openjournals.nl

Dutch Journal of Applied Linguistics
https://www.dujal.nl | E-ISSN 2211-7253

Structure-based versus Dynamic Usage-Based instruction: L2 French writing skills after six years of instruction in high school

Wim Gombert[1], Merel Keijzer[1] and Marjolijn H. Verspoor[1,2]
[1] *University of Groningen* | [2] *University of Pannonia*

Abstract In the Netherlands, teaching programs for French as a foreign language in secondary school usually involve an explicit focus on grammar. This is partially motivated by early findings in SLA research, which showed that explicit instruction is more effective in foreign language acquisition (Norris, & Ortega, 2000; Spada, & Tomita, 2010) and is even considered essential to achieve accuracy in advanced writing (Gunnarsson, 2012). The aim of this classroom study is to test these claims as it compares a structure-based (SB) method to a Dynamic Usage-Based (DUB) method in developing writing mastery in a pre-university program for French. The results suggest that both programs are equally effective in achieving grammatical accuracy and obtaining general text scores, but a DUB program seems more effective in achieving lexical complexity and fluency.

Keywords explicit-implicit, exposure, French, dynamic usage-based, structure-based

Article history
Received: August 5, 2021
Accepted: July 11, 2022
Online: December 8, 2022

Corresponding author
Wim Gombert, w.gombert@rug.nl

Author contributions
Wim Gombert (ORCID 0000-0002-7826-0692), conceptualization, methodology, investigation, writing – original draft; Merel Keijzer (ORCID 0000-0001-9041-8563), conceptualization, methodology, writing – review and editing; Marjolijn H. Verspoor (ORCID 0000-0001-8828-8405), conceptualization, methodology, writing – review and editing

Funding information
This paper is part of a research project which was funded by a Dudoc-Alfa grant for teachers.

Conflicting interests
None

Supporting information
None

1 Introduction

About 20 years ago, Long (2000) wrote his seminal paper on the differences between Focus on Form, Focus on Forms, and Focus on Meaning. In this article, he pointed at the lack of a widely accepted linguistic theory to form the foundation of communicative language teaching. At that time, generativism (Chomsky, 2009) was still the most commonly known and accepted theoretical linguistic framework, favoring a focus on grammar, even within a communicative approach. As a result, language instruction still predominantly focused on language as an object rather than a medium of communication. In the last decade or so, usage-based linguistic theories with its basic tenets of learning through use and exposure have found their way into the field of applied linguistics (see Tyler, &

How to cite Gombert, W., Keijzer, M. and Verspoor, M.H. (2022). 'Structure-based versus Dynamic Usage-Based instruction: L2 French writing skills after six years of instruction in high school', *Dutch Journal of Applied Linguistics* 11: DOI 10.51751/dujal10935

Ortega, 2018). Rather than a focus on forms, focus on form and meaning have become more and more important theoretically.

Against this backdrop, it needs to be pointed out that most foreign language (FL) teaching in secondary schools continues to be predominantly explicit and structure-based (SB) (Lightbown, & Spada, 2021, p. 166); language teaching practices in the Netherlands are no exception (West, & Verspoor, 2016). Although most FL course books used in Dutch schools claim to follow a communicative design inspired by the Communicative Language Teaching (CLT) approach and aim to develop communicative competence as advocated by different CLT proponents like Halliday (1970), Hymes (1972) and Widdowson (1978), a different picture emerges when observing classroom practices. The gradual acquisition of selected structural and lexical items constitutes the backbone of these course books, which often adopt an explicit and deductive approach to grammar, emphasize lexical and grammatical accuracy and impose the use of the L1 as the language of instruction (Popma, 1997; Hermans-Nymark, 2006; Dönszelman, 2019).

But during the last two decades, other language teaching approaches – more in line with usage-based theoretical perspectives on language learning and often labeled Dynamic Usage-Based (DUB) approaches – have been introduced in a number of Dutch secondary schools. The major tenets of a DUB approach are in line with a Focus on Form approach, but in addition, they include frequent and repeated exposure to meaningful and comprehensible input containing full chunks of language, active use of the FL, an implicit and inductive approach to grammar and a focus on productive skills in the classroom (Rousse-Malpat, 2019).

Dutch teachers are not eager to adopt such approaches. Even though they acknowledge that DUB approaches and a focus on oral production may help the learners in developing listening and speaking skills, they worry that they may do less well on reading (especially for the final exam) and writing skills (Rousse-Malpat, 2012; West, & Verspoor, 2016). Therefore, several studies have been conducted to compare the long-term effects of the more traditional SB teaching program to those of a DUB-inspired teaching program on all four language skills: listening, speaking, reading, and writing. The aim of this particular study is to compare the students' writing skills in the two programs after six years of instruction, building on Rousse-Malpat's work (Rousse-Malpat, 2019).

