System 94 (2020) 102333

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

Improving reference cohesion in young EFL learners' collaboratively written narratives: Is there a role for reformulation?

Yvette Coyle*, Pedro Antonio Férez Mora, Juan Solís Becerra

Faculty of Education, University of Murcia, Spain

ARTICLE INFO

Article history:
Received 4 February 2020
Received in revised form 23 July 2020
Accepted 27 July 2020
Available online 31 July 2020

Keywords:
Written corrective feedback
Reformulation
Young second language learners
Narrative writing
Reference cohesion

ABSTRACT

This article reports an empirical study in which we explore the role played by reformulation as a written corrective feedback technique on the use of reference cohesion in the collaboratively written narratives of young foreign language learners. The study was carried out with 11-to-12-year-old low proficiency pairs divided into a reformulation (RG) and a control group (CG), all of whom engaged in two multi-stage writing and feedback tasks over a three-month period. Features of reference cohesion were examined in four sets of narrative picture-story texts written by all the children before and after a treatment period during which the RG pairs completed a weekly writing task and received reformulated feedback, which they compared with their original drafts. The CG pairs also performed the weekly writing task and self-edited their texts. Results showed that the feedback proved useful in significantly promoting the accurate use of pronominal reference and sequence markers in the L2 by RG pairs and in contributing to a significant decrease in the misuse of articles in comparison to the CG. These findings are discussed and conclusions drawn for the role of reformulation in fostering cohesion in children's L2 writing.

1. Introduction

On the assumption that written corrective feedback (WCF) can contribute to second language (L2) learning, the present study explores the ways in which reformulation might enhance the accurate use of cohesion in children's collaboratively written L2 narratives. Most research has examined the effects of WCF with adult learners at advanced levels of proficiency. Fewer studies have focused on how WCF influences the written output of younger, less proficient learners. At the same time, less attention has been paid to more discursive WCF types such as reformulation or models, which are directed at learners' texts as a whole and offer a wider range of lexical, syntactic and discursive alternatives than error correction. This makes reformulations suitable for targeting concerns beyond word and sentence level. In addition, single treatment, 'one-shot' studies have been the norm in WCF research despite allegations that short-term improvements in accuracy cannot be taken as confirmation of language acquisition (Truscott & Hsu, 2008). The present study aims, therefore, to advance WCF research in a number of ways. Firstly, we explore the effects of reformulation on the use of reference cohesion, an aspect which is absent

* Corresponding author. Departamento de Didáctica de la Lengua y la Literatura, Facultad de Educación, Campus de Espinardo, University of Murcia, 30100, Murcia, Spain.
E-mail addresses: ycoyle@um.es (Y. Coyle), paferez@um.es (P.A. Férez Mora), jasolis@um.es (J. Solís Becerra).

https://doi.org/10.1016/j.system.2020.102333
0346-251X/

from most WCF studies to date despite its contribution to the quality of learners' writing. Secondly, we focus specifically on learners' collaboratively written L2 texts, a decision which acknowledges the widespread use of pair assignments in language classrooms, thus enhancing the ecological validity of the study. Thirdly, our research covers a three-month period, which enables us to track ongoing changes in the participants' writing in both revised and new texts. Finally, we target a population that is underrepresented in the field of L2 writing but increasingly present in instructional contexts worldwide. In this respect, we respond to Ferris's (2010) suggestion that "Classroom-based L2 writing designs should examine whether feedback and revision on one paper can help students writers exhibit greater control of those same features in subsequent writing" (p 194), and to Wigglesworth and Storch's (2012) claim that ".to examine the impact of feedback on language learning, whether that feedback is given to pairs or individual writers, we need to provide learners with several cycles of feedback treatments of the same type over time" (p 371).

2. Literature review

2.1. Reformulation as a written corrective feedback technique

Within the "writing-to learn language" strand of second language acquisition research (Manchón, 2011), providing learners with feedback on their written texts is believed to raise awareness of the "gaps" in their L2 knowledge through processes of noticing, hypothesis testing and metalinguistic analysis, which potentially leads to more accurate writing and the development of new knowledge (Schmidt, 2001). WCF can be delivered in a variety of ways that differ as regards their explicitness (e.g. direct or indirect) and focus (e.g. comprehensive correction of all errors or selective targeting of specific linguistic features). Despite numerous investigations into the merits of different WCF types (see Bitchener & Storch, 2016 for a review), the effective provision of feedback to language learners is still an issue of concern. Reformulation is a WCF technique whereby the learner's original text is rewritten by a native speaker or a teacher, maintaining the ideational content but correcting grammatical, lexical, cohesive and stylistic problems (Levenston, 1978). Since reformulations are specifically tailored to learners' own writing, it is believed that they will process the feedback more deeply, actively seeking out solutions to incorrect language use or focusing on better alternatives included in the reformulated text.

Reformulation research has attempted to establish connections between learners' noticing and subsequent uptake of lexical and grammatical features (Adams, 2003; Qi & Lapkin, 2001; Sachs & Polio, 2007). Findings have shown that reformulation can play an instrumental role in promoting accuracy in learners' subsequent writing. Most uptake detected in this research has been lexical and grammatical in nature. Qi and Lapkin (2001) found that two adult learners at different proficiency levels noticed and incorporated in their revisions more lexical and morphosyntactic features than discursive elements. Lapkin, Swain and Smith (2002) and Swain and Lapkin (2002) reported that their adolescent learners made no improvements in textual coherence or cohesion. In contrast, Kim and Bowles (2019) found that with undergraduate ESL learners, reformulated feedback promoted deeper processing of suprasentential errors, including sentence structure and paragraph organization, than direct error correction. The failure to pay attention to textual elements of the WCF could be attributed to the tendency of lower proficiency L2 learners to rely on lexical and semantic cues and sentence-level associative patterns of meaning and form in their attempts to process meaning from L2 input (Clahsen & Felser, 2006). Yet producing coherently written texts is crucial to effective communication. In school settings in particular, successful writing requires learners to process and produce language beyond sentence level by connecting words, grammar and syntax coherently through the use of appropriate cohesive devices. It would seem, then, that if cohesion is to be facilitated by reformulated feedback, specific cohesive devices might need to be specifically targeted in order to increase learners' chances of noticing and incorporating them into their written texts. This ties in with the suggestion by Qi and Lapkin (2001) that training learners to pay attention to particular features of a reformulation might increase their likelihood of benefitting from it.

