C. (2018). The effect of morphosyntactic training on multilingual fifth graders' spelling in French. *Applied Psycholinguistics*, 39(6), 1319–1343. https://doi.org/10. 1017/S0142716418000346
- Bodé, S., Serres, J., & Ugen, S. (2009). Similarities and differences of Luxembourgish and Romanophone 12 year olds' spelling strategies in German and French. Written Language and Literacy, 12(1), 82–96. https://doi.org/10.1075/wll.12.1.04bod
- Bosse, M.-L., Brissaud, C., & Le Levier, H. (2021). French pupils' lexical and grammatical spelling from sixth to ninth grade: A longitudinal study. *Language and Speech.*, 64(1), 224–249. https://doi.org/ 10.1177/0023830920935558
- Bourdin, B., Leuwers, C., & Bourbon, C. (2011). Impact des contraintes linguistiques et cognitives sur l'acquisition de l'accord en genre de l'adjectif en français écrit [Impact of linguistic and cognitive factors in the production of adjectival gender agreement in written French]. *Psychologie Française*, 56(3), 133–143. https://doi.org/10.1016/j.psfr.2011.08.001
- Brissaud, C., & Cogis, D. (2011). Comment enseigner l'orthographe aujourd'hui? [How to teach spelling today?] Paris: Hatier.
- Brissaud, C. & Fayol, M. (2018). Étude de la langue et production d'écrits [Language study and writing production] Cnesco. http://www.cnesco.fr/wp-content/uploads/2018/04/VDEF-rapport-Brissaud-Fayol.pdf
- Brooks, M.E., Kristensen, K., Van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M. & Bolker, B.M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero- inflated Generalized Linear Mixed Modeling. *The R Journal*, 9(2), 378–400. https://journ al.r-project.org/archive/2017/RJ-2017-066/index.html
- Cogis, D. (2004). Une approche active de la morphographie. L'exemple d'une séquence sur l'accord de l'adjectif. [An active approach to morphography. The example of a sequence on the agreement of the adjective.] *Lidil*, *30*, 73–86.
- Coulmas, F. (2003). Writing systems. An introduction to their linguistic analysis. Cambridge University Press.
- Cousin, M.-P., Largy, P., & Fayol, M. (2002). Sometimes early learned instances interfere with the implementation of rules: The case of nominal number agreement. *Current Psychology Letters: Behavior, Brain & Cognition*, 2(8), 51–65. https://doi.org/10.4000/cpl.512
- Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments & Computers*, 36(3), 488–499.

- Fayol, M. (2003). L'acquisition/apprentissage de la morphologie du nombre. Bilan et perspectives. [The acquisition/learning of number morphology. Summary and perspectives]. *Rééducation Orthophonique*, 213, 151–166.
- Fayol, M., Totereau, C., & Barrouillet, P. (2006). Disentangling the impact of semantic and formal factors in the acquisition of number inflections: Noun, adjective and verb agreement in written French. *Reading and Writing*, 19(7), 717–736. https://doi.org/10.1007/s11145-005-1371-7
- Granget, C. (2005). Développement de l'accord verbal avec un sujet pluriel dans les récits écrits d'apprenants germanophones scolarisés du français. [Development of verbal agreement with plural subject in compositions written by German learners of French]. In: Granfeldt, J. & Schlyter, S. (Eds.), Acquisition et production de la morphologie flexionnelle. Actes du Festival de la morphologie. PERLES, 20 : petites études romanes de Lund, 111–124.
- Kemp, N., & Bryant, P. (2003). Do Beez buzz? Rule-based and frequency-based knowledge in learning to spell plural-s. *Child Development*, 74(1), 63–74. https://doi.org/10.1111/1467-8624.00521
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behaviour Research Methods*, 42, 627–633. https://doi.org/10.3758/BRM.42.3.627
- Largy, P., Cousin, M. P., Bryant, P., & Fayol, M. (2007). When memorized instances compete with rules: The case of number-noun agreement in written French. *Journal of Child Language*, 34(2), 425–437. https://doi.org/10.1017/S0305000906007914
- Largy, P., Cousin, M.-P., & Fayol, M. (2004). Acquérir le pluriel des noms. Existe-t-il un effet de fréquence du nom? [Acquiring the plural of nouns. Is there a noun frequency effect?]. *Lidil, 30*, 39–54. https://doi.org/10.4000/lidil.663
- Largy, P., & Fayol, M. (2001). Oral cues improve subject-verb agreement in written French. International Journal of Psychology, 36(2), 121–131. https://doi.org/10.1080/00207590143000009
- Largy, P., Fayol, M., & Lemaire, P. (1996). The homophone effect in written French: The case of verbnoun inflection errors. *Language and Cognitive Processes*, 11(3), 217–255. https://doi.org/10.1080/ 016909696387178
- Lecocq, P. (2013). L'É.co.s.se une épreuve de compréhension syntaxico-sémantique. [L'É.co.s.se a test of syntactic-semantic reception]. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Lété, B., Peereman, R., & Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based Study. *Journal of Memory & Language*, 58, 952–977. https://doi.org/10.1016/j.jml.2008.01.001
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments, & Computers, 36*, 156–166.
- Le Levier, H., & Brissaud, C. (2020). Marquer le nombre du nom et de l'adjectif : une difficulté persistante dans deux corpus de dictées d'élèves avancés. [Marking the number of noun and adjective: a persistent difficulty in two corpora of dictations by advanced students]. SHS Web of Conferences, 78, 07012. https://doi.org/10.1051/shsconf/20207807012
- Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, 3(26), 772. https://doi.org/10.21105/joss.00772
- Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of reading acquisition. *Cognition*, 91, B11–B22. https://doi.org/10.1016/j.cognition.2003.09. 002
- MENJE (Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse). (2011). Plan d'études Ecole fondamentale. [Study plan—elementary school]. Retrieved from: https://men.public.lu/fr/ publications/courriers-education-nationale/numeros-speciaux/plan-etudes-ecoles-fondamentale. html
- MENJE (Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse) and University of Luxembourg (2015): Bildungsbericht Luxembourg 2015. Band 1: Sonderausgabe der chiffres clés de l'éducation nationale 2013/2014. [Education Report Luxembourg 2015. Volume 1: key data on national education 2013/2014]. Retrieved from: https://men.public.lu/en/publications/statistiquesetudes/themes-transversaux/15-bildungsbericht-band-1.html
- Morin, M. F., Alamargot, D., Diallo, T. M. O., & Fayol, M. (2018). Individual differences in lexical and grammar spelling across primary school. *Learning and Individual Differences*, 62(February), 128–140. https://doi.org/10.1016/j.lindif.2018.02.002
- Mousty, P. Leybaert, J., Alegria, J., Content, A., & Morais, J. (1994). BELEC. Batterie d'évaluation du langage écrit et de ses troubles. [BELEC. Written language and disorders assessment battery]. In: Grégoire, J. & Piérart, B. (1994). Évaluer les troubles de la lecture: Les nouveaux modèles

