This item is the archived peer-reviewed author-version of:

In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch

Reference:

Vandommele Goedele, Van den Branden Kris, Van Gorp Koen, De Maeyer Sven.- In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch
Journal of second language writing - ISSN 1060-3743 - 36(2017), p. 23-36
Full text (Publisher's DOI): <https://doi.org/10.1016/J.JSLW.2017.05.010>
To cite this reference: <https://hdl.handle.net/10067/1432830151162165141>

Manuscript Number: 3213R4

Title: In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch

Article Type: SI:SLW in CMC contexts

Keywords: multimodal composition; language learning in out-of-school settings; task-based language learning; writing development of beginner language learners

Corresponding Author: Dr. Goedele Vandommele, Ph.D.

Corresponding Author's Institution: KU Leuven

First Author: Goedele Vandommele, Ph.D.

Order of Authors: Goedele Vandommele, Ph.D.; Kris Van den Branden, Prof; Koen Van Gorp, Phd; Sven De Maeyer, Prof

Abstract: This study investigates the effects of a collaborative multimodal writing intervention in different settings on the development of writing skills of beginner learners of Dutch as a second language. The study extends previous research by using multiple measures to assess writing development and by comparing the impact of different settings, i.e., a task-based in-school intervention and an out-of-school leisure intervention. All participants in both interventions (N= 52) completed two writing tasks before and after the intervention (in a pretest-posttest design). Participants in a non-intervention group (N = 32) did the same before and after a two-week holiday in which they did not receive writing instruction. Participants' output was assessed by analytic CAF-measures as well as for communicative effectiveness. Multilevel analyses revealed significantly greater growth for both intervention groups compared to the non-intervention group. The in-school group showed growth for six measures (complexity and text length on one task and communicative effectiveness, content, lexical diversity and text length on the other). Also the out-of-school group revealed greater gains for six measures (complexity on one task and content, two measures of syntactic complexity, lexical diversity and text length on the other). In all, this study shows that multimodal composition can promote writing of beginner L2 learners, regardless of the exact context in which the writing takes place. Furthermore, the study shows the importance of using multiple measures and multiple tasks to evaluate writing development.

Thank you to reviewers

We would like to thank our reviewers for their thorough review and constructive feedback. Their comments have contributed to sharpen the paper and their efforts were greatly appreciated.

Sincerely,

The Authors

In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of
Dutch

Goedele Vandommele*, Kris Van den Branden*, Koen Van Gorp** and Sven De Maeyer***

*University of Leuven, **Michigan State University, ***University of Antwerp, Belgium

Author Note

Vandommele: Centre for Language and Education, University of Leuven, Blijde Inkomststraat 7, 3000 Leuven, Belgium (e-mail: goedele.vandommele@kuleuven.be); Van den Branden: Centre for Language and Education, University of Leuven, Blijde Inkomststraat 7, 3000 Leuven, Belgium (e-mail: kris.vandenbranden@kuleuven.be); Van Gorp: Michigan State University, 619 Red Cedar Road, Wells Hall, B-135F, East Lansing, MI 48824-1027, USA (When working on this study Van Gorp was affiliated to the University of Leuven) (e-mail: vangorpk@msu.edu); De Maeyer: University of Antwerp, Venusstraat 35, 2000 Antwerpen, Belgium (email: sven.demaeyer@uantwerpen.be)

In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch

Introduction

Since the beginning of the 21st century, digital multimodal writing has gained ground and its potential for developing L2 writing skills has been identified (Bloch, 2007; Kern & Schultz, 2005; Warschauer, 2009). Two main research strands can be distinguished in studies on multimodal composition: one strand investigates students' authentic language use through multimodal writing practices in outside-school contexts (e.g., Hull & Schultz, 2001), the other explores the incorporation of multimodal writing techniques in the classroom (e.g., Karlsson, 2015; Warschauer, 2009). Each strand has its own emphases and strengths. Studies in out-of-school settings most often look into how multimodal practices change people's notions of literacy, learning and authority (e.g., Ito et al., 2008). Inside-school research mostly focuses on how education can put these changed notions to good use to foster students' writing skills (e.g., Thorne & Black, 2007). However, few studies have compared the effects of multimodal writing in in-school and out-of-school settings on L2 writing development, even though better understanding of each setting's merits could assist in setting up more effective interventions for less proficient writers. In this study, we attempt to address this gap by examining the differential effects of multimodal writing practice in the two settings on the L2 writing development of adolescent, beginner learners of Dutch.