Recent research with school-aged children has shown that young EFL learners can also benefit from working with WCF. Findings from these studies suggest that the L2 learning potential of WCF seems to be influenced by a number of factors including (a) the nature of the feedback technique; (b) the length and timing of the intervention, and (c) the task modality (individual or collaborative).

Regarding different types of feedback, evidence of the positive effects of models on children's developing knowledge of lexis and their increasingly target-like use of morphological elements in narrative writing over a five-month period was reported by Coyle, Cánovas Guirao and Roca de Larios (2018), who also found that repeated exposure to reformulations led to a gradual increase in young EFL learners' use of more effective problem-solving strategies in their co-written picture-story narratives (García Hernández, Roca de Larios, & Coyle, 2017). In another comparative study of different feedback types, error correction (EC) and models were found to be equally useful for promoting children's attention to lexis. However, EC led to greater improvements in the acceptability and comprehensibility of young EFL learners' written texts (Coyle & Roca de Larios, 2014). We suggested that the perceptual saliency of the teacher's corrections in the EC condition, may have facilitated noticing and uptake more easily than the model which required the children to actively seek out differences with their own texts. An advantage for EC over implicit WCF (underlining) was also noted by Simard, Guenette and Bergeron (2015), whose high school ESL learners reported that implicit feedback caused greater misunderstandings and erroneous interpretations when they failed to understand the source of the error.

There is growing evidence to suggest, then, that error saliency may play a key role in promoting noticing and uptake (Sachs & Polio, 2007). In this respect, reformulated feedback in which the targeted feature(s) are explicitly highlighted, might be an

appropriate technique to use with young learners. The use of a textually enhanced reformulation (as in the present study) also aligns with the notion of input enhancement proposed by Sharwood Smith (1993), by which increasing the perceptual saliency of certain features in the input (through boldfacing, for example) can improve their noticeability and chances of being acquired. As an enhanced version of their own narratives, reformulations are also more likely to engage children affectively than explicit EC, which can often be discouraging for younger learners. Reformulation also seems more suitable than editing codes or metalinguistic explanations, which low proficiency learners have difficulty in understanding (Gorman & Ellis, 2019; Simard et al., 2015).

Regarding the length and timing of feedback interventions, most studies have generally involved a typical 'one-shot' design (initial draft, feedback provision and text revision or rewriting), carried out over a few weeks (e.g. García Mayo & Loidi Labandibar, 2017). Longitudinal designs in which learners receive feedback repeatedly over several weeks or months are less frequent. Yet evidence obtained from learners' ongoing experience with feedback, as proposed by Wigglesworth and Storch (2012), and attested by our studies of reformulation and models reported above, may provide better indication of progressive changes in the quality of cohesion in children's narrative writing than the analysis of a single text. This is particularly important when dealing with children for whom language learning in input-limited FL classrooms is a slow and gradual process (Muñoz, 2006). The timing of feedback interventions might also influence its language learning potential. This would explain, in part, why Gorman and Ellis (2019) failed to identify improvements over a five-week period in the accurate use of the present perfect tense by young ESL learners. It is possible that the brevity of the intervention, in which learners received 8-minute metalinguistic explanations on the targeted verb tenses or silently looked over direct feedback for the same length of time, was insufficient to make a difference to the children's knowledge of a complex grammatical feature. As noted by Sachs and Polio (2007), WCF processing, especially when it involves search and/or interpretation processes, requires time for thought.

Task modality may also play an instrumental role in the language learning potential afforded by WCF. Collaborative rather than individual writing may provide children with useful opportunities to give and receive feedback on their linguistic hypotheses and to pool their limited knowledge in the co-construction of a jointly negotiated text. This idea is in line with sociocultural perspectives on the role of collaborative dialogue in L2 learning, and with cognitive views of writing as a problem-solving activity. It is also supported by empirical research comparing collaborative and individual writers, which has generally found jointly produced texts to be superior in accuracy (Fernández-Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007). Likewise, studies of collaborative WCF processing with both adults, adolescents and children have suggested that joint deliberation on language use can alleviate the complex cognitive demands involved in responding to feedback (Coyle et al., 2018; Storch & Wigglesworth 2010; Swain & Lapkin, 2002).

2.2. Research into cohesion in children's L1 and L2 speech and writing

The lack of attention to children's development of cohesion in their L2 narrative writing is surprising given the popularity of storytelling activities in primary school classrooms. In telling stories, children need to learn how to tie together ideas across sentences and clauses through the appropriate use of linguistic devices such as connectives, reference or word overlap. Cohesion thus reveals itself as linguistic 'glue' (Celce-Murcia, 1991) that enables learners to transcend sentence-level concerns to create a well-organized text. Cohesive devices have been traditionally classified into five categories: lexical, reference, substitution, ellipsis and conjunction (Halliday & Hasan, 1976). Early research comparing the development of narrative discourse in children and adults with different first languages revealed broad developmental patterns in their use of cohesive devices (Berman & Slobin, 1994). Hickmann (2003) claimed that discourse functions for a variety of forms, including noun and pronoun systems, temporal-aspectual features or personal and spatial reference, often absent from children's cohesive repertoire before the ages of 6 or 7, increased as they matured cognitively and linguistically. Fitzgerald and Spiegel (1986), Cameron, Lee, Webster, Munro, Hunt and Linton (1995) and Zarnowski (1983) reported lexical cohesion as the most frequently used tie in monolingual children's writing, followed by reference and conjunction. Crowhurst (1987) also noted a developmental pattern in the use of cohesion by older children as a function of textual genre. Lexical repetition, recurrent in younger learners, decreased across grade levels, while pronouns, demonstratives and the definite article, appeared more frequently in the children's narrative writing than in their argumentative texts.