théoriques et leurs implications diagnostiques. De Boeck Supérieur. https://doi.org/10.3917/dbu. grego.1994.01.0127

- Mussar, R., Sénéchal, M., & Rey, V. (2020). The development of morphological knowledge and spelling in French. *Frontiers in Psychology*, 11, Article 146 https://doi.org/10.3389/fpsyg.2020.00146
- Nadeau, M. & Fisher, C. (2014). Expérimentation de pratiques innovantes, la dictée 0 faute et la phrase dictée du jour, et étude de leur impact sur la compétence orthographique des élèves en production de texte. Rapport de recherche FRQSC. [Experimentation of innovative practices, the zero error dictation and the dictated sentence of the day, and study of their impact on the spelling competence of pupils in text production. FRQSC research report] https://frq.gouv.qc.ca/app/uploads/2021/06/pt_ nadeaum_rapport-2014_dictee-impact-orthographique.pdf
- Nunes, T., Bryant, P., & Bindman, M. (1997a). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology*, 33(4), 637–649. https://doi.org/10.1037//0012-1649. 33.4.637
- Nunes, T., Bryant, P., & Bindman, M. (1997b). Learning to spell regular and irregular verbs. *Reading and Writing*, 9, 427–449. https://doi.org/10.1023/A:1007951213624
- Van Reybroeck, M., & Hupet, M. (2009). Acquisition of number agreement Effects of processing demands. *Journal of Writing Research*, 1(2), 153–172. https://doi.org/10.17239/jowr-2009.01.02.3
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. http://www. rstudio.com/
- Sandra, D. (2011). Spelling strategies in alphabetic scripts: Insights gained and challenges ahead. *The Mental Lexicon*, 6(1), 110–140. https://doi.org/10.1075/ml.6.1.05san
- Sandra, D., & Van Abbenyen, L. (2009). Frequency and analogical effects in the spelling of full-form and sublexical homophonous patterns by 12 year-old children. *The Mental Lexicon*, 4(2), 239–275. https://doi.org/10.1075/ml.4.2.04san
- Schmettow, M. (2021). New statistics for design researchers. A Bayesian workflow in tidy R. Springer Cham. https://doi.org/10.1007/978-3-030-46380-9
- Simoës-Perlant, A., Loury, F., Largy, P., Gunnarsson, C., & Soum-Favaro, C. (2013). L'effet de la liaison en production écrite chez l'enfant dyslexique et normo-scripteur. [The effect of liaison in written production in dyslexic and normally developing children]. ANAE - Approche Neuropsychologique Des Apprentissages chez L'enfant, 124, 327–333.
- Thévenin, M.-G., Totereau, C., Fayol, M., & Jarousse, J.-P. (1999). L'apprentissage / enseignement de la morphologie écrite du nombre en français. [The learning/teaching of the written number morphology in French]. *Revue Française de Péda*gogie, *126*, 39–52. https://www.persee.fr/doc/rfp_0556-7807_1999_num_126_1_1093
- Totereau, C., Barrouillet, P., & Fayol, M. (1998). Overgeneralizations of number inflections in the learning of written French: The case of noun and verb. *British Journal of Developmental Psychology*, 16, 447–464.
- Totereau, C., Brissaud, C., Reilhac, C., & Bosse, M. L. (2013). L'orthographe grammaticale au collège: Une approche sociodifférenciée. [Grammatical spelling in college: A socio-differentiated approach]. ANAE - Approche Neuropsychologique Des Apprentissages chez L'enfant, 123, 164–171.
- Weth, C. (2020). Distinguishing syntactic markers from morphological markers. A cross-linguistic comparison. Frontiers in Psychology, 11, 1–11. https://doi.org/10.3389/fpsyg.2020.02082
- Weth, C., Ugen, S., & Fayol & M., Bîlici, N. (2021). Spelling patterns of plural marking and learning trajectories in French taught as a foreign language. Written Language and Literacy, 24(1), 81–110. https://doi.org/10.1075/wll.00048.wet
- Wickham, H. (2016). ggplot2: Elegant data analysis. Springer. https://doi.org/10.1007/ 978-3-319-24277-4

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.