International Journal of Educational Research Open 2 (2021) 100102

Contents lists available at ScienceDirect

International Journal of Educational Research Open

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

Languaging and learning in independent EFL study compared to collaborative pairwork

Andrew E. Sampson*

University of Porto, Departamento de Estudos Anglo-Americanos, Faculdade de Letras, Via Panorâmica s/n 4150-564, Porto, Portugal

ARTICLE INFO

Keywords:
Languaging
LREs
Engagement
SCT
Self-study
Pairwork

ABSTRACT

Covid-19 has meant an increase in the time language learners spend doing tasks alone, either at home or in socially distanced classrooms once characterised by collaborative pairwork. The present classroom study compares independent language task performance with collaborative pairwork in terms of Language-Related Episodes (LREs) and the learning of topicalised forms. A mixed-methods analysis suggests that while individuals engage in LREs to a similar extent as each learner in dyads, the additive effect of two learners means greater languaging in pairwork. While LRE resolution and cognitive engagement does not differ significantly between the two settings, test scores suggest that forms focussed on individually are more memorable than those discussed dyadically. Pedagogical recommendations are proposed for both independent study and pairwork.

1. Introduction

Covid-19 has forced many English as a Foreign Language (EFL) teachers to move their activities online, either completely or within a hybrid / blended system in which learners spend some time attending classes face-to-face and the remainder studying at home. While some of the home component may involve synchronous online learning through Zoom or similar technologies, many learners are now expected to do more language tasks alone, without the peer collaboration that characterises face-to-face communicative classrooms. Those learners who have continued in face-to-face classrooms may be socially distanced from other students and find themselves individually completing tasks that would previously have been done in pairs. This shift towards more individual language work was in fact already evident before Covid-19, with the growth of flipped classroom models (Vitta & Al-Hoorie, 2020) in which independent study is used for language presentation and controlled practice in order to free up class time for communicative practice.

This shift necessitates an examination of the learning processes occurring in individual task performance compared to collaborative pairwork, in order to assess the benefits and constraints of increased amounts of independent study. While a large body of research (Philp, Adams & Iwashita, 2013 provide an overview) has assessed the potential benefits of collaborative pairwork, a gap in the literature remains regarding the thinking and learning processes that take place when learners work independently. The present study compares individual performance with pairwork on two language tasks by analysing learners' languaging, the "process of making meaning and shaping knowledge and experience through language" (Swain, 2006: 98) rooted in a Vygotskian (1978, 1987) sociocultural framework in which language mediates cognition and learning. Languaging is observable in learners' Language-Related Episodes (LREs), instances in which "students talk about the language they are producing, question their language use, or other- or self-correct" (Swain, 1998: 70). LREs encompass a range of behaviours associated with language learning, such as noticing the gap between students' or their partners' interlanguage and a target language feature (Gass & Mackey 2007), formulating recasts and participating in metalinguistic discussions (Kim & McDonough 2011) and engaging in hypothesis testing and self-repair (Gilabert & Barón 2013).

Only a small number of studies (Swain & Lapkin 1995 and Kim, 2008 are two of the few) have examined LREs when learners perform tasks individually, with most studies interpreting LREs as collaborative events. However, Swain and Lapkin's (1995) conceptualisation of LREs, and the Vygotskian sociocultural framework in which they are rooted – specifically Vygotsky's notion of inner speech – allow for LREs to be events that can occur in individual learners' thinking, and that may be observed by way of think-aloud protocols. By examining this thinking and the LREs it contains, a better understanding can be reached regarding independent study compared to the much more extensively researched pairwork condition. In this way, the present study contributes to the discussion regarding the hybrid nature of teaching and learning and the gradual dissolution of boundaries between online and offline learner roles, and has the potential to help inform teachers' decision-making regarding how much individual and pairwork to set by

* Corresponding author.
E-mail addresses: andrewesampson@yahoo.co.uk, asampson@letras.up.pt

https://doi.org/10.1016/j.ijedro.2021.100102
Available online 8 December 2021

deepening our understanding of the possible constraints and benefits experienced in each mode.

2. Theoretical framework

This study is informed by the work of Vygotsky (1978, 1987) and Sociocultural Theory (SCT), which proposes that knowledge construction occurs collaboratively between novices and experts, and by the related concept of scaffolding (Wood, Bruner & Ross, 1976), finely-tuned support provided to aid students' development from their current to potential level within the Zone of Proximal Development (ZPD). Since learners often have different levels of expertise in different areas of language and skills, peers in dyads can provide scaffolding to mediate each other's development (Donato, 1994; Ohta, 2000, 2001; Storch, 2002, 2005).

SCT is also an appropriate framework for examining language as a mediational tool when learners perform tasks alone, given Vygotsky's concept of inner speech (1987). By using language to think through problems, humans reach new insights, construct knowledge and, potentially, achieve independent problem solving. This process can involve self-scaffolding (Holton & Clark 2006; Knouzi, Swain, Lapkin & Brooks, 2009), a "form of internalized conversation in which the student interrogates their epistemic self" (Holton & Clark 2006: 128). While learners clearly cannot impart to themselves concepts they do not know, they can self-scaffold by breaking down problems into smaller parts, starting with simpler problems first, and making optimal use of available resources (Bickhard, 2005).

3. Literature review

While very few studies have compared collaborative with individual task performance from the sociocultural perspective of learner languaging, a number of cognitively-orientated studies have compared the two settings in terms of the accuracy, complexity and fluency of learner output. Findings generally suggest beneficial effects of peer collaboration over individual task performance. A comparison of collaborative and individual text reconstruction (dictogloss) was conducted by Basterrechea and García Mayo (2013) in two contexts, Content and Language Integrated Learning (CLIL) and EFL. In EFL, there was little difference between pairs and individuals in accuracy relative to the use of the 3rd person *-s* morpheme, but in CLIL there was a statistically significant difference in favour of collaborative reconstruction. Similarly, Wigglesworth and Storch (2009) found that pairs who collaboratively wrote an argumentative essay produced statistically significantly more error-free clauses than learners writing the same essay alone, and that pairs' LREs contained evidence of collective scaffolding which contributed to greater accuracy. In (LaPierre, 1994) study, learners who had incorrectly resolved LREs collaboratively in a dictogloss later applied this "incorrect knowledge" to a post-test, thus providing evidence of the retention of collaboratively constructed knowledge. Nassaji and Tian (2010) found that pairs demonstrated greater accuracy than individuals when completing cloze and text editing tasks seeded with phrasal verbs, although no statistically significant differences were found in learning gains.

These claims for the beneficial impact of collaboration are supported by studies investigating the impact of participant numbers in groups on LRE production. In Fernández Dobao (2012, 2014) greater attention to form occurred in small groups, compared to dyads doing the same task, with more LREs and greater LRE resolution. The author claims that more learners meant more cognitive resources, and a greater possibility for episodes to be resolved correctly. Similarly, Lasito and Storch (2013) found that while adolescent Indonesian learners performing a jigsaw task in triads produced fewer LREs than pairs doing a similar task, triadic learners were better able to resolve the LREs.