Out-of-school and multimodal writing as resources for struggling academic writers

For students today, advanced writing skills are a key competence and an essential requirement for educational success. Yet, writing is notoriously difficult. For adolescent newly-arrived migrants (“newcomers”), L2 academic writing is especially taxing. Many of them have experienced interrupted or limited schooling, resulting in poor academic skills, little familiarity with specific genres, or undue focus on linguistic elements at the expense of broader issues related to content and textual organization. In addition, adolescent L2 writers often feel disengaged from literacy practices in the classroom due to difficulties with academic language (Niiya, Warschauer, & Zheng, 2013). In short, many adolescents, newcomers in particular, struggle to meet the demands related to writing that are required in school.

In this respect, Graham and Perin (2007) call for innovative research that can help improve struggling writers' skills. They propose that such research should not be limited to in-school contexts but should also involve settings outside the school as most adolescents, including L2 newcomers, show considerable engagement in writing activities in contexts outside of the school (see Hull & Schultz, 2001; Ito et al., 2008, for reviews). Such activities include writing rap verses, letters, journal entries, diaries and poems, as well as writing activities that embrace new technologies (Lenhart, Arafeh, & Smith, 2008; Warschauer & Matuchniak, 2010), as seen in blogging, constructing websites, participating in online social networks or in gaming environments, creating music videos, e-mailing and using online chat (Black, 2005; Thorne, Fischer, & Lu, 2012).

Observations of these rich outside-school digital writing practices have transformed ideas of what counts as “literacy” and “writing” in the 21st century. As a result, not only print-based pen-and-paper writing activities but also screen-based writing activities are considered

worthwhile. Those products often contain text and simultaneously use sophisticated layout, graphics, photographs and images, and thus multiple “modes,” for meaning-making. To incorporate all these varied forms, the term “multimodal literacy” has been coined (Cope & Kalantzis, 2000). In view of the engagement of adolescents in multimodal writing, such writing practices may offer an interesting route to aid struggling writers.

Advantages of Multimodal Writing

In various ways, adolescents' engagement in multimodal writing practices can support the development of their L2 writing skills. In her review of 76 studies on multimodal composition in both in-class and out-of-class settings, Smith (2013) distinguishes six recurring themes, fairly naturally occurring and non-exclusive features of multimodal composition that promote writing skills: (1) enhanced motivation, (2) high-quality collaboration, (3) specific benefits for struggling writers, (4) opportunities for explicit instruction, (5) scaffolding, and (6) similarities between the multimodal composition process and print writing processes.

First of all, many studies reported learners' high levels of engagement during multimodal writing activities. These high levels of motivation have been attributed to factors such as the digital medium itself (Warschauer, 1996), the learners' autonomy, the authenticity of the writing product and personal value of the text (Jiang & Luk, 2016). Even though Smith (2013) does not distinguish authenticity as a separate factor, the advantages of authentic practice have been shown to go beyond motivational aspects. The real-world audience associated with the internet and the knowledge that writing products will be read (by someone different from the teacher) may not only increase writers' motivation but can also result in superior writing products (Lammers, Magnifico, Curwood, Pytash, & Ferdig, 2014; Magnifico, 2010) and enhanced writing development (Purcell-Gates et al., 2012). These benefits might result from learners being

pushed to produce messages explicitly and clearly (Ellis, 2005), or can be attributed to the fact that by publishing online, the audience of a writing task becomes “tangible” and can even provide online feedback (Magnifico, 2010), resulting in more goal-directed writing.

Second, in multimodal writing, there are numerous possibilities for collaboration among learners and between learners and a wide variety of supportive interlocutors. Online compositions easily allow people to share texts and interact about them, to divide labor, to build on a previous composer’s work and to receive feedback from authentic audiences (Jenkins, 2006; Murray, 2005). Consequently, by co-constructing target-like language use, writers may create new knowledge they did not possess prior to the task (Nassaji & Tian, 2010; Swain & Lapkin, 2001). Moreover, when collaborating in online environments, traditional power relations do not always apply. Age for example, does not necessarily imply authority, whereas digital skills might do so. Instances of successful, collegial collaboration between adults and youth have been observed in multimodal writing activities and coined “a pedagogy of collegiality” (Chávez & Soep, 2005)

Third, multimodal writing appears especially beneficial to struggling writers, such as L2 learners and newcomers. It provides these learners with alternative means to construct and communicate identities (Halverson, 2010) and invites them to share their unique backgrounds and perspectives (Kelly, 2012). In many cases, multimodal composing also allows writers to use more than one language, leading to the validation of language learners' unique linguistic repertoire. Implied in Smith’s (2013) review is that multimodal composition allows L2 learners non-linguistic alternatives to express meaning if linguistic resources are lacking (Sylvester & Greenidge, 2009). Such alternative means of expression may reduce demands of formulation and leave more attention for the processes of selecting, organizing and preparing ideas.