Studies of children's use of cohesion in L2 settings have focused overwhelmingly on oral skills. Álvarez (2006) found that the oral narrative ability of young EFL learners (8 and 11-years-old) proceeded through nine stages, beginning with word, phrase and clause-level syntax before textual organization emerged. Cohesion was limited initially to the use of lexical ties, mainly noun repetition, and the additive conjunction *and,* until referential ties for entity, spatial, and temporal deixis appeared. Once reference was consolidated, tenses, aspectual differences and subordinate conjunctive connectors were widely introduced. A second group of studies examined cohesion in children's oral narratives by comparing their L1 and the L2 storytelling skills. Based mainly on the examination of conjunction, Goto Butler, Liu and Kim (2017) and Minami (2011) revealed that Chinese L1 and EFL learners and Japanese-English bilinguals used more connectives the higher their grade level.

This research suggests that children's use of cohesive devices in their L2 oral narratives evolves gradually over time in line with their cognitive maturity and developing L2 knowledge. This raises an important empirical question regarding whether or not the collaborative comparison of reformulations with their own texts might push L2 learners beyond the constraints imposed by their limited L2 knowledge to improve reference cohesion in their written performance. Accordingly, the following research question was formulated: Are there any differences in the use and misuse of reference cohesion in the

narrative texts of young EFL learners who were provided with reformulated feedback and who engaged in a comparison of the reformulation with their own texts, and those who did not?

3. Method

3.1. Participants

The participants in the study were 28 young EFL primary school learners aged between 11 and 12 years old selected from different classes taught by the same teacher. The children had been learning English for eight years and had received on average 2 h of instruction per week, totalling around 490 h of exposure. In their EFL classes, the children followed a textbook, which emphasized explicit grammar and vocabulary teaching. Instruction, therefore, was traditional and mostly dedicated to the explanation and practice of language forms. Reading and writing were prioritized over oral skills so that opportunities for meaningful interaction in the L2 were rare. Consequently, the overall proficiency level of the children was low.

Fourteen sets of pairs were formed and randomly assigned to either a reformulation group (RG) or a control group (CG). Each group was composed of seven pairs chosen for their familiarity with each other and willingness to work together on the basis of the teacher's recommendations regarding their degree of mutuality and equality (Storch, 2002). This was considered important in order to avoid a potential imbalance in each child's individual contribution to the writing and WCF tasks.

3.2. Data collection

Data was collected over a period of twelve weeks. The procedure consisted of two cycles of a multi-stage task involving i) an initial writing stage, ii) a feedback comparison stage, and iii) the rewriting of the initial draft. In both cycles, the procedures followed were identical. At the writing stage, all the pairs were asked to jointly compose a short narrative text based on a six-frame picture prompt designed by the children's teacher. In the first cycle, the picture prompt depicted the story of a witch whose cat is turned into a bat after she places a spell on it, and in the second, a scientist who turns into a cat after swallowing a magic potion. No indications were given regarding the use of specific lexis or verb tenses when writing their texts and the children did not have access to dictionaries. One week later, at the feedback comparison stage, pairs in the RG were given their initial draft and a typewritten, reformulated version of their story and asked to compare both texts, identifying and discussing any differences they noticed between both versions. The CG pairs were also given their original texts and asked to self-edit them together. Finally, at the rewriting stage, the following week, the pairs from both groups were given the same picture prompts again and asked to rewrite their original stories. The learners did not have access to their original stories during the re-write. Following Hanaoka (2007), it was thought that rewriting (rather than revising) their original drafts might increase their chances of improving and expanding on their initial narrative. Time was not limited (Sachs & Polio, 2007) although most pairs spent an average of 10 minutes per task. The time lapse between task 2 and 3 was considered necessary to avoid the potential effects of memorization of the teachers' feedback on the children's rewritten texts (Storch & Wigglesworth, 2010). The dialogues of all the pairs were recorded during task performance in both cycles. The children were allowed to speak in Spanish while performing the tasks so that they could communicate their ideas more effectively.

The RG texts were reformulated maintaining the pairs' original ideas but specifically highlighting in bold personal pronouns, noun phrases with the definite or indefinite article and sequence markers (Table 1). The introduction of the first two features was intended to provide the children with evidence of a balanced form of character deixis, since young learners have been found to overuse the definite article (Zdorenko & Paradis, 2012). The sequence markers—*first*, *after that* and *finally*—were also included to raise learners' awareness of the basic sequence of a simple narrative (Hickmann, 2003; Álvarez, 2006). *First* and *finally* were introduced in all the narratives once, while the incorporation of *after that* varied according to the length of the children's writing. The maximum number was two per text. Grammatical, lexical and spelling mistakes were also corrected. The children were instructed to focus on the reformulation as a whole but to pay special attention to the bolded features (Qi & Lapkin, 2001).

Table 1
Examples of reformulations.

	Learners' output	Reformulation
Personal pronouns	The wich have a idea. The wich transform the cat in bat.	The witch has an idea. **She** transforms the cat into a bat.
Definite/ indefinite articles	*The* wicht eating *the* sanwich	**A** witch is eating **a** sandwich
Sequence markers	The cientific drink experiment and transform a cat. The dog see a cat and the dog kill the cat.	The scientist drinks his experiment and, **after that**, becomes a cat. **Finally,** the dog kills the cat.

3.2.1. Treatment period

Between both cycles, a six-week treatment period was implemented differently in both groups. The six weekly sessions were held separately for each group outside their classrooms during the lunch break in a quiet room in the school. In the RG, each week the children's original draft was handed out together with the reformulated version and the pairs were asked to discuss any differences between them. They were then given a different picture prompt and asked to write a new narrative text, which was collected and reformulated before the following session. In total, the children wrote six narratives and received six reformulations. The picture story prompts used during the treatment period were taken from two school textbooks which had been recently published at that time (*Cool Kids* and *Explorers*). Although slightly different in terms of plot, they were comparable to those used during the data collection in both cycles. Each story was made up of six frames without any written text, the lexical and grammatical content (everyday objects, animals, food and drink, action verbs in present tense, etc.) was considered appropriate for the children's proficiency level, and the end of each story generally involved the use of some kind of magic by the characters. The first story dealt with a genie in a lamp who used his magic to make two children smaller. The second text was a fantasy version of the three little pigs. The last four stories dealt with a magic pencil that turned drawings into real objects and how a group of children used the pencil to solve some problems. In the CG, the children also wrote six co-authored narratives using the same weekly prompts as the RG and then self-edited their texts so that they represented their best possible version. These texts were also collected at the end of the session. To ensure that the children in the CG received the same opportunities to benefit from the impact of the intervention as their peers, delayed sessions with reformulations were planned for the later in the school year once the study was completed.