It should be noted that the aforementioned studies did not observe LREs in individuals. The methodological challenge of observing LREs in independent study, compared with the relative ease of observing LREs in dialogue, is one reason for the scarcity of studies observing individual languaging. Furthermore, LREs are often assumed to be events that occur in collaborative dialogue with an interlocutor, rather than individually (Fortune, 2013). However, Swain asserts that languaging, the process in which LREs arise, is "made visible as learners *talk through with themselves or others* the meanings they have, and make sense of them" (2006: 95, my italics). In (Swain and Lapkin, 1995), for example, transcripts of think-aloud protocols produced by students of French as they wrote a composition revealed a range of LREs, indicating that learners engage in languaging even when no external feedback is available from a peer or teacher.

The only published study attempting to compare collaborative with individual LREs is Kim (2008). Her findings suggest that while Korean as a Second Language (KSL) learner dyads completing a dictogloss were able to pool their knowledge and correctly resolve most LREs, individual learners tended to leave LREs unresolved, since they had no resources to draw on other than their own knowledge, the gap in which had given rise to the LRE. Pairs showed statistically significantly higher gain scores than individuals on immediate and delayed post-tests, which suggests learning advantages for collaborative over individual task performance. However, individuals seemed reluctant to vocalise their thinking, leading Kim to suggest that having learners think aloud in L2 had created an additional cognitive demand not experienced by dyads.

While the research has generally examined LREs in terms of their quantity, linguistic focus and the correctness of resolution, few studies to date have observed the quality of learner's participation or involvement within episodes; that is, their level of engagement. Engagement has been defined as "a state of heightened attention and involvement" (Philp & Duchesne 2016: 52) that "requires energy and effort" and "drives learning" (Christenson, Reschly & Wylie, 2012: 817). Of the three types of engagement – behavioural, emotional and cognitive – described by Fredericks, Blumenfeld & Paris in their seminal 2004 paper, the type that is of key interest in the present study is cognitive engagement, which "draws on the idea of investment; it incorporates thoughtfulness and willingness to exert the effort necessary to comprehend complex ideas and master difficult skills" (Fredricks, Blumenfeld & Paris, 2004: 60).

Specific observable learner behaviours that operationalise cognitive engagement relate to self-regulated learning strategies, that is, metacognitive strategies learners use to plan, monitor and assess their thinking. Fredricks et al. (2004) draw on multiple studies to identify such strategies, which include rehearsing, summarising and elaborating information in order to remember, organize and understand (Corno & Madinach 1983; Weinstein & Mayer 1986), remaining on task and avoiding distractions (Corno, 1993; Pintrich & De Groot 1990) and creating connections between concepts and ideas (Weinstein & Mayer 1986). Further indicators of cognitive engagement include completing peer utterances and making gestures and facial expressions (Helme & Clarke 2001), comparing, asking questions and drawing inferences regarding the target language (Svalberg, 2009), and interactive support and positive attitudes (Baralt, Gurzynski-Weiss & Kim, 2016).

The relative newness of engagement as a concept in empirical language research in the classroom is evident in the small number of studies (Baralt et al., 2016 and Lambert, Philp & Nakamura, 2017; Storch, 2008 are amongst the few) that have attempted to measure engagement in LREs. The related construct of noticing (Schmidt, 1990), which relates to conscious attention to the form and meaning of language items in input – and may therefore be considered an example of cognitive engagement – has been much more extensively researched, with associations generally found between increased noticing and learning gains. In Leow (1997), for example, elaborate noticing, compared to simple noticing, led to better receptive knowledge of verbs, and more accurate productive ability as measured by a cloze text. Likewise, in Qi and Lapkin (2001)), learners performing a re-write remembered and incorporated items that had been subject to substantive noticing in a collaborative pre-task more often than the items that had been noticed perfunctorily.

In one of the first studies to employ the concept of engagement within LREs, Storch (2008) found that greater learning and consolidation of forms occurred following LREs characterised by elaborate rather than limited engagement. Elaborate engagement was operationalised as deliberation over language items, seeking and providing confirmation and explanations, and suggesting alternatives, whereas limited engagement was operationalised as learners simply stating a linguistic item without further deliberation.

Given the lack of studies comparing collaborative pairwork with individual task performance from the perspective of learner languaging or cognitive engagement within LREs, the present study aims to contribute to the literature by investigating the following research question:

How do Language-Related Episodes differ between individual task performance and pairwork in terms of i) number; ii) correctness of resolution; iii) level of cognitive engagement; and iv) learning of forms topicalised?

4. Methodology

4.1. Participants

Participants were 45 L1 Spanish adult learners (all names in the discussion below are pseudonyms) studying at a private language school in Spain. They comprised 30 learners (17 female, 13 male, average age 31.6) in 15 learner-learner dyads in face-to-face classes, and 15 learners (11 female, 4 male, average age 32.8) following the same course material independently at home. All participants had a similar level of English, having studied upper-intermediate (Common European Framework of Reference level B2) general English for the same period and achieved marks of between 80% and 90% on the same institutional progress test, taken two months prior to the study, which included multiple choice and cloze grammar and vocabulary items, reading and listening comprehension, an open-ended written composition and a speaking interview.

4.2. Tasks

In the present study all participants completed the same two tasks. The first was a language-focussed passage editing task (Task 1, Appendix 1) in which they had to correct an email to a university admissions officer written in informal language rather than a more appropriate formal register. Passage editing tasks have been shown to draw learners' attention to a range of language forms (Storch, 1997) and lead to discussions and reflections on language choices and hypothesis testing (García Mayo 2002). The passage was seeded with a total of 30 errors and inappropriacies (6 gmatical errors, 6 lexical errors, 6 errors in spelling, punctuation and capitalisation, and 12 inappropriacies in register) relating to forms studied in the course. The second task, completed two weeks later, was a meaning-focussed written composition (Task 2, Appendix 2) on the topic of banning smoking. Written compositions have been shown to effectively elicit metatalk by providing opportunities for emerging Focus on Form (Swain & Lapkin 1995).

Participants in dyads talked together to complete the tasks collaboratively, and were audio recorded. Individual participants completed the tasks alone at home, thinking aloud, and audio recorded themselves. Individual learners saw a video model of a think-aloud protocol prior to the task.

4.3. Post-test

One week after Task 1, all learners individually completed a post-test, which took the form of an isomorphic task (Appendix 3) seeded with 30 errors of the same kinds as those described above in Task 1. The use of an isomorphic post-test was based on the assumption that if learners had participated in an LRE about a form in the first task and had either learned something new or consolidated existing knowledge in the episode, they would be able to recognise and correct a similar or identical form in the post-test.