2 Structure-based versus Dynamic Usage-Based instruction

In line with generativist linguistic theory, an SB approach sees the FL as a system of structurally related elements that encode meaning. Language learning within an SB approach is inherently rule-driven (Lightbown, & Spada, 2013) and the development of FL proficiency benefits from learning to apply grammatical, morphological and phonological rules and focusing on accuracy. The structural elements of different subsystems

are usually presented from simple to complex, making language structure an essential backbone of SB course books.

In line with usage-based theory, a DUB approach (see Verspoor, 2017) views language as a complex dynamic system that itself consists of different interacting sub-systems like the lexicon, syntax and morphology. The language system is complex and dynamic in the sense that it is connected to other systems in its environment, like for instance the general cognitive system and the affective system, and that its subsystems (e.g., lexicon, syntax, phonology) are not modular but interact. Furthermore, language as a complex system is dynamic in the sense that it evolves with a changing environment and changing input patterns. Second language development is the result of a dynamic interplay between internal resources like general aptitude, degree of motivation, eagerness to learn, and attention as well as external resources, including the degree of exposure or the effectiveness of an instructional approach (De Bot, & Larsen-Freeman, 2011). Language thus emerges through use and in interaction with different sub-systems that themselves foster change over time.

Key elements in language learning from a DUB perspective are repetitive exposure to meaningful input and authentic language use (Ellis, 2008; Langacker, 2000; Tomasello, 2003). When sufficiently exposed to the target language for regularities to become noticed, the learner comes to use distributional information to bootstrap knowledge, resulting in language acquisition (Onnis, 2012). Linguistic units are learned as they are "heard and used frequently and therefore entrenched, which is the result of habit formation, routinization and automatization" (Verspoor, & Schmitt, 2013, p. 354).

A fundamental difference between SB and DUB perspectives is the type of instruction it assumes to be needed for language learning. Operationalizing a SB perspective into a teaching program typically results in an explicit approach with a focus on understanding language structure and producing accurate language. Operationalizing a DUB perspective into a teaching program, on the other hand, usually involves an implicit approach to morphosyntax and an emphasis on exposure and active use of the FL and fluency in language production.

One of the language domains that has been strongly associated with SB instruction is writing. The distance between the writer and the reader in both time and space implies the absence of feedback from the reader and therefore necessitates the creation of coherent and understandable texts. FL writers generally spend a great deal of time on the three cognitive sub-processes involved in writing: planning, formulation and revision (see Hayes, & Flower, 1980). Indeed, both Fayol (1997) and Barbier (1997) have shown that the process of writing is much slower than the process of speaking, and it has been argued that especially in the area of accuracy in French FL writing, implicit knowledge might not be sufficient: "in order to ensure accuracy in the low-level aspects of the text, writers also use their explicit knowledge, especially in the case of writing in L2 French" (Gunnarson, 2012, p. 249).

Based on these opinions, it is reasonable to assume that writing skills are more likely to be affected than other skills in a paradigm shift from SB explicit approaches to DUB implicit approaches. However, this has not been supported by a number of recent longitudinal classroom studies conducted in Dutch secondary schools with free response data.

In a classroom study comparing explicit instruction and implicit instruction, Andringa et al. (2011), who present an excellent overview of the explicit-implicit debate, tested eighty-one 12–18 year old learners of L2 Dutch on their use of explicit knowledge in a free written response task and found that after four months of L2 Dutch instruction in which exposure was controlled for, there was no advantage of explicit instruction over implicit instruction on a free writing response task.

In a three-year longitudinal classroom study in the first three years of a Dutch high school (with participants aged 12–15), Rousse-Malpat (2019) conducted a large cohort study with 229 students examining the development of speaking and writing proficiency of French as a FL in several Dutch schools using the same teaching programs as in the present study: one structure-based and the other dynamic usage-based. After three years, the DUB learners outperformed the SB learners in both skills, which may have been due to the large difference in FL exposure. In one SB group, though, the teacher spoke FL French for the greater part, and in Rousse-Malpat (2019) this group was compared to a DUB group of learners with the same scholastic aptitude level on writing skills. The SB group and DUB group scored the same on holistic scores, but the DUB wrote significantly longer texts, suggesting a greater ease in language production. The DUB group also wrote longer sentences, suggesting a higher overall complexity, and used significantly more advanced morphological forms. A previous study by Rousse-Malpat and Verspoor (2012) showed that the SB groups outperformed the DUB group in terms of accuracy after one year, but this difference disappeared after two years, most probably because DUB learners need more time to discern the morphological patterns.

The classroom study by Piggott et al. (2020) involved 416 Dutch learners of English as a foreign language and investigated the effectiveness of a two-year program with explicit grammar instruction and a program without explicit grammar instruction. All 416 students used the same coursebooks, but in the implicit condition, the grammar explanations were removed, and the time left was used for more listening and reading tasks from the book. The study showed that the effectiveness of the program was associated with different aspects of language performance: while the explicit group performed better on accuracy measures in general, the implicitly taught group performed better on complexity and fluency measures. Holistic ratings showed no difference for vocabulary, but the explicit group outperformed the implicit group in relation to the grammar ratings that the experts provided.