3.3. Data analysis

The data consisted of the pairs' written narratives, totalling 56 written texts across four time periods, as well as transcriptions of the recorded dialogues while writing and, in the case of the RG, while discussing the feedback. The time periods corresponded to the initial writing (T1 and T3) and rewriting stages (T2 and T4) of both cycles. In order to identify improvement in the use of cohesive devices over time, Halliday and Hasan's (1976) model of reference cohesion was employed. The entire data set was coded for three referential features: personal pronouns, articles and temporal sequence markers by the three researchers, who were also experienced teachers, working independently. Cohesive features and errors, including instances when a feature should have been used, were identified in the pairs' texts. The mean frequencies for the correct and incorrect use of the three cohesive devices were computed for both groups across the four time periods. Interrater reliability at 92% was calculated by simple percentage agreement. Given the brevity of the children's narratives (between 40 and 80 words), ratios were adjusted by counting the total number of words in the four texts, dividing the number of cohesive ties/errors identified by the total number of words and multiplying by 10, following Van Beuningen, De Jong and Kuiken (2012). For the quantitative analysis, given the limited sample size and non-normal distribution of the data, as shown by a Shapiro Wilk test, non-parametric statistics were run on the data. Mann-Whitney tests with a Bonferroni correction were used to check for intergroup differences in the use and misuse of cohesive reference devices across the four time periods. A Friedman test with a post hoc Bonferroni correction was also carried out to check for intragroup differences.

4. Results

The results showed that the RG pairs significantly outperformed the CG over time in the accurate use of two of the targeted cohesive ties: personal pronouns and sequence markers (Table 2). Improvements in pronominal use were apparent in the RG pairs' revised texts in the first WCF cycle immediately after receiving reformulated feedback (T2) and were maintained over time after the treatment period when writing a new text (T3) and its revision (T4). On the contrary, none of the CG pairs used pronouns accurately at any of the four time periods (as reflected by the zeros in Table 2). Significant differences in favour of the RG were also found for sequence markers at T4. As for the misuse of reference cohesion, the CG consistently made more errors than the RG pairs with both the pronominal and article systems, significantly so at T4 for pronouns and at T3 and T4 for articles. These results suggest a positive impact for reformulated feedback on the children's use of cohesive ties. These findings are discussed below in greater detail and illustrative examples from the children's texts and dialogues are provided where appropriate.

4.1. Use and misuse of pronominal reference

At T1, the only accurate form of character reference deployed by all pairs were noun phrases comprising a definite article plus a noun. Pairs in the RG began using personal pronouns to statistically significant proportions at T2 immediately after WCF provision ($z = -2.972$; $p = .003$), and continued to do so when writing a new text at T3 ($z = -3.343$; $p = .001$) and its revision at T4 ($z = -3.343$; $p = .001$) (Table 2). The effect sizes of these incorporations were large for T2 ($d = -1.38$), T3 ($d = -1.91$) and T4 ($d = -2.51$). Accurate use of personal pronouns was not reported at any of the four writing stages by the CG pairs, who relied solely on noun phrases.

From an intragroup perspective, the use of personal pronouns increased significantly in the RG, with large effect sizes at T3 ($z = -2.643$, $p = .001$, $d = -2.6$) and T4 ($z = -1.929$, $p = .031$, $d = -1.9$) with respect to the initial measure at T1 (Table 3). Table 4 shows that personal pronouns were not used in any of the four time periods by the CG.

Table 2
Intergroup comparison of the use and misuse of referential cohesion.

		Reformulation Group (n = 7 pairs)		Control Group (n = 7 pairs)		*p* value	Cohen's *d*
		Mean	Sd	Mean	Sd		
Personal pronouns (accurate use)	T1	0.00	0.00	0.00	0.00	1.00	–
	T2	0.39	0.28	0.00	0.00	0.003**	−1.38
	T3	0.65	0.28	0.00	0.00	0.001**	−2.51
	T4	0.46	0.25	0.00	0.00	0.001**	−1.91
Articles (accurate use)	T1	1.17	0.63	0.87	0.41	0.225	−0.41
	T2	1.30	0.53	0.96	0.71	0.406	−0.34
	T3	0.87	0.61	1.35	0.35	0.277	0.51
	T4	1.15	0.70	1.17	0.68	0.898	−0.35
Sequence markers (accurate use)	T1	0.00	0.00	0.00	0.00	1.00	–
	T2	0.04	0.07	0.00	0.00	0.142	−0.51
	T3	0.30	0.30	0.02	0.05	0.063	−0.81
	T4	0.50	0.20	0.00	0.00	0.001**	−2.61
Personal pronouns (inaccurate use)	T1	0.16	0.33	0.25	0.36	0.710	0.23
	T2	0.00	0.00	0.26	0.35	0.062	−1.90
	T3	0.00	0.00	0.19	0.32	0.142	−1.50
	T4	0.00	0.00	0.37	0.35	0.025*	−2.20
Articles (inaccurate use)	T1	0.50	0.56	0.99	0.38	0.109	0.23
	T2	0.36	0.32	0.56	0.17	0.224	0.24
	T3	0.32	0.09	0.77	0.42	0.025*	2.27
	T4	0.15	0.25	0.68	0.35	0.008**	1.60

T: Time; *$p < .025$; **$p < .010$.

Table 3
Intragroup comparison for the RG of the use and misuse of referential cohesion.