4.4. Data analysis

I transcribed learner talk in the two tasks. In the transcriptions below, pauses of one second or longer are indicated by …, and overlapping speech is indicated using indentation. I then identified LREs in the transcriptions. Following Swain (1998)), each LRE was an instance in which learners talked about the language they were producing and / or other- or self-corrected. Correction of the text was considered a form of other correction and therefore constituted an LRE. Reading out loud was not considered an LRE, since there was no talk about language; also not classed as LREs were general comments such as "this is OK" or "this is fine", without reference to any particular form, as these contain no discussion about language, only a judgement.

Within each LRE, firstly I identified whether it was resolved correctly, resolved incorrectly, or left unresolved. In dyads, I also noted who initiated and resolved the LRE.

Secondly, I coded the cognitive engagement in each LRE as elaborate or limited. Following Storch (2008), limited engagement was operationalised as instances in which a linguistic item was stated without further deliberation, including when, in student-student or student-teacher dyads, there was some phatic utterance such as "OK" or "yeah", but no further evidence of cognitive engagement. Elaborate engagement LREs, conversely, showed evidence of a metacognitive self-regulation strategy, such as elaborating on linguistic choices made (e.g. by seeking and / or providing justifications, noticing, and reflecting on forms), generating options from which to choose, creating connections (e.g. by hypothesis testing or generating rules), and attempting to go further than the requirements of the task. In dyads, I identified elaborate engagement in participant 1 only, in participant 2 only, or in both participants.

Thirdly, I identified if there was evidence of learning of forms topicalised in each LRE. Evidence of learning could take one of two forms: the first was microgenetic development (MGD), that is, occasions where, based on a qualitative analysis of the transcript alone (i.e. without consulting the post-test), there was evidence of change in one or both of the participants' language knowledge within the duration of the task. To be coded MGD, some indication of uptake within the interaction itself had to be evident, beyond a phatic response such as "Oh", in the form of a more extended response or further use of the item. The second form of evidence of learning was each participant's post-test responses, which were compared to the transcript of his or her original passage editing task to check if test items corresponded to LREs. I marked these post-tests according to the system described in Table 1.

Appendix 4 contains examples of coding decisions.

A subset of 20% of the transcripts and post-tests was randomly selected and marked by a second rater. An acceptable inter-rater reliability score of 91% was achieved using the Miles and Huberman (1994: 64) inter-rater reliability formula.

Data for the dependent variables (numbers of LREs; resolution of LREs; cognitive engagement in LREs; test scores) were tested on Q-Q plots to determine the normalcy of distributions. I performed unpaired *t*-tests (for normally distributed data) or Mann-Whitney U tests (for non-normally distributed data) to determine whether differences between the two modes (pairwork and individual) were statistically significant. Tests were two-tailed since there was no directional hypothesis, and unpaired since data for each condition came from different groups, given the study's between-subjects design.

5. Results and discussion

5.1. Number of LREs

Table 2 presents the descriptive statistics for numbers of LREs in passage editing (PE) and written composition (WC).

Table 1
System for comparing post-tests with LREs in task 1 transcripts
.

How many **post-test items** related to the LRE in task 1?0, 1 or 2?If this figure is 1 or 2, go to next step. If this figure is 0, stop.

↓

How many of the **post-test items relating to each LRE in task 1** were **attempted**?0, 1 or 2?If this figure is 1 or 2, go to next step. If this figure is 0, stop.

↓

How many of the **attempted post-test items** were resolved in **agreement** with the resolution of the LRE in task 1?0, 1 or 2?

↓

How many of the **attempted post-test items** were resolved in **disagreement** with the resolution of the LRE in task 1?0, 1 or 2?

↓

How many of the **attempted post-test items** were resolved when the LRE in task 1 had been left **unresolved**?0, 1 or 2?

↓

How many of the **attempted post-test items** were only **partially** resolved i.e. underlined or circled, but not corrected?0, 1 or 2?

Table 2
Number of LREs.

		LREs	M	SD	Kurtosis	Skew
Passage Editing	Pairs (n = 15)	406	27.1	7.9	−1.0	0.4
	Individuals (n = 15)	235	15.7	4.4	3.5	1.6
Written Composition	Pairs (n = 15)	234	15.6	7.9	0.2	0.4
	Individuals (n = 15)	129	8.6	4.1	−0.2	0.3

Table 3
Identity of initiator of LRE.

		LRE initiations	M
P1 initiates in pairs (n = 15)	Passage Editing	240	16.0
	Written Composition	97	6.5
P2 initiates in pairs (n = 15)	Passage Editing	166	11.1
	Written Composition	137	9.1
Individual initiates (n = 15)	Passage Editing	235	15.7
	Written Composition	129	8.6

Independent-samples t-tests revealed a statistically significantly higher number of LREs at the $p < .05$ level in pairs than individuals, in both PE (t (28) = 4.48, p = .00012) and WC (t (28) = 3.04, p = .0051). While this finding supports the majority of studies in the literature, which claim benefits for pairwork compared to individual performance, it should be reiterated that most of those studies compare the two modes in terms of the accuracy, complexity and fluency of output, rather than languaging.

Despite the significantly lower number of LREs in individual performance, it is important to note that the number of LREs produced by individuals was not significantly different from dyadic LREs *initiated* by participant 1 or participant 2, as revealed by independent-samples t-tests for PE (t (43) = 1.19, p = .24) and WC (t (43) = 0.56, p = .58). Table 3 presents the total number of LRE initiations, and the mean per participant, in each task:

Individuals therefore identified language problems and initiated LREs to a similar extent as each one of their dyadic counterparts – the difference is that in dyads, the sum of these initiations means there is a statistically significantly greater total number of LREs. This result supports findings from Fernández Dobao (2012, 2014) and Lasito and Storch (2013), where increasing the number of participants resulted in more linguistic resources being pooled, and significantly greater numbers of LREs. In other words, two heads appear to be better than one (Storch, 1999) in terms of the amount of languaging taking place.

5.2. Resolution of LREs

Table 4 presents the descriptive statistics for LRE resolution.

Mann-Whitney U-tests revealed no statistically significant differences between settings at the $p < .05$ level in the proportion of correctly resolved episodes (PE: U(28) = 87.5, z = 1.02, p = .31; WC: U(28) = 93.5, z = 0.77, p = .44), incorrectly resolved episodes (PE: U(28) = 80, z = 1.33, p = .19; WC: U(28) = 89.5, z = 0.93, p = .35), or unresolved episodes (PE: U(28) = 94.5, z = 0.73, p = .47; WC: U(28) = 86, z = 1.08, p = .28).