To summarize, most CLT programs in the Netherlands are still inspired by SB views, and teachers tend to favor explicit instruction on morphosyntax and accuracy, using the L1 as the language of instruction. This reminds us of a Focus on Forms approach. Yet,

several longitudinal studies in the Netherlands with free response data have shown that accuracy can also be achieved with implicit, high exposure approaches.

The current study (part of a larger study in which all four language skills are examined) explores the effect of two teaching programs for French as a FL, the same as those in Rousse-Malpat (2019), but rather than comparing the results after three years of instruction, the current study focuses on the writing results after 6 years of instruction. The research question is whether a DUB program (without explicit attention to forms) is as effective as a SB program (with a great deal of explicit attention to forms) in achieving accuracy and developing the writing skills of Dutch VWO (pre-university) students in a 6-year teaching program. Our expectation, based on the long-term studies mentioned above, is that with enough exposure to and active use of the FL, a DUB approach might work as well as an SB approach in developing writing skills, not only on overall scores but also on morphological accuracy.

3 Method

3.1 Participants

The 56 learners in the current study are from two different cohorts: one group started in 2010 and graduated in 2016 and the other started in 2011 and graduated in 2017. All participants started at the age of 12 as true beginners and they left school at the age of 17–18. They were tested in their final year on their writing skills. The SB group consisted of 24 learners (5 male; 19 female). The DUB group consisted of 32 learners (6 male; 26 female). Their first language was Dutch, and all learners were enrolled in Dutch Voortgezet Wetenschappelijk Onderwijs (pre-university program), which is the highest secondary educational level in the Netherlands. During the first three years, French was compulsory, and students had different teachers, while in the last three years French was optional, and all students had the same teacher.

3.2 Teaching programs

In both the SB and DUB programs, students had two 50-minute lessons per week in the first three years, in which French is compulsory for all students, and three 50-minute lessons per week in the last three years, in which French is optional. The amount of total classroom instruction time can thus be estimated at 450 hours.

For the SB-program, two similar course books were used: Grandes Lignes (Bakker et al., 2005) in the first three years, and Libre Service (Breek et al., 2003) in the final three years. In these textbooks, which are widely used in the Netherlands, there was a focus on reading, writing and listening activities (one third of the average course book), but a substantial amount of time (again one third of the average course-book)

was spent on the acquisition of grammatical and lexical knowledge through explicit rules and word lists to be memorized. As target language use with these kinds of activities is considered to reduce the students' comprehension and henceforth make learning less effective (Van Compernolle, 2015), the use of the FL was limited because much time – via L1-medium instruction – was spent on explaining grammar, on teaching reading and listening strategies and on testing reading and listening comprehension.

For the DUB program, two complementary communicative, DUB-inspired methods were used. In the first three years, the Accelerative Integrated Methodology (AIM; Maxwell, 2001) was used. AIM is a story-based program in which teacher scripts are used for controlled, oral input activities at the beginning of each lesson and meaningful use of the foreign language takes place during different activities in small groups that follow. A focus on oral skills precedes the development of written skills and in the first six months students are exposed to spoken language only. After six months, written text is introduced. Classroom activities focus on meaning and repetition without explicit grammar instruction. Key to the AIM method is target language use and the use of gestures (including gestures for some grammatical features such as gender) to enhance multimodal learning and to facilitate meaningful use of the FL (See Arnott, 2011, for more details about AIM). In the final three years, an extended version of AIM, labeled 'AIMe', was used as the AIM materials are only available for the first three years of secondary education. This extension of AIM provides authentic input through online sources (short videos and texts from the internet) and more controlled input through FL magazines (texts only) amongst others to facilitate a large amount of exposure to the language, followed by lesson activities with a main focus on speaking.

3.3 General writing instruction in both programs

In the SB program writing instruction was in line with a Focus on Forms approach. It focused on the teaching of grammar through closed type[1] and translation tasks[2] around communicative themes and writing skills were mainly tested by the same kinds of tasks. As tense use is known to be one of the most difficult aspects of French as a FL (Klein et al., 1995), explicit tense use instruction is an important part of writing instruction. Students in the SB program were never engaged in any type of extended writing until the final year, when they were usually invited to write a formal letter, which was assessed on the basis of language accuracy and other aspects of writing like formality conventions, and punctuation.

In the DUB program writing instruction was most in line with a Focus on Form approach. It started after the first six months of FL instruction and focused on guided writing tasks like story retelling and story extension on the basis of known stories and later on continued with free narratives. In the last three years writing instruction con-

sisted mainly of argumentative writing tasks. Writing activities were usually done in class and peer-assistance was used as a means to develop linguistic awareness. The tests were assessed on both content and language proficiency.