Fourth, in schools, multimodal composition provides an occasion for explicit instruction. Teachers have been found to explicitly instruct learners in technological issues, the multimodal design and the metalanguage related to their multimodal products (Smith, 2013).

Fifth, multimodal composition activities elicit various scaffolding practices. To support composition, outside-school learners often start by studying real-world examples and models. In schools, learners may profit from teacher support. Teachers for example assist learners in the sequencing, viewing and analysis of multimodal products. This may take the form of drawing attention to the selection, structure and formulation of ideas, linguistically or otherwise (Li, 2013; Smith, 2013) or of pushing learners' output further and extending learners' repertoire of skills and genres, which may lead them to move beyond superficial task performance (e.g., repetitive use of structures or vocabulary, less cognitively challenging content) that is often observed in multimodal composition outside school (Mills, 2010a). Often, teachers also present traditional forms of writing during multimodal composition, for example, in the form of written narratives or storyboards. These traditional forms of writing are often helpful, but are sometimes experienced by learners as stifling (Smith, 2013).

Sixth, through multimodal composition learners engage in complex and recursive composition processes, which also encourage reflection. Multimodal composition has distinct stages, reminiscent of traditional writing, which transpire in a non-linear and iterative manner: brainstorming for ideas, selecting and editing texts, pictures, drawings, photographs, music and video sequences (Bruce, 2009). Moreover, video composition has been shown to be a valuable site for the development of meta-awareness about composition processes, which has the potential to feed back into other (written) compositions (VanKooten, 2016).

Multimodal Writing In School

Many educators are in favor of incorporating multimodal practices into the L2 classroom as a way of harnessing the potential advantages of multimodal writing for academic purposes and integrating the literacy practices of the 21st century (Kern & Schultz, 2005; Warschauer, 2009). Studies have shown that the incorporation of multimodal composition in classroom settings is both feasible and worthwhile (Hafner, Chik, & Jones, 2015). Of the 76 studies reviewed in Smith (2013), 50% explored multimodal composition in school contexts. They showed that multimodal practices combined with good instruction displayed many of the above-mentioned benefits. The studies found greater learner motivation and participation (Warschauer, 1996) and enhanced possibilities for interaction and collaboration (Jewitt, 2006). Furthermore, studies found more extensive or better writing through multimodal composition practices such as blogging (Bloch, 2007; Karlsson, 2015) and digital storytelling (Ohler, 2013; Yang & Wu, 2012). These benefits also positively affected academic writing competences (Yang & Wu, 2012).

In addition, the presence of a competent language teacher may increase the value of multimodal composition for L2 writing. Teachers can monitor composition processes, make sure that all learners profit from the writing experience, push learners' output (Swain & Lapkin, 1995) and provide scaffolding (Gibbons, 2002) and relevant explicit instruction (Norris & Ortega, 2000). A promising approach to multimodal composition in the classroom is task-based language teaching (Chapelle, 2005; Fuente, 2003; Van den Branden, Bygate, & Norris, 2009). The integration of multimodal composition provides learners with opportunities to engage in language learning through meaning-making and active collaboration, which constitutes the core of task-based language teaching. In this respect, it is hardly surprising that multimodal

composing is increasingly used in task-based environments to promote L2 writing development (Dal, 2010; González-Lloret & Ortega, 2014; Oskoz & Elola, 2014).

At the same time, the integration of multimodal composition in (task-based) classroom settings faces several hurdles. Obvious ones relate to teachers struggling to understand and incorporate new technology (e.g., Tan & McWilliam, 2009), or schools' lack of digital devices or internet access (Sadik, 2008). Undeniably, schools and teacher education have been slow in exploring opportunities provided by new technologies (Kessler, 2007). Other implementation issues are related to the transition to a new pedagogy. For example, Miller's (2013) synthesis of research on the implementation of digital video composition pointed to the varied ways in which teachers trivialized learning through multimodal composition, refused to relinquish control, opposed students' active roles, or experienced difficulties in balancing direct instruction and learner-centered scaffolding.