This finding is contrary to Kim (2008), in which dyads resolved a significantly higher proportion of episodes correctly than individuals, and where pooling linguistic resources appeared to result in greater ability to correctly resolve LREs. Individual learners in Kim's study tended to leave LREs unresolved since they had no resources to draw on other than their own knowledge, the gap in which had given rise to the LRE in the first place. This association between pooling resources and correct LRE resolution is supported by findings from Donato (1994), Storch (2005) and Fernández Dobao (2014), although it should be noted that of these, only Kim (2008) compared LRE resolution between individuals and learner-learner dyads.

One explanation for the results in the present study is that the proportion of correctly resolved episodes in individuals, while not statistically significantly different from pairs, was based on significantly fewer total LREs. This suggests that individual learners did not even attempt to initiate episodes – at least not vocally – that they knew they would not be able to resolve, preferring instead to focus on items they felt they had the linguistic resources to correct. It is possible that individual learners did in fact *silently* initiate more episodes than they verbalised, but preferred not to begin vocalising these if they were unsure as to how to resolve them. There was some evidence of this in the protocols, such as this instance in Illanca's think-aloud while completing her written composition:

Illanca:	"OK let me think… yes this is OK".

Table 4
LRE resolution.

			LREs	%*	M	SD
Correctly resolved	Passage Editing	Pairs (n = 15)	290	71.4%	19.3	5.7
		Individuals (n = 15)	159	67.7%	10.6	5.4
	Written Composition	Pairs (n = 15)	196	83.8%	13.1	6.6
		Individuals (n = 15)	115	89.1%	7.7	3.8
Incorrectly resolved	Passage Editing	Pairs (n = 15)	48	11.8%	3.2	2.7
		Individuals (n = 15)	24	10.2%	1.6	2.4
	Written Composition	Pairs (n = 15)	27	11.5%	1.8	1.9
		Individuals (n = 15)	7	5.4%	0.5	0.5
Unresolved	Passage Editing	Pairs (n = 15)	68	16.7%	4.5	3.6
		Individuals (n = 15)	52	22.1%	3.5	2.5
	Written Composition	Pairs (n = 15)	11	4.7%	0.7	0.7
		Individuals (n = 15)	7	5.4%	0.5	0.9

(* percentage of total LREs in that mode and task).

It seems highly likely that there were unspoken thoughts in the pause between "think" and "yes", and these may have constituted an LRE. The use of the think-aloud may have constituted a limitation that should be considered when interpreting data from individuals.

Incorrect LRE resolution was relatively uncommon, suggesting that on the whole, learners had sufficient linguistic resources to resolve the LREs that arose. Proportions of unresolved LREs were higher, with the highest unresolved figure occurring in individual PE, where over a fifth of all episodes were left unresolved. Individual PE data revealed episodes in which a problem was identified, but no alternative form was suggested. This is illustrated in the following extract from Ilroy's PE think aloud protocol:

Ilroy:	like Hi is not formal … and me, and if this was, like… thanks a million, this is not formal… if, and here brilliant, OK this brilliant here we can't say brilliant is like, this… and then give you a buzz, If I give you a buzz, this buzz is definitely, definitely not formal, so we can't say this in an e-mail to a university…

A similar tendency was observed in Ida's response:

Ida:	OK, so probably, I wouldn't start with er "Hi", if is to a, a, I don't know, a teacher from a university, I would start probably in a different way, then, I think that "thanks a million" is not, like, the perfect way to say, to thank a teacher from a university

In dyadic interaction, conversely, if one participant identified an error without correcting it, the interlocutor could suggest or elicit a correction. While dyadic learners also left PE LREs unresolved, this occurred less frequently than in individuals. In some cases this was for the same reason as in individuals, that is, errors were identified but no resolution was proposed. However, learner-learner interaction also contained examples of one participant asking a question about a form, and this being either ignored or not understood by the second participant, who moved the discussion on and left the previous item unresolved. Paul's question regarding the capitalization of *UK*, for example, was ignored by Patricia, who was focusing on a subsequent structure, *looking forward to + in*:

Paul	It's correct, "I'm sure the information", I'm sorry, "the formation will be brilliant, I'm really looking forward to studying in the UK" UK is here is right? in capital letters, here may I,
Patricia	looking forward
Paul	Here we have the
Patricia	Yeah but to study, erm
Paul	studying, Just studying, I'm writing

The nature of dyadic dialogue meant that on occasion, the LRE got lost in the flow of interaction and remained unresolved.

5.3. Cognitive engagement

Table 5 presents the descriptive statistics for cognitive engagement in LREs.

Mann-Whitney U-tests revealed no statistically significant difference at the $p < .05$ level between settings in the proportion of episodes characterised by limited cognitive engagement, in PE ($U(28) = 62$, $z = 2.07$ $p = .051$) or in WC ($U(28) = 67.5$, $z = 1.85$, $p = .064$). Individual learners engaged cognitively to a similar extent as learners in pairs. Many instances of individual elaborate engagement took the form of a justification for a correction based on the degree of formality of the expression, as demonstrated in Irene's PE response:

Irene:	which reminds me, could you, because can is quite informal so could you, give me

Similarly, Ingrid justified an alternative for "with you" based on her perception of the formality of register:

Ingrid:	In this sentence is "if I would come to study with you, how much would I need to pay in total", it's, is not a correct form, because it's very informal to say to speak with the university so I think it's better if we put for example if I would come to study in your university

The following extract from Paola and Pedro was characterised by elaborate engagement, evidenced by a discussion and justification of the expression "which reminds me":

Paola	Yes, these languages in your university, which reminds me
Pedro	No, it's not remind
Paola	No no
Pedro	Which
Paola	Remind me… *me recuerda* [it reminds me] remind me, erm, because it's plural languages, it's plural so is it's remind me
Pedro	No Because remind me er is you say remind me something, I forgot to close the door
Paola	Remind me that I go to the bakery or something like that
Pedro	So doesn't make sense here we can say in another, in another way
Paola	*Me recuerda*, [it reminds me] which reminds me

The engagement data may be explained by Vygotskian sociocultural theory. LREs characterised by elaborate engagement may be considered evidence that concepts had been internalised by learners; that is, they had developed from being spontaneous concepts that learners were able to utilise without fully understanding their form, to scientific concepts, about which learners had some formal awareness. In the previous example, Paola demonstrated spontaneous knowledge of "remind" by indicating an awareness of the meaning (she provided an L1 translation), but also demonstrated scientific awareness by proposing that the verb form should be singular. Had there been no elaborate engagement evidenced by the presence of metalanguage, it would have been more difficult to demonstrate that the form had yet moved beyond a spontaneous concept.

Table 5
Cognitive engagement in LREs.