3.4 Testing writing skills after six years

To control for topic effects, the same 30-hour program was used during the final two years in both teaching programs to introduce seven academic topics such as Migrants, Tattoos, Abortion, etc. During 6 to 7 lessons, a topic was introduced through a video-documentary in French with various exercises entailing repeated exposure to the language used. Free response writing assignments were given to the students to enable them to practice for the assignment used in the current study.

3.5 Testing procedure

The writing test for both cohorts took place in Year 6. The teacher selected four topics for students to prepare and during the test, students were offered two of those topics and they were asked to write an essay of a minimum of 200 words on one of these topics within 45 minutes. The students wrote the essay in the computer lab at their school and handed them in digitally. Supportive tools (e.g., spelling- and grammar checker) were not turned off during the test, but they were hardly used given the time restrictions (the teacher was able to monitor all screens by using specialized software). The anonymized essays from both cohorts were assessed at the same time through holistic ratings by expert teachers, machine-mediated morphosyntactic profiling and by means of analytical CAF measures (see below).

3.6 Holistic ratings by expert teachers

To rate the texts holistically, the same method was used as in Verspoor et al. (2012). A group of 9 French teachers were asked to rank 10 texts in terms of general proficiency in several rounds until consensus was reached. These texts were used as benchmarks and rubrics were created to describe the benchmarked texts (see Appendix). After this five-hour session, the nine raters were divided into three groups of three raters each. The texts were divided in 4 batches of 12 and a final batch of 8. To avoid bias, the raters were randomly divided into new groups with each new batch of essays. Agreement among the raters was high. In SPSS (version 27), the Intraclass Correlation Coefficient (two way random, consistency with a 95% confidence interval) on the ratings produced by three groups of three independent raters was $r = 0.893$ ($p < .001$).

3.7 Machine-mediated morphosyntactic profiling

Direkt Profil (DP) is one of the few corpus tools available for FL French and was developed at the University of Lund in Sweden (Granfeldt et al., 2006). In total, the software bases its analyses on 142 different analytic text measures in profiling the morphosyntactic content of any (learner) text including subject-verb agreement, tense use, number of conjunctions, etc. Three different algorithms (Granfeldt et al., 2006) produce a profile based on the six morphosyntactic stages of development, from beginner to native speaker, identified and defined by Bartning and Schlyter (2004) and the results were validated by Granfeldt and Ågren (2014).

3.8 Analytical text measures

To support the holistic and overall Direct Profil judgments mentioned above analytically, a number of text measures was used. Different broad complexity, accuracy and fluency (CAF) measures that showed almost linear change across proficiency levels from beginner up to intermediate proficiency levels (CEFR level B1) in English (see Verspoor et al., 2012) were chosen to support the findings of human-mediated and machine-mediated ratings. As coherence and cohesion already played an important role in the holistic judgments (see Appendix A) and the writing samples were too short to provide useful data, this text measure was not included in the analysis.

For complexity Tense Use, Guiraud Index and Sentence Length were chosen. For Tense Use, the relative use of tenses other than the Present Tense was seen as a measure of verb phrase complexity (Granfeldt, & Ågren, 2014). Tense Use was computed on a 10-point scale ranging from 0 (only Present Tense) to 10 (only other tenses), where a score of 6 implied that 6 out of 10 tenses were tenses other than the present tense. The Guiraud Index was chosen as measure of lexical diversity for texts containing more than 200 words (Van Hout, & Vermeer, 2007).

Although Biber et al. (2011) claim that phrasal complexity (i.e., the use of different modifiers in a noun phrase) is a better indicator of writing proficiency than clausal complexity, average sentence length (full sentence, including coordinate and subordinate clauses) was taken as the third complexity measure (Norris, & Ortega, 2009; Oh, 2006; Yoon, 2017). The data investigated by Biber et al. (2011) consists of research articles, written by proficient (academic) writers, while the participants in the Rousse-Malpat (2019) study were beginners (CEFR level A1-A2) and the participants in this study had an intermediate level (CEFR level B1). Consequently, the language investigated in both studies can be considered as conversational French and sentence length is an appropriate measure for examining the linguistic quality of writings.

Accuracy measures were Subject-Verb Agreement (SVA) and Determiner-Noun Agreement (DNA) as they are expected to contribute significantly to accuracy in FL French (Ågren et al., 2012). Both SVA and DNA were calculated on a scale ranging from 0 (no

agreement) to 10 (100% agreement). Finally, text length, operationalized as the total number of tokens in the text, was taken as a fluency measure. Direct Profil provided information on tense use and accuracy. Vocabprofilers (Cobb, 2018) was used to calculate the Guiraud Index, average sentence length and text length.

3.9 Statistical design

Holistic ratings, morphosyntactic profiling scores, and CAF measures were inserted in SPSS (Version 27) and after assumptions of normal distribution and homogeneity of variance were checked, Independent Samples t-tests were conducted ($p < .05$) on all variables.

There were no outliers in the holistic and morpho syntactic profiling scores: the scores of participants for each teaching program were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p > .05$). The mean score and the Standard Deviation were computed, and the effect size was calculated, using Cohen's d (Cohen, 1988).