Finally, there are cases in which the advantages of school settings conflict with the potentially beneficial features of multimodal composition. Teachers' scaffolding practices, for example, sometimes get in the way of learners' enjoyment of composition practices. In Gilje's (2010) study, the storyboard and written narrative that were introduced by the teacher to better prepare the students for multimodal work proved counterproductive. Steeves (2015) found that the compulsory critical reading of texts met with resistance from the learners. Even the mere act of replicating interesting outside-school activities into a classroom context may challenge their appeal. Lotherington, Neville-Verardi and Ronda (2009), for example, showed how students opposed the incorporation of elements from their digital culture into classroom contexts, claiming that such acts were "culturally invasive" and of "suspect educational value" (p. 30).

This study

In summary, a considerable body of research suggests that multimodal writing practices in both in-school and out-of-school settings might be beneficial for L2 writing development. In in-school settings, the advantages of scaffolding and explicit instruction can be added to the list of potential benefits of multimodal composition in out-of-school settings. However, the characteristics of multimodal composition in school (scaffolding and explicit instruction), together with implementation issues related to the incorporation of new technology and implementation of innovative pedagogical approaches, may weaken the impact of multimodal composition. So far, few studies have measured the effects of multimodal composition on L2 writing skills, as studies in this domain focus on analyzing and describing the characteristics of the (collaborative) multimodal writing process or on identifying affordances of multimodal composition (e.g., Smith, 2013). A few studies, however, reported general learner and teacher perceptions of positive development in writing (e.g., Early & Marshall, 2008). Empirical evidence has been particularly scarce when it comes to contexts outside of the classroom. As a result, researchers have called for evaluative efforts concerning multimodal composition (Adawu, 2012; Ware & Hellmich, 2014).

The main purpose of this study, therefore, is to gain a better understanding of the differential effects of multimodal composition in in-school and out-of-school settings on adolescent beginners' L2 writing development. The research question guiding the study is the following:

Does the incorporation of meaningful multimodal writing activities in in-school and out-of-school settings differentially influence the development of writing skills by adolescent L2 learners of Dutch of low proficiency levels? If so, in what ways?

Based on the theoretical considerations and research described, we hypothesize that engaging learners in meaningful multimodal writing activities will have positive effects on learners' writing skills. More specifically, learners participating in multimodal writing in out-of-school and in-school settings will outperform those in a non-intervention group who do little writing and are without L2 contact during a similar period.

Furthermore, the literature review makes a case for differential effects in in-school and out-of-school settings, as affordances of these settings differ (Smith, 2013). The literature on multimodal composition, however, provides little guidance on the way in which these effects go. To hypothesize on the directions of differential effects, research on L2 acquisition and L2 writing has to be taken into account. Thus, we expect the following differential effects: First, the opportunities for meaning-making in multimodal composition in in-school settings may promote communicative effectiveness and improve the content of writing products. Additionally, the accuracy, complexity and lexical richness of the learners' output may be promoted in school, as a result of the balanced focus on collaborative meaning-making and attention paid to explicit knowledge and rule-based competence in in-school settings (Ellis & Shintani, 2014). Moreover, teachers' capacities for monitoring, scaffolding and pushing output further may enhance these particular aspects (Graham & Perin, 2007). Secondly, the rich, authentic meaning-making opportunities in writing in the out-of-school setting and the collaboration between peers may be expected to have a similar positive effect on communicative effectiveness and content. As meaningful language use outside of school settings seldom focuses on accuracy, we do not expect the out-of-school experiences to positively affect this aspect of writing. Regarding lexical richness and syntactic complexity, the available research is inconclusive: Some studies found meaning-focused tasks to contribute to lexical growth (e.g., Laufer & Hulstijn, 2001) and to the

use of more complex syntax (e.g., Long, 2014, as a result of negotiation for meaning), especially when learners were involved in the tasks at hand; other studies indicated that lexical and syntactic development required an explicit language focus (e.g., Skehan, 1998), which is more likely to occur in an in-class context.