			LREs	%*	M
Limited	Passage Editing	Pairs (n = 15)	177	43.6%	11.8
		Individuals (n = 15)	84	35.7%	5.6
	Written Composition	Pairs (n = 15)	144	61.5%	9.6
		Individuals (n = 15)	94	72.9%	6.3
Elaborate	Passage Editing	Pairs (n = 15)	126	31.0%	8.4
		Individuals (n = 15)	151	64.3%	10.1
	Written Composition	Pairs (n = 15)	54	23.1%	3.6
		Individuals (n = 15)	35	27.1%	2.3
Elaborate + Limited	Passage Editing	Pairs (n = 15)	103	25.4%	6.9
	Written Composition	Pairs (n = 15)	34	14.5%	2.3

*percentage of total LREs in each mode and task.

Table 6
Instances of MGD.

	Instances of MGD	M	SD
Pairs (n = 15)	16	1.1	1.2
Individuals (n = 15)	1	0.1	0.3

5.4. Learning

5.4.1. Microgenetic development

Table 6 presents instances of microgenetic development (MGD) observed in learners' task transcripts.

The Mann-Whitney U test revealed statistically significantly more instances of MGD at the $p < .05$ level in pairs than in individuals, $U(28) = 50$, $z = 2.57$, $p = .010$. Qualitative analysis of the transcripts revealed that MGD in peer interaction frequently co-occurred with peer scaffolding. In the following extract, Priscila collaborated with Pamela to support Pamela's understanding of the past form in second conditional structures. Pamela raised the question of which form to use, past or present, and Priscila confirmed her belief it should be the past. Pamela asked again, seeming unsure whether the information provided by Priscila was correct, and Pamela provided specific support contingent on Pamela's apparent lack of sureness in the form of a metalinguistic explanation. Pamela then appeared to have a 'lightbulb' moment in which she remembered how to form conditional sentences. Priscila continued to provide more support in the form of a further example, and asked for confirmation. Pamela's confirmation of the correct answer in this analogous example was evidence of MGD:

Pamela	Here, he's talking about, er "If I pay a deposit now, how much time shall I have to pay the rest of the money?"… but is pay? Or better in the past, "paid"? Or "if I have to pay a deposit now"… this about money all this thing…
Priscila	er… paid, if I paid
Pamela	past?
Priscila	Yes…. Is not past in the, er meaning, is past in the form only, is con, conditional…
Pamela	Ah conditional sentences, OK Like, "if
Priscila	I give you a buzz on the phone number you put in your email, are there a chance you can tell me more?"… we need past?
Pamela	Yes, is similar, if I give, gave, gave you a buzz

If, as the data suggest, observable MGD is associated with uptake following correction or scaffolded input by a peer, then it is unsurprising that there were almost no instances of observable MGD in the individual mode, as there was no interlocutor. The only instance of MGD in individuals occurred in Ibrahim's think-aloud protocol, where he thought through and verbalised a problem relating to prepositions of place. By drawing on his knowledge of the analogous prepositional structure "at + school", he was able to resolve the episode and produce "at + university". The evidence of microgenetic development is in his application of this constructed knowledge to a subsequent problem involving the same form:

> "just writing to say"… "formation in your university"… now I'm not, not sure but I think it's not at, in your university, but language formation at your university, I'm not sure but I think it's at not in, because it's like at school, so at your university… same mistake erm… another time, these languages at, "so it would be really cool to study these languages in your university"… erm, I think… in your university, at your university, no in your university…

Ibrahim's strategy of drawing on existing knowledge to help resolve a new problem may be considered an example of self-scaffolding. Ibrahim interrogated himself about what he did not understand, then resolved the episode through self-explanation in a process akin to the interrogation of the epistemic self described by Holton and Clark (2006). Ibrahim self-scaffolded heuristically by making optimal use of available resources (Bickhard, 2005), in this case his knowledge of analogous forms.

Table 7
Post-test items that corresponded to LREs in Task 1.

	Items that corresponded to LREs	As a percentage of total LREs
Pairs (n = 30)	614	75.6%
Individuals (n = 15)	201	85.5%

Table 8
Post-test items attempted.

	Items that corresponded to LREs	Items attempted	Items attempted as a percentage of items that corresponded to LREs
Pairs (n = 30)	614	249	40.6%
Individuals (n = 15)	201	103	51.2%

5.4.2. Post-test responses

Table 7 presents numbers of post-test items corresponding to LREs in task 1, and expresses this number as a percentage of total LREs in each mode.

The Mann-Whitney U test revealed no statistically significant difference between settings in the number of items, expressed as a percentage of total LREs, corresponding to LREs in Task 1 at the $p < .05$ level, $U(43) = 752$, $z = 0.67$, $p = .50$. That items corresponded to between 76% and 86% of LREs indicates that the post-test managed to gather data relating to most of the languaging that occurred.

If a participant attempted to correct a form in the post-test, this was marked as a test item attempted. Table 8 presents the numbers of test

Table 9
Post-test items corrected i) in agreement with LRE resolution, ii) in disagreement with LRE resolution, iii) when the LRE had been unresolved and iv) only partially.

	Items attempted	Items corrected in **agreement** with LRE resolution	Items corrected in **agreement,** as a proportion of items attempted	Mean items per participant
Pairs (n = 30)	249	182	73.1%	6.1
Individuals (n = 15)	103	73	70.9%	4.9
	Items attempted	Items corrected in **disagreement** with LRE resolution	Items corrected in **disagreement,** as a proportion of items attempted	Mean items per participant
Pairs (n = 30)	249	37	14.9%	1.2
Individuals (n = 15)	103	4	3.9%	0.3
	Items attempted	Items corrected when there had been **no resolution**	Items corrected when there had been **no resolution,** as a proportion of items attempted	Mean items per participant
Pairs (n = 30)	249	11	4.4%	0.4
Individuals (n = 15)	103	10	9.7%	0.7
	Items attempted	Items corrected **only partially**	Items corrected **only partially,** as a proportion of items attempted	Mean items per participant
Pairs (n = 30)	249	17	6.8%	0.6
Individuals (n = 15)	103	16	15.5%	1.1

items attempted and expresses this number as a percentage of test items that corresponded to LREs:

The Mann-Whitney *U* test revealed a statistically significantly higher proportion of post-test items attempted by individuals than dyadic learners, U(43) = 129.5, z = 2.29, p = .022. This suggests that participants found forms focussed on individually more memorable, and therefore easier to identify as errors in the post-test, than forms focussed on in dyads. This may support Swain's (2013) observation that in peer interaction, not all talk is social, but may in fact be private, for the self. Learners may sometimes appear to be talking "to each other, but are in fact following their own agenda" (Swain, 2013: 201). Such an assertion relates to Vygtosky's (1987) concept of private speech, in which inner speech, that is, speech that has become internalised as a tool for the purposes of self-regulation, surfaces in order to aid the speaker in the resolution of cognitively complex tasks. In the following extract, Paul vocalises a series of language problems but resolves these himself. His speech is not, it would seem, socially directed. Paul follows his own agenda and decides on the words to write in order to complete the task:

Paul	Erm, this idea… "but apart from the studies, time for making leisure activities is also a priority for me" where? Whereas?
Patricia	*aunque o algo así, no sé como decirlo* [although, or something like that, I don't know how to say it]
Paul	whereas *mientras que* [whereas]
Patricia	Ah vale [ah OK]… *con esto* [with this] then
Paul	OK "this is important, whereas"
Patricia	Erm we could erm talk er we could say that erm, we
Paul	Time for make
Patricia	Yes
Paul	Making
Patricia	We?
Paul	Time for making leisures activities
Patricia	ah OK, or
Paul	Or OK
Patricia	Or maybe, like
Paul	OK

Learners in pairs may not always listen to each other's languaging, and as a result, this may not always constitute a learning opportunity for the interlocutor.