As to the CAF measures, there were minor outliers in the data. For most measures, the scores were normally distributed as assessed by Shapiro-Wilk's test ($p > .05$) and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p > 05$). For Average Sentence Length, there were only minor violations in the SB group (one outlier and no normal distribution). Hence, a non-parametric test was not assumed necessary for analyzing these data and we decided to use only the regular parametric test.

For text length, however, there were outliers in both groups (one higher and three lower) and both assumptions of normal distribution and homogeneity of variance were violated. Therefore, after having observed symmetry between the shape of distribution of both groups, we used the Mann-Whitney U test for analyzing scores on text length.

For the interpretation of effect sizes of both holistic and analytical measures Plonsky and Oswald's (2014) SLA field-specific benchmarks were used: small ($d = 0.4$), medium ($d = 0.7$), and large ($d = 1.0$).

4 Results

The holistic scores by expert teachers and the level of writing proficiency provided by a computer program for morphosyntactic analyses of written FL French are summarized in Table 1.

Table 1 Overview of holistic scores by expert teachers and morphosyntactic profiling

	SB program *N* = 24 Mean (*SD*)	DUB program *N* = 32 Mean (*SD*)	Cohen's *d*	Significance (2-tailed)
Holistic scores by expert teachers (scores 1–4)	2.50 (0.95)	2.58 (0.92)	0.09	.757
Morphosyntactic profiling by Direkt Profil (scores 1–6)	4.05 (1.22)	4.25 (1.14)	0.17	.534

Although the DUB learners scored higher on both measures, the differences were not significant and effect sizes were low.

Table 2 provides an overview of the results on CAF measures.

Table 2 Overview of CAF measures

	SB program *N* = 24 Mean (*SD*)	DUB program *N* = 32 Mean (*SD*)	Cohen's *d*	Significance (2-tailed)
Complexity:				
Guiraud Index	8.97 (1.12)	8.80 (1.00)	0.16	.541
Tense Use Ratio	2.68 (1.93)	2.43 (1.05)	0.16	.584
Average Sentence Length	14.72 (3.73)	17.33 (2.78)	0.79	.001***
Accuracy:				
Subject-Verb Agreement	7.33 (1.63)	7.90 (1.28)	0.39	.147
Determiner-Noun Agreement	9.15 (0.83)	8.83 (0.99)	0.35	.213
Fluency:				
Text Length	293 (75)	356 (128)	0.60	.044*

* significant at $p < .05$
*** significant at $p < .001$

The DUB students produced significantly longer and more complex sentences and produced longer texts overall with medium effect sizes (Plonsky, & Oswald, 2014). For a better understanding of these results, the first paragraphs of two essays of two students will be given as an example:

> Example 1: A DUB-student (Respondent 713, total number of words: 770, 17 chunks)
> *Conforme à l'émission, aujourd'hui plus et plus des femmes se font avorter au Chili. Les interventions qu'elles ont subi, ont été clandestine, dangereuse. Souvent, les femmes se font avorter à domicile avec des médicaments vendus sur le marché noir. Ces médicaments causeront peut-être maladies entre les femmes. De plus, les femmes sont en danger et vivent dans la peur, parce qu'avorter est un crime dans son pays. Chaque jour de ses vies, les femmes peuvent être poursuivie par la justice. Dans l'émission on raconte Marie, elle est une de ses femmes qui risque d'aller au prison.*

> Example 2: A SB student (Respondent 607, total number of words: 243, 13 chunks)
> *Les migrants sont tout le monde. Et tout le monde parle des migrants. J'ai vu une émission des migrants au dérivé. C'était encore un sujet difficile. Dans la documentation j'ai vu plus des migrants désespérées. Les migrants travaillent sur un bateau avec beaucoup de personnes. La même chose que l'animal. Le travail sur le bateau est trop cher pour les migrants. Environ 5000 euros par personne. Plus de migrants est réfugiés. Ils cherchent la sécurité et en particulier: ils cherchent la paix. Parce qu'il y a une situation dangereuse dans leur pays. L'environnement n'est pas paisible.*

These extracts clearly show both significant differences reported in Table 2. The DUB student not only outperformed the SB student with regard to the total number of words, but also to the average sentence length. Moreover, the sentences produced by the SB student were predominantly simple (no subordinate clauses) while the sentences of the DUB student were more complex (more subordinate clauses). But the extracts not only show that sentence complexity contributes to a higher perceived quality of the text but also the use of chunks. In another sub study (Gombert et al., submitted) it appeared that chunk coverage (percentage of words in chunks out of total words) is much higher in the extract of the DUB student (52%) when compared to the extract of the SB student (35%).

As far as accuracy is concerned, the DUB students performed slightly better than SB students when conjugating verbs, but SB students performed slightly better than DUB students when applying agreement rules pertaining to gender and number. In neither of these cases, however, the differences were significant.