Method

Participants

All participants in this study were adolescent newcomers, that is L2 learners of Dutch who had arrived in the city of Antwerp one to 13 months prior to the intervention. Before the intervention, all participants were in school. They were enrolled in reception classes, a specific and temporary form of education which aims to prepare newly-arrived migrants for Dutch-medium mainstream secondary education, by teaching the Dutch language and promoting social integration. During the summer holidays, many newcomers participate in summer schools, aiming to further enhance their Dutch L2 skills.

Out of 92 newcomers that were enrolled in a summer school program, 58 learners met three criteria which were directly relevant for this study. Firstly, we focused on learners from low socio-economic backgrounds (SES), because this group underachieves in the Flemish educational system (OECD, 2011) and because of the alleged advantages of multimodal writing for struggling writers. SES was measured by means of the educational level of the mother and was reported by most newcomers to be at primary school level or below. Only one fourth of the mothers completed secondary school. Secondly, to participate in meaning-making through multimodal composition, a minimal proficiency in Dutch was required, which was assessed by means of an oral picture-description task. This resulted in a rather heterogeneous learner group of mostly basic users with proficiency levels ranging from A1 to A2 (of the CEFR, Council of

Europe, 2001), as well as some of B1 proficiency considered as intermediate (independent users) on this scale. Thirdly, prior to the study, no participants had attended Flemish reception classes for more than 12 months. Finally, learners were between 12 and 16 years old.

All learners were randomly assigned to either of two interventions integrating multimodal composition: one a task-based in-school setting, the other an out-of-school program. Due to absence at times of assessment, only 52 (20 female) of the selected 58 learners were included in this study.

A non-intervention group of 32 newcomers was drawn from four intact classes from two schools in the Antwerp area. Learners in the non-intervention group met the same selection criteria, except for age: this group included five learners aged 17 and 18. However, this did not lead to a statistically significant difference in mean age between the non-intervention and intervention groups ($t(82) = 1.31$; $p = .20$). Furthermore, the non-intervention group participants had been attending Dutch reception classes for a statistically significantly longer time than those in the intervention groups ($t(82) = 7.51$; $p < .001$); this difference represents a small-sized effect ($r = .41$) according to field-specific benchmarks provided by Oswald and Plonsky (2014). By systematically estimating the effect of the two variables “time in Dutch reception classes” and “age”, we incorporated the age and instruction differences into the analysis of our data. Table 1 summarizes the means and standard deviations of the above-mentioned background variables for the learners in the three groups.

For all learners, also those in the non-intervention group, the quantity of outside-intervention/school opportunities for language learning was assessed by means of a questionnaire. The results indicated that all learners had little exposure to spoken Dutch outside the intervention or during their holidays, possibly owing to their low proficiency levels and little contact with

native speakers. Reading, watching television, or surfing the internet outside school was largely done in the L1.

Table 1: Background variables of participants

	Task-based instruction	Out-of-school project	Non-intervention group	Total
Number of participants	26	26	32	84
Country of origin	Afghanistan, Brazil, Bulgaria, Cameroon, Ecuador, Ethiopia, Freetown, Ghana, Guinea, Iraq, Kosovo, Morocco, Pakistan, Senegal, Sierra Leone, Somalia, Sri Lanka, Tibet	Afghanistan, Brazil, Burkina Faso, Cameroon, China, Djibouti, Guinea, Guatemala, India, Iraq, Iran, Liberia, Madagascar, Mongolia, Morocco, Nepal, Pakistan, Senegal, Sri Lanka, Togo, Turkey	Afghanistan, China, Ghana, Guinea, Hungary, Iraq, Morocco, Nepal, Pakistan, Portugal, Romania, Rwanda, Somalia, Spain, Sri Lanka, Turkey, Ukraine	30 countries in intervention groups, 6 additional countries in non-intervention group
Gender	F: 8, M: 18	F: 12, M: 14	F: 11, M: 21	F: 31, M: 53
Age	14.7 (1.04)	14.7 (1.10)	15.1 (1.60)	14.9 (1.30)
Dutch reception classes in months	3.75 (1.84)	4.15 (2.34)	8.13 (2.95)	5.54 (2.46)

Note: Standard deviations are given in parentheses.

Design

A pretest-posttest intervention with a comparative design was set up during a summer program. All participants were L2 learners from reception classes, who had not yet attended mainstream education. The school-based intervention took place inside a classroom and integrated L2 multimodal composition with a task-based approach. The out-of-school

intervention was set up in a leisure-camp environment. For both settings, the intervention was spread over two weeks. The L2 learners in the non-intervention group were also in reception classes and were tested before and after their two-week Easter holidays. During these holidays, they were in their L1 environment at home. This means they did not receive any training in Dutch and went largely without L2 contact. This lack of L2 contact was verified by means of a survey.