Post-test items attempted were further categorised as i) in agreement with the original LRE resolution in task 1, ii) in disagreement with the LRE resolution, iii) there had been no LRE resolution, and iv) only partial, i.e. the post-test item was circled or underlined but no alternative form proposed. Table 9 presents these data, also expressed as percentages of items attempted.

Mann-Whitney U tests revealed no statistically significant differences at the $p < .05$ level between dyadic learners and individuals in proportions of test items i) resolved in agreement with LRE resolution, U(43) = 796, z = 0.014, p = .99, ii) in disagreement with LRE resolution, U(43) = 366, z = 0.90, p = .37, iii) when the corresponding LRE had been left unresolved, U(43) = 369, z = 0.59, p = .56, or iv) resolved only partially, U(43) = 321, z = 0.36, p = .72.

Between 71% (in individuals) and 73% (in pairs) of test items attempted were resolved in agreement with the LRE resolution. This suggests associations between LREs and learning, with new knowledge constructed (or existing knowledge consolidated) in the LRE surfacing again on the isomorphic post-test. No statistically significant differences were found between modes in test items resolved in agreement with LRE resolution, mirroring findings from Nassaji and Tian (2010), who found that although pairs demonstrated greater accuracy than individuals when completing cloze and text editing tasks seeded with phrasal verbs, there were no statistically significant differences in learning gains.

Also of interest are test items resolved in agreement with LRE resolution when the LRE had itself been *incorrectly* resolved. Pamela and Priscila, for example, settled on the incorrect form "Dear Mister" as an alternative to "Hi" in Task 1:

Pamela	OK, well this is formal no?
Priscila	Yes is formal because it's a student and he is, er, Andy, he is writing to, to
Pamela	To a university no? So this "Hi"… first, this is not right
Priscila	No it is not right, it needs to be much more formal, maybe Esteemed, er,
Pamela	I think Dear
Priscila	OK yes, Dear, Dear Mrs or Dear Mister

Both participants then reproduced the incorrect form "Dear Mister" in the post-test. As in LaPierre's (1994) study, participants apparently retained incorrect knowledge constructed collaboratively. (Swain, 1998) argues that retention of incorrect knowledge may be a greater indicator of learning associated with languaging than the reproduction of correct knowledge, since learners may have known the correct forms prior to the episode.

Relatively few test items were resolved in disagreement with LRE resolution in the task: just 4% in individuals, and 15% in pairs. This suggests that in both modes there exists a relationship between decisions made during LREs and subsequent receptive awareness of forms topicalised. Despite the lack of statistically significant difference, it is

noteworthy that the higher of these two figures is for pairs. As discussed above, even when LREs were resolved a certain way, dyadic learners may have been silently following their own agenda, which sometimes only became apparent in the post-test. Dyadic learner Penelope, for example, went on to produce a post-test in which over half of the items attempted were corrected in a way that differed from LRE resolutions during the task with Paco. In the following task excerpt, Penelope participated in an LRE regarding the formality of the adjective "cool", which was resolved by Paco, who decided on "great".

Penelope	I'm sure the course
Paco	the course
Penelop	Will be
Paco	Will be… *pero tenemos que utlilzar palabras más, más palabras porque* [but we need to use words that are more, more words because]
Penelope	*Más formal* [more formal]
Paco	*Otro vocabulario, un diferente vocabulario, todo es muy simple,* [another vocabulary, a different vocabulary, it's all too simple] … *yo pienso* [I think]
Penelope	Will be, will be
Paco	Will be great, I'm sure, the course will be great, "I'm really looking forward"
Penelope	"Really looking forward"

In the post-test, Penelope corrected the word "cool", but instead of "great" wrote "excellent". This suggests she may in fact have preferred "excellent" during the task, but was happy to let Paco decide on "great".

While there were very few instances of test items resolved when the corresponding LRE had been left unresolved (4% of items attempted in pairs, 10% in individuals), that the highest of these figures corresponds to individuals may indicate a further potential limitation of the think-aloud methodology. If individual learners made corrections on the test, but had not resolved LREs about these same forms during the task, this indicates a lack of verbalisation about forms learners in fact knew. It is plausible that lack of familiarity with verbal protocol procedure, or the perceived pressure of being recorded alone, may have inhibited the verbal resolution of episodes.

Conclusion

The present study was subject to a number of limitations which should be acknowledged in order to contextualise possible conclusions and pedagogical recommendations. Firstly, the quasi-experimental design necessitated a narrow observational focus within pair and individual work in order for comparisons to be drawn. Individual learners completed tasks alone, without interaction with peers or teachers or access to resources, and dyads completed tasks without teacher assistance or resources. The reality, of course, is that in classroom pairwork there are opportunities for the teacher to participate in learners' interaction while monitoring, acting as a resource and a facilitator when questions and difficulties arise; individual learners working alone, on the other hand, are able to consult online resources or contact their peers for assistance or pair up with a classmate through Zoom. Real learning contexts are more fluid than those under investigation here, which have been necessarily restricted in order for comparisons to be drawn between individual and dyadic languaging.

Secondly, it is possible that the statistically significant difference between pair and individual LRE numbers in the present study relates to the use of think–aloud methodology to collect data from individuals. As suggested by Kim (2008), think-alouds may produce an additional cognitive demand not experienced by dyads, which may negatively impact the number of LREs produced. In this case, the act of verbalising may be reactive (Ellis, 2001; Jourdenais, 2001) to the task, altering the cognitive processes taking place, and lending support to Vygotsky's position that "thought is restructured as it is transformed in speech. It is not expressed but completed in the word" (1987: 150). If thought and speech interact as they co-occur then the use of verbal protocols to observe cognition may be inherently problematic. As in any study employing a think-aloud protocol as a data collection instrument, then, there is a risk that participants may not have verbalised everything they were thinking, and results must therefore be considered in this context.

Thirdly, that the post-test was isomorphic – containing the same number of the same sorts of errors as task 1 – meant, by definition, that it was very similar to the task. One possible consequence of this similarity is the possible effect of task repetition: repeated exposure to the same or very similar tasks may improve learners' accuracy with forms contained within (Gass, Mackey, Alvarez-Torres and Fernández-García, (1999); Hawkes, (2011)). However, it is worth highlighting here that the post-test was only one of two ways of measuring learning: the other was the observation of in-task microgenetic development, and triangulating these two data sets provides a more complete picture of learning processes than examining post-test scores alone.