5 Discussion and conclusion

In the Dutch context, teachers prefer to hold to traditional ideas about FL instruction with a great deal of explicit attention to forms because they fear that students will be less accurate if they do not. Especially for writing, the argument has been that explicit attention to morphosyntax is needed to achieve accuracy (Fayol, 1997; Barbier, 1997; Gunnarson, 2012; Ellis, & Wulff, 2015; Hulstijn, 2015). The current study was designed to test this assumption and compared learners on French writing skills in two different FL teaching programs – a structure based (SB) and a dynamic usage-based (DUB) program – after six years of high school instruction in the Netherlands.

In line with Long (2000), the SB program could be considered a Focus on Forms approach with explicit explanations in the L1 on French grammar and relatively little true exposure in the FL, especially in the first three years. The DUB program could be considered a Focus on Form approach with implicit attention to form and a great amount of FL exposure and use.

The most interesting finding was that the SB and DUB groups did not differ much at all in holistic ratings given by a group of experts, in the proficiency level score produced by Direct Profil, which included many accuracy measures, nor in the two specific accuracy measures focused on a great deal in the SB classes: subject-verb agreement and determiner-noun agreement. With regard to specific CAF measures, the results show a difference between the approaches. The DUB learners produced longer texts, which may be a general indicator for fluency, and longer sentences, which is a general indicator of sentence complexity. In a study on the same students (submitted), we also found differences in chunks. The DUB learners used longer and more lexically based chunks, and therefore relatively more words that could be classified as part of chunks than the students in the SB condition ($d = 0.73$, $p < 0.05$). This was clearly illustrated by the two examples in the previous section which showed a higher chunk coverage of the DUB student.

When we relate these findings to the literature reviewed, the results of this study clearly align with previous classroom studies conducted in Dutch secondary schools (see Andringa et al., 2011; Rousse-Malpat, 2019; Piggott et al., 2020) with regard to complexity and fluency. In all studies, the implicit teaching programs appear to be as effective as explicit teaching programs. As for accuracy in writing, Piggott et al. (2020) reported that students in the explicit condition performed better on accuracy measures, while in this study students in the implicit condition performed better on complexity and fluency measures and equally well on accuracy measures. This seems logical as the implicit group in the Piggott et al. (2020) study was tested after two years, while in this study, students were tested after six years, and as Rousse-Malpat and Verspoor (2012) showed, the implicit learners seem to take longer to internalize the more subtle morphosyntactic patterns.

Of course, the current study has its limitations. Empirical studies usually require full experimental control to allow generalization of findings, but classroom research requires

high ecological validity to support teaching practice. Moreover, if this type of classroom research is widely replicated, as advocated by DeKeyser and Botano (2019), this lack of methodological rigor will be compensated as results will become more robust and more relevant for practitioners.

This study also has several strengths: First, although results were only obtained at one point in time (after six years) and development of writing skills cannot be determined over time, its duration was necessary to facilitate a comparison of two teaching programs which differ with regard to the degree of implicitness. Second, this study has a high ecological validity with participants who are not just learning for the sake of the experiment (DeKeyser, 2003), and French as a FL. As the majority of the experiments in this field so far have been conducted with English as a FL, it is important to conduct experiments with other languages as well, especially those where extramural exposure is limited. Indeed, in most existing effect studies on English, there is generally no control over informal learning at home, which makes French in the Netherlands an excellent choice for such a study as there is almost no extramural exposure normally. Despite these strengths, the findings of this explorative study should be interpreted with caution and more long-term and ecologically valid studies are needed to confirm the findings.

The current long-term classroom study suggests that general learning mechanisms like statistical learning are capable of facilitating FL acquisition through FL exposure and active use, if given sufficient time. It seems that a predominantly implicit, DUB FL teaching program might be as effective as a predominantly explicit, SB approach in terms of achieving morphological accuracy and even more effective in achieving complexity and fluency. Apparently, students do not need a great deal of explicit instruction to become accurate in writing; therefore, it might be much more useful to spend time in class on oral skills, which are seriously neglected in most FL French classes in the Netherlands (Michel et al., 2021). As another study on the same students has shown (submitted), the DUB students outperform the SB students significantly with a very large effect size (d = 2.05) on speaking skills.

Notes

1 Closed-type exercises, in which students have to select from a distinct set of pre-defined responses, often appear in SB programs and focus on one specific grammatical rule, for instance the right tense (*Hier il … son depart* (*a annoncé/annonçait*))

2 Translation tasks are often used in a SB program and focus on syntax and on differences between the L1 and the L2 (*He gives* ***me*** *all his money* = *Il* ***me*** *donne tout son argent*).

References

Ågren, M., Granfeldt, J., & Schlyter, S. (2012). The growth of complexity and accuracy in L2 French. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 95–119). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.05agr

Andringa, S., De Glopper, K., & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, 61(3), 868–903. http://dx.doi.org/10.1111/j.1467-9922.2010.00623.x

Arnott, S. (2011). Exploring the dynamic relationship between the Accelerative Integrated Method (AIM) and the core French teachers who use it: Why agency and experience matter. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, 14(2), 156–176. https://journals.lib.unb.ca/index.php/CJAL/article/view/19862

Bakker, T., Beijer, S., Brouwer, J., Heringa, J., & Morice, I. (2005). *Grandes Lignes* (3rd ed.). Noordhoff.