Treatment

The newcomers in the two interventions were followed for 42 hours spread over 10 days in two weeks. We used lapel microphones, which were worn by the supervisors (teachers or coaches), as well as by four learners, in each intervention. In addition, digital cameras were used to collect comprehensive shots. One camera offered a panoramic view of the classroom interaction, while a second was used to shoot close-up images of peer and student-supervisor interaction.

In both interventions, groups of newcomers were engaged in project work that led to a multimodal goal, that is the design of a website introducing future newcomers to Flanders. The newcomers generated texts, photos, music and videos to share online. The interventions also differed regarding the main interlocutor and his or her areas of expertise (teachers or youth workers/free-lance artists), and the autonomy with which learners could create multimodal products. These differences will be elaborated further for each of the conditions.

Multimodal composition in a task-based language class. To create optimal conditions for multimodal composition in school, the key features of task-based instruction and school-based multimodal composition were integrated. Therefore, we teamed up two teachers experienced in language teaching (five years of teaching experience or more) with two teacher trainees. We

provided a task-based syllabus with different tasks contributing to a project goal (a website) and a website environment to upload various multimodal products. Each teacher team was in charge of fifteen students, divided into an older (15 - 16 years) and younger (13 - 14 years) group.

Before the start of the intervention, a training session was organized in which the teachers worked around the principles and practices of task-based language teaching (TBLT). During this session, the teachers shared their experiences with TBLT, exchanged ideas on the use of the task-based syllabus and familiarized themselves with the aims of the project.

The teachers carried out all 18 tasks from the syllabus. Together, these tasks were geared to designing the website, incorporating seven multimodal compositions: an online quiz, an online memory game, a digital photo-comic, a video-based interview with Antwerp citizens on their life in Antwerp, an illustrated invitation, a video-commercial for the website, and pictures of Antwerp which incorporated contrasts. In Appendix A, screen shots of the website illustrate the multimodal products generated in school. While performing the tasks, the learners were supported by scaffolding activities for conceptualizing ideas: interactive brainstorming, reading other multimodal compositions, actively investigating features of Belgium, and reacting to each other's experiences of life in Antwerp. Revision was built in as well. Near the end of the intervention, participants were invited to select, revise and finalize the compositions that would be published on the website.

Multimodal composition in an out-of-school program. In the out-of-school project, we considered ways to promote learner motivation, meaningful, authentic communication, peer collaboration and learner autonomy in our choice of the project goal (designing the website), location (a youth club) and interlocutors (peers, youth workers and free-lance artists). The selection of four youth workers and four free-lance artists (subsequently labelled “coaches”) was

based on different criteria from that of the teachers: on their experience in youth work (two years or more), skills in web design, photography or film and absence of training in language pedagogy (to avoid typical teaching behavior such as monitoring, expert scaffolding or explicit instruction). To enhance learner autonomy, learners could self-select content, set goals and decide for themselves how to make use of the available resources in order to create this website. The newcomers generally worked in two age-related groups. One group comprised 13 students of 13 -14 years old, the other 14 students of 15 - 16 years old. However, because of their different interests, these groups often split up in order to work towards different end goals. Depending on group sizes, the newcomers were supported by one or more coaches, on a ratio of one coach per four newcomers approximately.

Before the start of the intervention, we discussed the website with the coaches. We asked the coaches to draft ideas for composition activities. In a brainstorming session that followed, the coaches came up with ideas such as collecting sounds of Antwerp, selecting participants' favorite music, introducing newcomers through written texts and drawings. Coaches could start from this brainstorm, but were aware of the learners' autonomy and the fact that activities should be challenging and full of opportunities for meaningful language use and collaboration. No explicit language learning goals were mentioned; the task-based intervention was not mentioned either. The coaches also formulated ideas for incorporating films or photographs of people's lives in Antwerp, be they imaginary (e.g., a love story, an action movie) or real (e.g., a documentary). Although the youth workers were not focused on "writing" activities, these ideas clearly involved numerous opportunities for multimodal composition. The activities that actually transpired included the shooting, selecting and arranging of the footage for the following multimodal compositions, in small groups: a documentary featuring newcomers in Antwerp who

discuss their life and hobbies; a love story of two newcomers in Antwerp; an action movie set in Antwerp; a compilation of pictures featuring different people the newcomers encountered in Antwerp. In Appendix B, several screen shots of the website made by the out-of-school groups are presented to illustrate these multimodal products.