		T1	T2	T3	T4	X^2	*p*
Personal pronouns (accurate use)	M	0.00 0.00[bc]	0.39 0.28	0.65 0.28[b]	0.46 0.25[c]	16.391	0.001
	Sd						
Articles (accurate use)	M	1.17 0.63	1.30 0.53	0.87 0.61	1.15 0.70	2.304	0.512
	Sd						
Sequence markers (accurate use)	M	0.00 0.00[c]	0.04 0.07[e]	0.30 0.30	0.50 0.20[ce]	16.667	0.001
	Sd						
Personal pronouns (inaccurate use)	M	0.16 0.33	0.00 0.00	0.00 0.00	0.00 0.00	6.000	0.112
	Sd						
Articles (inaccurate use)	M	0.50 0.56	0.36 0.32	0.32 0.09	0.15 0.25	2.169	0.538
	Sd						

Note. a: significant differences between T1-T2; b: significant differences between T1-T3; c: significant differences between T1-T4; d: significant differences between T2-T3; e: significant differences between T2-T4; f: significant differences between T3-T4; $p < .05$.

Table 4
Intragroup comparison for the CG of the use and misuse of referential cohesion.

		T1	T2	T3	T4	X^2	*p*
Personal pronouns (accurate use)	M	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0	1.000
	Sd						
Articles (accurate use)	M	0.87 0.41	0.96 0.71	1.35 0.35	1.17 0.68	2.829	0.419
	Sd						
Sequence markers (accurate use)	M	0.00 0.00	0.00 0.00	0.02 0.05	0.00 0.00	3.000	0.392
	Sd						
Personal pronouns (inaccurate use)	M	0.25 0.36	0.26 0.35	0.19 0.32	0.37 0.35	1.432	0.698
	Sd						
Articles (inaccurate use)	M	0.99 0.38	0.56 0.17	0.77 0.42	0.68 0.35	5.057	0.168
	Sd						

The exclusive use of noun phrases for character referencing in the texts written by the CG pairs at T1 and T2 conveys the plot as a collection of disjointed episodes (example 1). The hybrid deployment of character referencing prompted by the reformulations, as illustrated in the narratives of RG pairs at T2 (example 2), suggests that improvement in their L2 writing performance does not correspond to the total number of specific features used, which is higher at T1, but to the variety of features supplied, which from T2 onwards combined both noun phrases and personal pronouns. The combined approach to

character referencing adopted by the RG gives the impression of better quality narratives (even if there are other language problems) which are more fluid and less repetitive.

Example 1

The wicht eating the sanwich. *The* wicht eating the sanwich and *the* cat drinking the orange juice. *The* wicht looking the orange juice and *the* cat eating the sanwich […] (CG, T1, pair 11).

The wicht eating the sanwich. *The* wicht eating the sanwich and *the* cat drinking the orange juice. *The* wicht looking the orange juice and *the* cat eating the sanwich […] (CG, T2, pair 11).

Example 2

1) *The* wich have dinner at nine o'clock[1].
2) *The* cat drink the juice the wich.
3) *The* which look at juice, *the* cat eat the sanwich at half past nine.
4) *The* wich look the cat.
5) *The* wich hate the cat.
6) The bat love the cat. *The* wich eat de sanwich and juice, ten o'clock (RG, T1, pair 8).

A wich eating a sanwich, a cat look the juice. *The* cat drink the juice and *she* hungry eating a sanwich. *The* wich look the juice and cat eating the sanwich. *The wich* see the cat and *she* have a idea. *She* transforms the cat into a bat (RG, T2, pair 8).

Evidence from this RG pair's discussion[2] both during analysis of the feedback (excerpt 1) and while composing their revised version (excerpt 2), suggests that the reformulation may have helped raise their awareness of the pronoun 'she', which they incorporated into their rewritten narrative at T2, albeit in a different place to where it was originally noticed.

Excerpt 1 (RG, feedback comparison stage, pair 8)

A: Underline that. Here they've written 'but *she* doesn't see it'.
B: Yes (reading) 'but she doesn't see it'. And we didn't write that. Ok.
A: Underline it.

Excerpt 2 (RG, T2, pair 8)

A: 'And she … and the wich … '
B: '*She*', because we are talking about her.
A: 'And she … eh … have …. .a idea' (T2)

The misuse of personal pronouns by pairs in both groups generally consisted of using a pronoun to refer to an entity which was mentioned in the narrative for the first time (example 3). In the CG, as compared to the RG, there was a significant increase with a large effect size for this type of error at T4 ($z = -2.241$, $p = .025$, $d = -2.20 \pm 1.60$) (Table 2). The intragroup analysis of the RG shows that personal pronoun errors occurred only at T1 and were no longer reported after that (Table 3), whereas in the CG there was a similar error ratio across the four time periods (Table 4).

Example 3

She is angry […] (sentence initial use of third person pronoun) (CG, T1, pair 13).

4.2. Use and misuse of articles

The intergroup comparison of the children's accurate use of articles was statistically non-significant at all times (Table 2). The same was true for intragroup comparisons (Tables 3 and 4). This may have been accentuated by the varying length of the children's texts and by the incorporation of pronominal reference by the RG pairs which often resulted in fewer definite articles being used (see example 2 above).

Statistically significant intergroup differences in the misuse of articles appeared at T3 ($z = -2.236$, $p = .025$) and at T4 ($z = -2.660$, $p = .008$) when the CG continued to make frequent errors. The size of these effects were large at both T3 ($d = 2.27 \pm 1.43$) and T4 ($d = 1.60 \pm .91$) (Table 3). Articles were misused in two main ways. The most frequent cohesive error occurred when the definite article was used inappropriately the first time a character was introduced in the narrative (example 4). Further errors stemmed from children using the indefinite article to refer to entities that had already been introduced (example 5).

Example 4

The doctor make un líquido [a liquid] and his dog are slipping […] (RG, T3, pair 4).

[1] It was the children themselves who decided to number their narrative

[2] The children's discussions, which were originally in Spanish, have been translated into English. They used English only when referring to their written texts or to the reformulation.

Example 5

A witch is breakfast and a cat drink a juice orange and the cat eat a sandwich. The witch look *a* cat […] (RG, T1, pair 4).

The slight decrease in these type of errors at T3 and T4 by the RG pairs suggests that during the treatment period they may have begun to notice how articles were used for initial and subsequent noun reference in the reformulations. Evidence of their emerging awareness showed up in the joint deliberations of some of the pairs while composing a new text at T3 (excerpts 3 and 4).