Barbier, M.L. (1997). *Rédaction de texte en langue première et en langue seconde*. Aix-Marseille 1.

Bartning, I., & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French language studies*, 14(3), 281–299. http://dx.doi.org/10.1017/S0959269504001802

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *Tesol Quarterly*, 45(1), 5–35.

Breek, L., Polinder, K., Groenen, G., & Van der Drift, M. (2003). *Libre Service* (2nd ed.). ThiemeMeulenhoff.

Chomsky, N. (2009). *Syntactic structures*. De Gruyter Mouton. http://dx.doi.org/10.1515/9783110218329

Cobb, T. *Compleat Web VP v.2* [computer program]. Accessed 15 February 2018 at https://www.lextutor.ca/vp/comp/

Cohen, J. (1988). The effect size index: d. *Statistical power analysis for the behavioral sciences*, 2, 284–288.

De Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a dynamic systems theory perspective. In M.H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development* (pp. 5–24). John Benjamins. http://dx.doi.org/10.1075/lllt.29.01deb

DeKeyser, R.M. (2003). Implicit and explicit learning. In C.J. Doughty & M.H. Long, *The handbook of second language acquisition* (pp. 313–348). Blackwell Publishing Ltd. http://dx.doi.org/10.1002/9780470756492

DeKeyser, R.M. & Botano, G.P. (2019). Current research on instructed second language learning: A bird's eye view. In R.M. DeKeyser & G.P. Botana, (Eds.), *Doing SLA research with implications for the classroom: Reconciling methodological demands and pedagogical applicability* (pp. 1–8). John Benjamins. http://dx.doi.org/10.1075/lllt.52

Dönszelmann, S. (2019). *Doeltaal-Leertaal. Didactiek, professionalisering en leereffecten*. Parthenon.

Ellis, N.C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The modern language journal*, 92(2), 232–249. http://dx.doi.org/10.1111/j.1540-4781.2008.00716.x

Ellis, N.C., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 75–93). Routledge. http://dx.doi.org/10.4324/9780203628942

Fayol, M. (1997). *Des idées au texte. Psychologie cognitive de la production verbale, orale et écrite*. PUF. http://dx.doi.org/10.3406/prati.2000.2418

Gombert, W., Keijzer, M. & Verspoor, M. (submitted). *Exposure and chunks in L2 French writing*.

Granfeldt, J., & Ågren, M. (2014). SLA developmental stages and teachers' assessment of written French: Exploring Direkt Profil as a diagnostic assessment tool. *Language Testing*, 31(3), 285–305. https://doi.org/10.1177/0265532214526178

Granfeldt, J., Nugues, P., Ågren, M., Thulin, J., Persson, E., & Schlyter, S. (2006, May). CEFLE and Direkt Profil: A new computer learner corpus in French L2 and a system for grammatical profiling. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation* (pp. 565–570).

Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 247–276). John Benjamins Publishing Company. http://dx.doi.org/10.1075/lllt.32

Halliday, M.A. (1970). Language structure and language function. In J. Lyons (Ed.). *New horizons in linguistics* (pp. 140–165). Penguin.

Hayes, J.R., & Flower, L. (1980). Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg (Eds.). *Cognitive processes in writing* (pp. 3–30). Erlbaum.

Hermans-Nymark, L.D. (2006). *English in the EFL classroom: Why not? Classroom discourse patterns and teachers' beliefs*. Radboud Universiteit Nijmegen.

Hulstijn, J.H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 25–46). John Benjamins. https://doi.org/10.1075/sibil.48.02hul

Hymes, D. (1972). On communicative competence. *Sociolinguistics, 269293*, 269–293.

Klein, W., Dietrich, R., & Noyau, C. (1995). *The acquisition of temporality in a second language*. John Benjamins.

Langacker, R.W. (2000). *Grammar and conceptualization*. Walter de Gruyter. http://dx.doi.org/10.1515/9783110800524

Lightbown, P.M., & Spada, N. (2021). *How Languages Are Learned* (5th ed.). Oxford university press.

Long, M.H. (2000). Focus on form in task-based language teaching. In R.D. Lambert & E. Ahohamy, *Language policy and pedagogy: Essays in honor of A. Ronald Walton* (pp. 179–192). John Benjamins. http://dx.doi.org/10.1075/z.96.11lon

Maxwell, W. (2001). *Evaluating the effectiveness of the accelerative integrated method for teaching French as a second language*. University of London Institute.

Michel, M., Vidon, C., De Graaff, R., & Lowie, W. (2021). Language learning beyond English in the Netherlands: A fragile future?. *European Journal of Applied Linguistics*, 9(1), 159–182. http://dx.doi.org/10.1515/eujal-2020-0020

Norris, J.M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, 50(3), 417–528.