Test Description

We used two traditional, non-multimodal, writing tasks (before and after the intervention) to measure academic writing development. These tests allowed us to assess writing skills by means of a format commonly valued by teachers and familiar to participants. Also, this format reduced the impact of the individual variation in newcomers' computer skills. Both tasks traced the development of academic writing skills. One involved a narrative genre, the other a persuasive genre (both of which are mentioned explicitly in the official school curriculum). The persuasive text was directly linked to the texts produced by the learners during the intervention (in terms of genre, topic and goal), while the narrative was not. We opted for two tasks because multiple products are needed to produce reliable and generalizable results (Bouwer, Béguin, Sanders, & van den Bergh, 2014; van den Bergh, De Maeyer, van Weijen, & Tillema, 2012). Both tests were administered on the workday before the start, and after the completion of the intervention and were identical at each time of measurement.

“Ice cream” (Nar). This narrative writing task was not related to the goal of the interventions and had been used in previous studies to investigate the writing development of child L2 learners of Dutch. It involves describing a story presented in non-verbal, comic mode (Verheyden, Van den Branden, Rijlaarsdam, van den Bergh, & De Maeyer, 2010). As in Verheyden et al.'s study, learners were asked to describe the events in the comic strip “Ice cream” (see Appendix C) to a person who is not familiar with the story (and not allowed to see the

comic). By presenting the content visually, the potential impact of learners' prior knowledge and creative skills is minimized.

“Folder of Antwerp” (Per). This persuasive writing task corresponds well with the content of the activities the newcomers carried out during the interventions. It requires learners to complete a tourist brochure of Antwerp (see Appendix D) by describing interesting things-to-see and providing information on means of transport. An introductory whole-class discussion took place in all groups and provided the participants with some ideas to complete this task. A pilot study with native speakers and L2 learners of Dutch indicated that the task was well-suited to distinguish between learners of different levels of language proficiency (Versteden, 2012).

Output Measures

To assess the quality of the participants' performance in the pretest and posttest, we used a variety of measures, tapping into different aspects of the performance, linguistic as well as content-based, quantitative as well as qualitative. Because of the heterogeneity in learner profiles and the open-ended tasks that were set, we selected general, rather than specific measures. All learners' written samples were transcribed using Codes for the Human Analysis of Transcripts (CHAT) (MacWhinney, 2000b), which enabled us to code and analyze for several measures by means of Computerized Language Analysis (CLAN) (MacWhinney, 2000a). Below, we report on the measures of performance we used: seven analytic measures and one score for communicative effectiveness.

Communicative effectiveness (Com. Eff). Communicative effectiveness assesses how well speakers achieve the communicative goal of the task (Pallotti, 2009). With the help of benchmarks, trained raters (six for the narrative, three for the persuasive task [as three raters were not available for the second rating session]) holistically scored the written products on a 7-

item-scale. Each text was rated twice. We calculated the means from these double ratings and for the few instances where disagreement between the raters was larger than two points (3% of the narrative products, 1% of the persuasive products), the first author re-assessed the products and used the nearest scores to calculate the means. A sample of 10% was scored by all raters. Krippendorff's α revealed good to excellent interrater reliability (16 products, $\alpha = .85$ for the narrative; 16 products, $\alpha = .92$ for the persuasive task).

Content: aptness and completeness (Cont). This measure captured the relevance and completeness of the content covered in the newcomers' descriptions. It quantifies whether newcomers mentioned crucial elements for achieving the main goal of the tasks [i.e., having a reader follow a story line (narrative) or getting useful tourist information on a given topic (persuasive)]. All crucial elements were described in advance in terms of propositions yielding a maximum score of 20 for the narrative task (Verheyden et al., 2010) and 12 for the persuasive task. "Content" was scored by two raters, with a sample of 15 % scored by both. Krippendorff's α for the narrative task was good (interval, 30 products, $\alpha = .85$). Krippendorff's α for the persuasive task, however, was marginally acceptable (interval, 32 products, $\alpha = .64$).