Excerpt 3 (RG, T3, pair 7)

A: One day … one day…
B: The cientific ….
A: The cientific … I'm not sure if that's right … Eh one day a cientific…
B: invented
A: created a pócima (potion)

Excerpt 4 (RG, T3, pair 4)

A: I'll write.
B: Científico. How do you say 'científico'?
A: Eh … Cientific. I think it's like that. (writing) The cientific.
B: No, one day, *a* cientific.

4.3. Use of sequence markers

Neither of the groups used any form of temporal markers to create a cohesive narrative at T1. However, significant differences were found in favour of the RG at T4 ($z = -3.343$, $p = .001$) with a large effect size ($d = -2.61 \pm .64$) (Table 2). Intragroup comparisons revealed that, while in the RG the use of sequence markers increased significantly with large size effects at T4 with respect to T1($z = -2.286$, $p = .006$, $d = -.2.3 \pm 1.2$) and T2 ($z = -2.000$, $p = .023$, $d = -2 \pm 1.4$) (Table 3), in the CG, sequence markers were used only minimally at T3 (Table 4).

The use of sequence markers by the RG pairs at T4 helps to create the impression of the children's narratives as interrelated and well-sequenced stories. Although this sense of structural coherence began to take shape through the combined use of personal pronouns and noun phrases for character reference, the incorporation of sequence markers amplifies the sense of textual cohesion. Example 6 shows how a RG pair's initial narrative evolves from a series of numbered sentences listing noun phrases at T1 to a temporally cemented and cohesive narrative at T4 through the incorporation of three different sequence markers.

Example 6

1) The witch eating cook and water
2) The witch lend the food with the cat
3) The cat eating his biscuit
4) The witch have got a idea
5) The witch make magic under the cat
6) The bat woman loves to the cat (RG, T1, pair 7)

One day one cientific in the laboratory make a potion and his dog sleep. *After that* the cientific drink a potion and don't like the potion. *After that* the cientific turns one cat. *Finally* the dog get up and see to the cat and the dog killer the cat (RG, T4, pair 7).

Analysis of this pair's discussion of the reformulation immediately before producing the rewritten version at T4 suggests that the feedback had successfully drawn their attention to these features.

Excerpt 5 (RG, feedback comparison stage, pair 7)

B: Eh … *Then* and *after that*, wouldn't they mean the same?
A: Yes.
B: Like afterwards or later. *After that* I think means later, right?
B: And *Then*, afterwards I think
A: Yes
B: Shall I continue?
A: Yes.
B: (reading) "*After that, he turns into a cat and the dog wakes up*"

In contrast, no such development was found in the CG narratives where a fragmentary accumulation of sentences beginning with noun phrases was maintained across writing stages, so that the text structure remained essentially the same at T1 and T4 (example 7).

Example 7

The witch eating sandwiches and have got a juice in the home. The cat drink a juice and witch eating. The cat eating sandwiches. The witch has a idea. The witch transformer at the cat in bat girl and witch is happy eating. The bat love the bat girl eating sandwiches (CG, T1, pair 10).

The cientific work potion, the dog is slep in the table. The cientific drink the potion and the dog is busy. The cientific is silly. The potion fall. The dog scared at look the cientific transformer to the cat. The dog look at the cat. The cat fall and the dog run to the cat (CG, T4, pair 10).

5. Discussion

The present study aimed to expand research on the language learning potential of reformulation in relation to the use of reference cohesion in children's L2 collaborative writing. Collectively considered, the results show that exposure to reformulated WCF had a greater effect on the accurate use of reference cohesion in children's written L2 narratives than writing practice alone, and that, over time, this led to improvement in specific referential features, particularly pronouns and sequence markers. These findings add to a growing body of research which has confirmed the positive effects of reformulation on improving the accuracy of L2 texts written by learners of different ages and proficiency levels, both individually (Qi & Lapkin, 2001; Sachs & Polio, 2007) and in collaboration (Adams, 2003; Lapkin et al., 2002; Storch & Wigglesworth, 2010; Swain & Lapkin, 2002; Wigglesworth & Storch, 2012). Previous research has shown that when learners notice differences between their text and a reformulated version, when they discuss and reflect on these differences with a partner, their engagement with the WCF is often deeper. As a result, the chances of incorporating what is noticed into their writing increases. The evidence we obtained from the RG children's dialogues provides some support for these claims, since in their joint attempts to make sense of the feedback and to combine their knowledge to produce a single narrative text, they often reflected, albeit in a very rudimentary way, on the cohesive features highlighted in the reformulations. This, in turn, may have influenced the improvements in accuracy that characterized their writing over time in comparison to the CG pairs who had no additional input other than their own limited linguistic resources. These findings suggest a positive role for collaborative writing and feedback processing in L2 writing even with younger, less proficient learners.

Additional factors which may help to explain the differences found between both groups in the use and misuse of reference cohesion include the nature of the WCF, its linguistic focus, and potential practice effects with the feedback. The usefulness of the reformulation may have been enhanced by making the cohesive devices more salient for the learners. The rationale for doing so is supported theoretically by the idea that that noticing (and uptake) may be facilitated when certain features in the input learners are exposed to are textually enhanced in some way, generally by using bold type or by colour coding (Sharwood Smith, 1993). When it comes to WCF, Sachs and Polio (2007) found that learners recalled direct error corrections more easily than a reformulation due to the visual saliency of the cross-outs, which made them easier to remember. The enhanced reformulations we gave to the children were intended to direct their attention to specific cohesive features within the familiar context of a narrative story. This may have increased the explicitness of the WCF and presumably facilitated the noticing of the targeted features without making excessive demands on the children's underdeveloped L2 knowledge. The fact that other linguistic errors (e.g. subject-verb agreement) not specifically targeted by the reformulation remained unchanged (as attested in the examples of the children's texts), despite having been corrected, would seem to provide preliminary evidence in support of the usefulness of textually enhanced WCF.