To conclude, the present study set out to compare individual task performance with pairwork in order to investigate the potential benefits and drawbacks of the increased periods of independent language study found in online and socially distanced EFL classrooms. Drawing on Vygotskian Sociocultural Theory and Swain's concept of languaging, the analysis of dyadic interaction and individual think-alouds revealed the presence of LREs in both individual study and pairwork. While individuals produced statistically significantly fewer LREs than dyads, individual numbers were similar to LREs initiated by each learner in dyads, suggesting individuals identified language problems to a similar extent as their dyadic counterparts; the difference is that in pairwork, the additive effect of two learners leads to greater amounts of languaging. This finding complements the significant body of research to have found collaborative pairwork to be beneficial for learners (Philp et al., 2013). Proportions of correctly resolved LREs were similar between settings, although the individual proportion was based on fewer LREs, suggesting individuals did not initiate episodes they could not resolve. The extent to which LREs were characterised by limited engagement (linguistic preferences were stated without further deliberation) or elaborate engagement (with evidence of a cognitive self-regulation strategy) did not differ significantly between settings. Microgenetic development within LREs was observed more in dyadic interaction, and learners in both settings generally responded to post-test items in agreement with LRE resolution, suggesting associations between languaging and learning. However, dyadic learners attempted fewer test items relating to their LREs, suggesting forms languaged individually are more memorable.

Some tentative pedagogical recommendations can be proposed. Learners who find themselves spending more time doing language tasks alone, either at home or in socially distanced classrooms, can be reassured that while they may not engage in as much visible languaging as learners working in pairs, the languaging they do produce is generally characterised by elaborate engagement, and is as effective a learning opportunity, in terms of memorability for posterior recall, as languaging with a partner. In order to increase languaging in individual work, learners can be encouraged to take a more questioning role of their own language production: to monitor their own language and be on the lookout for gaps, considering how to make their language more accurate, appropriate or sophisticated. Individual learners can self-scaffold their own development by fully considering the range of linguistic resources available to them, thus widening their range of action. In this way, the amount of individual languaging may more closely approximate that of dyads.

The statistically significantly lower numbers of LREs and instances of MGD observed in individuals strongly suggests that the presence of an interlocutor is associated with languaging and learning. Teachers could therefore recommend that individual learners at home seek out an interlocutor to work with, either in person or via Zoom, in order to complete language tasks. Practice with an interlocutor would not only favour languaging and learning, but would help improve pronunciation and speaking skills.

Pedagogical implications for learners in dyads may also be proposed in order to deepen the current understanding of peer scaffolding. Given the evidence in the present study that learners do not always listen to

or learn from each other's languaging in pairwork, teachers could provide guidance on how learners should interact with peers. Encouraging learners to adopt a more questioning role that invites their interlocutor to consider more appropriate or sophisticated language forms, and to make clarification requests and confirmation checks such as "so what you're saying is… so if I understand you correctly… so do you mean this?" may help improve the quality of interaction and opportunities for peer learning. An exploration of this proposal would make for potentially fruitful future research.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendices

Appendix 1: Task 1 - passage editing task

Read this email from a student to a University in the UK, and correct any problems / errors.

Remember to consider the full range of possible errors. These may include:

- Grammar
- Vocabulary
- Spelling
- Punctuation
- Style (formal / informal)

Hi Mrs Horowitz,

Just writing to say thanks a MILLION for your email about language formation in your university. The language learning is really important for students here in spain, not just English but other languages too, at my country it is imposible to find good courses in Chinese or the Russian, although it depends of the place, so it'll be really cool to study these languages in your university. Which reminds me, can you give me an aproximate cost of the courses? If I would come to study with you, how much would I need to pay in total? If I pay a deposit now, how much time shall I have to pay the rest of the money? I'm sure the formation will be BRILLIANT, I'm really looking forward to studying in the uk, but apart from the studies, time for making leisure activities is also a priority for me. There were something in your email about what students can do in their free time at the weekends – if I give you a buzz on the phone number you put in your email, are there a chance you can tell me more?

Bye for now and see you soon!

Andy

P.S. Any recommendations for good places on the city to visit at night-time? We really want to take full advantage of our time in England!

Appendix 2: Task 2 - written composition

Write a letter to your local newspaper giving your opinion about this topic:

"Should we ban smoking everywhere – even at home?"

You might want to include comments about the following:

- Health issues related to smoking
- The importance of individual freedom
- Taxes on cigarettes
- Plus any ideas of your own.

First, make notes and decide which ideas will go into each paragraph. Then write your letter, and try to give emphasis to your opinions. Finally, read and check your letter for mistakes.

Appendix 3: Post-test (isomorphic passage editing task)

Read this email from a student to a University in the UK, and correct any problems / errors.

Remember to consider the full range of possible errors. These may include:

- Grammar
- Vocabulary
- Spelling
- Punctuation
- Style (formal / informal)

Hi Mrs. Horowitz,

Just letting you know that I've now received the extra information you sent me about language formation on England, thanks a MILLION, once again. The university studies at spain are BRILLIANT for subjects like Enginneering, for the languages I think it's better in the UK, so it'll be really cool to study there. Any recommendations for an english certification to acredit previous formation? I have seen that we would make an English test in the first week, but what does it consist in? Before I leave Spain I'll check your website again to see if there is things I need to bring, and I should give you a buzz if I have any questions – any chance you can confirm if there are a phone number on your webpage?

Bye for now and see you soon,

Andy

Appendix 4: Examples of LRE coding decisions

Excerpt 1:

Patricia	It's better "not just English but other languages too, at my country" is In Country
Paul	yeah I think so
Patricia	in, in my country, the preposition
Paul	yeah, I think in, or on
Patricia	or in, on …
Paul	yeah, I, I dunno, in, I think..
Patricia	in my country

Patricia and Paul together decide on the correct form "in my country", so the LRE is coded as **correctly resolved**. Patricia uses metalanguage to describe the type of error (preposition) and Paul proposes an alternative form (on), so for both participants cognitive engagement in this LRE is coded as **elaborate**. There is no evidence of learning of forms within the episode, as demonstrated by further use of the correct form, so the LRE is coded as **not demonstrating microgenetic development**.

Excerpt 2

Patricia	"you put in your email, if I give you a buzz on your phone number are there a chance you tell me more", a chance, will be there a chance
Paul	yeah …
Patricia	"you can tell me more?"
Paul	OK

Patricia attempts to correct the error "are there a chance" by proposing "will be there a chance", and Paul agrees with this. As they have decided on an incorrect alternative form, the LRE is coded as **incorrectly resolved**. There is no discussion of other possible forms or justification for the correction, so cognitive engagement in both participants is coded as **limited**. There is no evidence of learning of forms within the episode, as demonstrated by further use of the form decided upon, so the LRE is coded as **not demonstrating microgenetic development**.