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578. http://dx.doi.org/10.1093/applin/amp044

Oh, S. (2006). *Investigating the relationship between fluency measures and second language writing placement test decisions*. University of Hawaii.

Onnis, L. (2012). The potential contribution of statistical learning to second language acquisition. *Statistical learning and language acquisition*, 203–235. http://dx.doi.org/10.1515/9781934078242.203

Piggott, L., Tribushinina, E., & De Graaff, R. (2020). The Icing On the Cake? Effects of Explicit Form-Focused Instruction after Two Years of Implicit EFL Learning. In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer & R. Steinkrauss (Eds.), *Usage-Based Dynamics in Second Language Development* (pp. 249–270). Blue Ridge Summit. https://doi.org/10.21832/9781788925259-014

Plonsky, L., & Oswald, F.L. (2014). How big is "big"? interpreting effect sizes in L2 research: Effect sizes in L2 research. *Language Learning*, 64(4), 878–912. https://doi.org/10.1111/lang.12079

Popma, G. (1997). Mag het ietsje minder zijn? Het gebruik van Nederlands in de Engelse les. *Levende Talen Magazine*, 84(517), 70–73. https://lt-tijdschriften.nl/ojs/index.php/ltm/article/view/1087/1056

Rousse-Malpat, A. (2012). *Effectiveness of focus on form versus focus on meaning* [Master's thesis, University of Groningen]. https://research.rug.nl/en/publications/

Rousse-Malpat, A.L.M.J. (2019). *Effectiveness of explicit vs. implicit L2 instruction: A longitudinal classroom study on oral and written skills*. Rijksuniversiteit Groningen.

Rousse-Malpat, A., & Verspoor, M. (2012). Measuring effectiveness in focus on form versus focus on meaning. *Dutch Journal of Applied Linguistics*, 1(2), 263–276. http://dx.doi.org/10.1075/dujal.1.2.07rou

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language learning*, 60(2), 263–308

Tomasello, M. (2003). *Constructing a language*. Harvard University Press. http://dx.doi.org/10.2307/j.ctv26070v8

Tyler, A., & Ortega, L. (2018). Usage-inspired L2 instruction. An emergent, researched pedagogy. In A. Tyler, L. Ortega, M. Uno, & H. Park (Eds.), *Usage-inspired L2 Instruction. Researched pedagogy* (pp. 3–26). John Benjamins. https://doi.org/10.1075/lllt.49.01tyl

Van Compernolle, R.A. (2015). *Interaction and second language development*. John Benjamins Publishing Company. http://dx.doi.org/10.1075/lllt.44

Van Hout, R., & Vermeer, A. (2007). Comparing measures of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 93–115). Cambridge University Press.

Verspoor, M.H. (2017). *Cognitive linguistics and its applications to second language teaching*. Springer. https://doi.org/10.1007/978-0-387-30424-3_140

Verspoor, M., & Schmitt, N. (2013). Language and the lexicon in SLA. *The Routledge encyclopedia of SLA* (pp. 353–360). Routledge.

Verspoor, M., Schmid, M.S., & Xu, X. (2012). A dynamic usage-based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239–263. http://dx.doi.org/10.1016/j.jslw.2012.03.007

West, L., & Verspoor, M. (2016). An impression of FL teaching approaches in the Netherlands. *Levende Talen Tijdschrift*, 17(4), 26–36. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/1631/1239

Widdowson, H.G. (1978). *Teaching language as communication*. Oxford University Press.

Yoon, H.J. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, 66, 130–141. http://dx.doi.org/10.1016/j.system.2017.03.007

Appendix: Rubric based on benchmarked texts for the assessment of writing skills

Level 4 (high):

- Good use of complex subordinate clauses
- Choice of words is varied and good
- Not many mistakes that might annoy
- Message is clear
- Good collocations
- Text is well structured. Good use of connecting words
- Good syntax
- Many correct sentences.

Level 3:

- Longer, complex sentences are used but short ones as well.
- Tense use predominantly correct
- Verbs are often well conjugated
- Message is mostly clear. Sometimes re-reading is needed.
- Word use is varied and mostly correct
- Sometimes a lack of coherence necessitates re-reading and analysis
- Grammatical errors occur but do not annoy.

Level 2:

- Basic vocabulary
- Errors in tense use are frequent
- Numerous grammatical errors occur
- Verbs are often conjugated poorly or not at all
- Syntax is not very good. Requires rereading the text again.
- Text is not always coherent.

Level 1 (low):

- A lot of grammatical errors
- Longer sentences are not coherent and often not quite clear.
- No text cohererence. Re-reading is needed for comprehension.
- Few verbs used
- Hardly any correct sentence
- Choice of words is poor and regularly causes a lack of comprehension.
- Numerous errors in verb use (conjugation and tenses)
- "Dutch syntax"