The impact of the reformulation might also have been influenced by the linguistic feature itself. Our study provides some insight on the amenability to change of specific cohesive features. Pronoun reference and temporal discourse markers proved more responsive than the article system. This finding aligns with research by García Mayo, Lázaro and Liceras (2005) who found that use of the L2 pronominal system emerged in young Basque-Spanish EFL learners around the age of 12. Before this age, younger Spanish-speaking children tend to rely on noun phrases to express pronominal reference, since transfer from the L1 leads them to treat pronouns as verbal morphemes, which delays the acquisition of the latter (Lázaro Ibarrola, 2011). For the 11-12-year-olds in the present study, for whom pronominal reference was not initially part of their linguistic repertoire, exposure to enhanced reformulated feedback may have helped accelerate this developmental process. Whether this might be true for other linguistic features would seem an issue worth exploring.

Improvements in the children's texts were also evident in their attempts to narrate the story as a sequence of events linked by connectors. The reformulations seem to have raised the children's awareness of story structure and helped them integrate connectors appropriately in their revised texts. In doing so, the learners began to treat the picture story as a sequence of events rather than as a set of isolated episodes. In this sense, the use of a narrative task, which children would have been familiar with, most likely enhanced the influence of the reformulation. The delayed appearance of connectors in the children's texts might also be attributed to a trade-off effect by which the learners focused initially on personal pronouns before incorporating sequence markers into their stories. This argument is congruent with the developmental stages outlined by Álvarez (2006) for oral performance, according to which the deployment of personal pronouns occurs just before the emergence of temporal discourse markers. In contrast to our findings, the use of temporal connectors (*then, after, when*) by the young Chinese EFL learners investigated by Goto Butler et al. (2017) was still exceptional even at the age of 11-12-years-

old, a characteristic associated with the influence of the children's L1. In Chinese, temporal adverbs (e.g. *yesterday, now, in the evening*) are used instead of tense markers to indicate temporality. A possible L1 transfer effect would explain why adverbial phrases were more frequently used in these children's oral narratives in English than connectors, which our young Spanish speakers incorporated more readily after exposure to WCF.

The accurate use of article reference did not improve significantly, although some promising evidence of a gradual reduction in article errors over time by children in the RG was noted. Even though the picture stories and reformulations provided numerous opportunities to notice and use both article functions to describe characters and objects, articles were not used consistently in the children's texts. These findings conform to universal developmental trends described for adult and child L2 learners, which suggest that learners have difficulty in using articles successfully until later stages of acquisition. Despite the greater complexity of the Spanish article system, Zdorenko and Paradis (2012) have pointed out that young Spanish-speaking L2 learners have difficulty in mastering the semantic distinctions between definite and indefinite contexts in English. Evidence from the children's dialogues confirmed their struggles to make the correct article choice, which often ended in the incorrect suppliance of the definite article in an indefinite context. On discussing the reformulated versions of their stories, some pairs did notice the use of the indefinite article to introduce the story characters and other objects for the first time. However, rather than engaging in further discussion on the nature of their misuse of the article, the children simply pointed out the difference in relation to their own text (e.g. "*we forgot 'a'"; "The witch, we put 'the', we've started wrong*"). The quality of their noticing, then, was more superficial than substantive, thus reducing the likelihood of contributing directly to writing improvement (Qi & Lapkin, 2001).

This finding does not coincide with research on adult learners (Bitchener, 2008; Shintani & Ellis, 2013) which has reported increased accuracy in the use of definite and indefinite articles after focused WCF provision. It should be noted, though, that, aside from the higher proficiency level of the learners, in these studies only one structure was specifically targeted, which would have increased the probability of its being noticed and incorporated in subsequent writing. However, in research examining more than one linguistic structure, different combinations of WCF had no lasting effects on the accurate use of the indefinite article by adult EFL learners (Shintani, Ellis & Suzuki, 2014; Suzuki, Naasaji & Sato, 2019). The authors attribute this to its lack of saliency in comparison to the more noticeable hypothetical conditional and past perfect structures, both of which contributed more to the overall meaning of the task.

The decline in the misuse of article reference by children in the RG over time could be indicative of subtle progress in the acquisition of this feature stimulated by their continued practice with the enhanced reformulations. In fact, studies of oral task repetition with younger learners have shown that performing the same or similar tasks on several occasions can impact positively on children's linguistic performance (Pinter, 2007; Shintani, 2012). The repetition of writing tasks is also believed to be useful for helping to redirect attention away from task procedure and more towards formal aspects of the language (Manchón, 2014). In this sense, the improvements identified in the RG children's narrative writing suggest there may be greater benefits for task repetition with reformulation than for writing practice alone.

6. Conclusion

The present study suggests that collaborative writing and processing of reformulations were useful in helping children to improve the cohesive structure of their L2 narratives. This finding is unique in broadening the focus of research attention beyond the analysis of purely linguistic categories (lexis, grammar, spelling) to consider how this WCF technique influences an underexplored component of school-based writing. The study also expands research on the merits of collaborative writing and feedback processing with younger learners since the improvements identified in the pairs' L2 writing offer preliminary evidence to show that even inexperienced language learners are capable of working together to encode meaning in the L2 and to use reformulations to resolve linguistic problems. The time frame of the research, carried out over a three-month period, adds a new dimension to previous 'one-shot' reformulation studies, most of which are shorter and lack a comparison group against which to compare their results. Finally, the nature of the feedback itself, which combined the personalized reformulation of the children's original texts with enhanced cohesive devices, represents a promising attempt at directing learners' attention to certain features of language. The study also suggests which cohesive features might be more readily susceptible to reformulated feedback and which might require more focused and sustained attention if learning is to be facilitated. From a methodological perspective, the combined analysis of the children's texts and their interactions during the writing and WCF sessions enabled us to account for outcomes which might otherwise have been more difficult to interpret.