References

Baralt, M., Gurzynski-Weiss, L., & Kim, Y. (2016). The effects of task complexity and classroom environment on learners' engagement with the language. In M. Sato, & S. Ballinger (Eds.), *Peer interaction and second language learning. Pedagogical potential and research agenda* (pp. 209–240). Amsterdam: John Benjamins. 10.1075/lllt.45.09bar.

Basterrechea, M., & García Mayo, P. (2013). Language-related episodes during collaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough, & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 25–44). Amsterdam: John Benjamins. 10.1075/lllt.34.05ch2.

Bickhard, M. H. (2005). Functional scaffolding and self-scaffolding. *New Ideas in Psychology, 23*, 166–173. 10.1016/j.newideapsych.2006.04.001.

Christenson, S., Reschly, A., & Wylie, C. (2012). *Handbook of research on student engagement*. New York, NY: Springer. 10.1007/978-1-4614-2018-7.

Corno, L. (1993). The best-laid plans: Modern conceptions of volition and educational research. *Educational Researcher, 22*, 14–22.

Corno, L., & Mandinach, E. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist, 18*, 88–108.

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf, & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.

Ellis, R. (2001). Introduction: Investigating form-focussed instruction. *Language Learning, 51*, 1–46. 10.1111/j.1467-1770.2001.tb00013.x.

Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair and individual work. *Journal of Second Language Writing, 21/1*, 40–58. 10.1016/j.jslw.2011.12.002.

Fernández Dobao, A. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research, 18/4*, 497–520. 10.1177/1362168813519730.

Fortune, A. (2013). The role of metalanguage in the performance of a sequence of collaborative output tasks by five L2 learner dyads. In K. Roehr, & G. A. Ganem-Gutierrez (Eds.), *The metalinguistic dimension in instructed second language learning* (pp. 171–194). London: Bloomsbury. 10.5040/9781472542137.ch-008.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74/1*, 59–109. 10.3102/00346543074001059.

Gass, S., & Mackey, A. (2007). Input, interaction and output in second language acquisition. In B. Van Patten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175–200). Mahwah, NJ: Lawrence Erlbaum.

Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*, 549–581. 10.1111/0023-8333.00102.

Gilabert, R., & Barón, J. (2013). The impact of increasing task complexity on L2 pragmatic moves. In K. McDonough, & A. Mackey (Eds.), *Second language acquisition in diverse educational contexts* (pp. 45–70). Amsterdam: John Benjamins. 10.1075/lllt.34.06ch3.

Hawkes, M. (2011). Using task repetition to direct learner attention and focus on form. *ELT Journal*. 10.1093/elt/ccr059.

Helme, S., & Clarke, D. (2001). Identifying cognitive engagement in the mathematics classroom. *Mathematics Education Research Journal, 13/2*, 133–153.

Holton, D., & Clarke, D. (2006). Scaffolding and metacognition. *International Journal of Mathematical Education in Science and Technology, 37*, 127–143. 10.1080/00207390500285818.

Jourdenais, R. (2001). Cognition, instruction and protocol analysis. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 354–376). Cambridge: CUP. 10.1017/cbo9781139524780.014.

Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal, 92/1*, 114–130. 10.1111/j.1540-4781.2008.00690.x.

Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research, 15/2*, 183–199. 10.1177/1362168810388711.

Knouzi, I., Swain, M., Lapkin, S., & Brooks, L. (2009). Self-scaffolding mediated by languaging: Microgenetic analysis of high and low performers. *International Journal of Applied Linguistics, 20/1*, 23–49. 10.1111/j.1473-4192.2009.00227.x.

Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. *Language Teaching Research, 21/6*, 665–680. 10.1177/1362168816683559.

LaPierre, D. (1994). *Language output in a cooperative learning setting: Determining its effects on second language learning* MA thesis. Ontario Institute for Studies in Education: University of Toronto.

Lasito, & Storch, N. (2013). Comparing pair and small group interactions on oral tasks. *RELC Journal, 44/3*, 361–375. 10.1177/0033688213500557.

Leow, R. P. (1997). Attention, awareness, and foreign language behaviour. *Language Learning, 47/3*, 467–505.

García Mayo, P. M. (2002). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research, 37/3-4*, 323–341. 10.1016/s0883-0355(03)00008-9.

Miles, M. B., & Huberman, M. A. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Newbury Park, CA: Sage.

Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research, 14/4*, 397–419. 10.1177/1362168810375364.

Ohta, A. S. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese classroom. In J. K. Hall, & L. Verplaeste (Eds.), *The construction of second and foreign language learning through classroom interaction* (pp. 47–71). Mahwah, NJ: Erlbaum.

Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. Mahwah, NJ: Lawrence Erlbaum. 10.4324/9781410604712.

Philp, J., Adams, R., & Iwashita, N. (2013). Peer interaction and second language learning. doi:10.4324/9780203551349

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics, 36*, 50–72. 10.1017/s0267190515000094.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82/1*, 33–40. 10.1037/0022-0663.82.1.33.

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10/4*, 277–303. 10.1016/s1060-3743(01)00046-7.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11/2*, 129–158. 10.1093/applin/11.2.129.

Storch, N. (1997). The editing talk of adult ESL learners. *Language Awareness, 6/4*, 221–232. 10.1080/09658416.1997.9959931.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System, 27/3*, 363–374. 10.1016/s0346-251x(99)00031-7.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*, 119–158. 10.1111/1467-9922.00179.

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14/3*, 153–173. 10.1016/j.jslw.2005.05.002.

Storch, N. (2008). Metatalk in a pair work activity: level of engagement and implications for language development. *Language Awareness, 17/2*, 95. 10.2167/la431.0.

Svalberg, A. M. L. (2009). Engagement with language: Interrogating a construct. *Language Awareness, 18/3-4*, 242–258. 10.1080/09658410903197264.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). Cambridge: CUP.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contribution of halliday and vygotsky* (pp. 95–108). London: Continuum. 10.5040/9781474212113.ch-004.

Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching, 46/2*, 195–207. 10.1017/s0261444811000486.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16/3*, 371–391. 10.1093/applin/16.3.371.

Vitta, J. P., & Al-Hoorie, A. H. (2020). The flipped classroom in second language learning: A meta-analysis. *Language Teaching Research*, Article 136216882098140. 10.1177/1362168820981403.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1987). *Thinking and speech*. New York: Plenum Press.

Weinstein, C., & Mayer, R. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching and learning* (pp. 315–327). New York: Macmillan.

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26/3*, 445–466. 10.1177/0265532209104670.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89–100. 10.1111/j.1469-7610.1976.tb00381.x.