Our research is not without its limitations and suggestions for future investigations are warranted. Since we did not include a final individual writing task in the design, no firm claims can be made regarding the development of reference cohesion by individual learners. Future research might include an additional measure to fully reflect the impact of collaborative processing of feedback on children's learning. Also, the inclusion of a reformulation group whose texts were not visually enhanced might provide more robust evidence to either confirm or refute our findings, as would a comparison between cohesive features and other linguistic errors that were not specifically targeted, especially those subject to developmental and maturational constraints. A larger sample of learners of different ages, proficiency levels and language backgrounds would certainly enhance the generalizability of the results, Future studies might also consider different feedback types, cohesion categories or writing genres, all of which might yield different results. This would shed light on the potential effects of a greater number of learner-internal and external variables on the use and development of cohesion in L2 writing.

Acknowledgement

This study was funded by the Spanish Ministerio de Ciencia e Innovación (research grant PID2019-104353GB-I00).

References

Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research, 7*, 347–376.
Álvarez, E. (2006). Rate and route of acquisition in EFL narrative development at different ages. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 127–155). Clevedon: Multilingual Matters.
Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*, 102–118.
Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol/Buffalo/Toronto: Multilingual Matters.
Cameron, C. A., Lee, K., Webster, S., Munro, K., Hunt, A. K., & Linton, M. J. (1995). Text cohesion in children's narrative writing. *Applied PsychoLinguistics, 16*(3), 257–269.
Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *Tesol Quarterly, 25*(3), 459–480.
Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied PsychoLinguistics, 27*(1), 3–42.
Coyle, Y., Cánovas Guirao, J, & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing, 42*, 25–43.
Coyle, Y, & Roca de Larios, J (2014). Exploring the role played by error correction and models on children's reported noticing and output production in an L2 writing task. *Studies in Second. Language Acquisition, 36*(3), 451–485.
Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English, 21*(2), 185–197.
Fernández-Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*, 40–58.
Ferris, D. (2010). Second language writing research and written corrective feedback in SLA. Intersections and practical applications. *Studies in Second Language Acquisition, 32*, 181–201.
Fitzgerald, J., & Spiegel, D. L. (1986). Textual cohesion and coherence in children's writing. *Research in the Teaching of English, 20*(3), 263–280.
García Mayo, P., Lázaro Ibarrola, A., & Liceras, J. (2005). Placeholders in the English interlanguage of bilingual (Basque/Spanish) children. *Language Learning, 55*(3), 445–489.
García Mayo, & Loidi Labandibar, U. (2017). The use of models as written corrective feedback in EFL writing. *Annual Review of Applied Linguistics, 37*, 110–127.
García Hernández, F. J., Roca de Larios, J., & Coyle, Y. (2017). Exploring the effect of reformulation on the problem-solving strategies of young EFL writers. In M. P. García Mayo (Ed.), *Learning Foreign Languages in Primary School: Research Insights* (pp. 193–222). Bristol. Multilingual Matters.
Gorman, M., & Ellis, R. (2019). The relative effects of metalinguistic explanation and direct written corrective feedback on children's grammatical accuracy in new writing. *Language Teaching for Young Learners, 1*(1), 57–81.
Goto Butler, Y., Liu, Y., & Kim, H. (2017). Narrative development in L1 and FL: A longitudinal study among young Chinese learners of English. In M. P. García Mayo (Ed.), *Learning foreign languages in primary school: Research insights. (* (pp. 49–176). Bristol: Multilingual Matters.
Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman Group.
Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 459–479.
Hickmann, M. (2003). *Children's discourse: Person, space and time across languages*. Cambridge: Cambridge University Press.
Kim, H. R., & Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think-alouds. *Tesol Quarterly*, 1–26.
Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *The Modern Language Journal, 86*, 485–507.
Lázaro Ibarrola, A. (2011). Faster and further morphosyntactic development of CLIL vs EFL Basque-Spanish bilinguals learning English in high school. *IJES, 12*(1), 79–96.
Levenston, E. A. (1978). Error analysis of free comparison: The theory and the practice. *Indian Journal of Applied Linguistics, 4*(1), 1–11.
Manchón, R. M. (2011). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning to write and writing to learn in an additional language* (pp. 61–82). Amsterdam: John Benjamins.
Manchón, R. M. (2014). The internal dimensions of tasks. In H. Byrnes, & R. Manchón (Eds.), *Task-based language learning - insights from and for L2 writing* (pp. 27–53). Amsterdam: John Benjamins.
Minami, M. (2011). *Telling stories in two languages: Multiple approaches to understanding English-Japanese bilingual children's narratives*. Charlotte, NC: Information Age Publishing.
Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning*. Clevedon, UK: Multilingual Matters.
Pinter, A. (2007). Some benefits of peer–peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research, 11*(2), 189–207.
Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*, 277–303.
Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition, 29*(1), 67–100.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165–179.
Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research, 16*(2), 253–279.
Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22(3)*, 286–306.
Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning, 64*(1), 103–131.
Simard, D., Guènette, D., & Bergeron, A. (2015). L2 learners' interpretation and understanding of written corrective feedback: Insights from their metalinguistic reflections. *Language Awareness, 24*(3), 233–254.
Storch, N. (2002). Relationships formed in dyadic interaction and opportunity for learning. *International Journal of Educational Research, 37*(3), 305–322.
Storch, N. (2005). Collaborative writing: Product, Process, and students' reflections. *Journal of Second Language Writing, 14*, 153–173.
Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 155–177). Clevedon, UK: Multilingual Matters.
Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition, 32*, 303–334.
Suzuki, W., Nassaji, H., & Sato, K. (2019). The effects of feedback explicitness and type of target structure on accuracy in revision and new pieces of writing. *System, 81*, 140–150.
Swain, M., & Lapkin, S. (2002). 'Talking- it- through': Two French immersion learners' responses to reformulation. *International Journal of Educational Research, 37*, 285–304.

Truscott, J., & Hsu, A. Y. P. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292–305.
Van, B., G, C., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning, 62*(1), 1–41.
Wigglesworth, G., & Storch, N. (2012). Feedback and writing development through collaboration: A socio-cultural approach. In R. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 69–97). Berlin: Mouton de Gruyter.
Zarnowski, M. (1983). Cohesion in student narratives: Grades four, six, and eight. *Unpublished research report (ERIC Document Reproduction Service No. ED 247 569)*, 1–35. https://files.eric.ed.gov/fulltext/ED247569.pdf.
Zdorenko, T., & Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language, 32*(1–2), 38–62.