Lexical diversity (D). Lexical diversity (i.e., the quality of the vocabulary range) was operationalized by the measure "D" and was calculated with the help of the program VocD in Clan. Different from most type-token measures, D takes text length into account as it generates a type-token curve fitting model of lexical diversity. It has been shown to be effective in reducing effects of sample size (i.e., output length; Malvern & Richards, 2002). We selected D as measure of lexical diversity since the variation between and within learners in text sample sizes was large and only D can accommodate for such variation.

Syntactic complexity (MLT and CT). We included both length of T-units and subordination measures to accommodate for the multidimensionality of the construct (Norris & Ortega, 2009). Hence, we calculated the mean length of T-unit (MLT) and the number of clauses per T-unit (CT), two common types of syntactic complexity measures indicative of general grammatical complexity (Bulté & Housen, 2012; Wolfe-Quintero, Inagaki, & Kim, 1998).

Accuracy (Acc). Accuracy of the written texts was determined by using two types of overall accuracy scores. Following Wolfe-Quintero et al. (1998) and taking into account the lower proficiency levels of our learners, we opted for a measure of error-free production as well as for one that takes into account the number of errors. Both accuracy measures have been widely advocated as being sensitive to experimental conditions (Foster & Skehan, 1996; Vercellotti, 2012) and seemed appropriate for our purposes as no specific structures were targeted in either intervention. The first accuracy score is often used owing to ease of identification. It is based on marking the proportion of error-free T-units to the number of T-units (EFT). The identification of the total number of errors per T-unit (ET) for the second measure is more time-consuming. Importantly however, since we are dealing with basic learners, the second measure takes into account how errors are distributed and differentiates between units containing a single error, versus multiple errors, by analyzing how many errors occur in relation to T-unit production (Wolfe-Quintero Inagaki & Kim, 1998). In addition, this last measure has been claimed to easily pick up change (Polio, 2001; Tonkyn, 2012). Errors for both accuracy measures included lexical, morphological and syntactic errors and omissions of article, verb and subject (Michel, Kuiken, & Vedder, 2007). Lexical errors were included only when the choice of phrasing obscured word meaning. Errors in spelling and punctuation were ignored, given the learners' beginner proficiency levels.

Text Length (Len). We also calculated text length, operationalized as product fluency measured by text quantity (i.e., the total number of words written (Words)). This measure has been shown to be robust in the meta-analysis of different measures in writing (Wolfe-Quintero et al., 1998). Moreover, text length measured by number of words has been shown to distinguish well between different levels of writing ability and to document L2 learners' writing development (Sasaki, 2004).

Analysis

The objective of this study is to evaluate the differential effects of two settings, which constituted two different interventions, on L2 learners' development of writing skills. To this end, we collected two scores for each learner, on each task, and for each measure. One score was generated from the composition written before, and one from the composition written after the intervention took place. We used multilevel modelling (*MLM*) for the analysis of the data, as we considered both scores as "nested" within a learner and wanted to take this hierarchical structure into account. Also, MLM allowed us to abandon the assumption of equal variance at both times of measurement; in other words, it allows differences between learners to increase or decrease. Our choice for MLM has the added advantage of achieving greater power in fixed effects and having less restrictive assumptions for a valid assessment (Hoffman & Rovine, 2007). We used MLWin 2.31 (Rasbash, Charlton, Jones, & Pillinger, 2014) to analyze our data and to determine which model fitted the data best.

Model Construction

Following Barr, Levy, Scheepers, & Tily's (2013) recommendations on confirmatory hypothesis testing, we chose a multilevel model reflecting our research question concerning the effects of the two interventions. Our main interest lies in the evolution of writing scores in each

intervention, that is, in the fixed effect of the interaction between occasion and condition. To this end, we tested a multilevel model containing “occasion” (pretest and posttest), “condition” (no intervention, inside[-school] and outside[-school]) and the “interaction between occasion and condition” as fixed effects. In the random part of the model we allowed each learner to have a random intercept and random slope for occasion and thus for “growth” to vary individually.

Three additional control variables were included in our multilevel model to assess their potential influence: a centered measure for instruction in Dutch measured in months prior to the intervention, followed by a centered measure for age and the categorical variable “gender” (reference category = boy). Fit-of-model for the three control variables was examined through log-likelihood ratio, before its inclusion into or exclusion from the final model.¹

Additionally, we calculated the sizes of growth between tests, by dividing the difference between pre- and posttest estimates by the standard deviation based on the total variance. The results indicate the strength of associations. Following Plonsky and Oswald (2014), we will use the field-specific benchmarks of small ($d = .40$), medium ($d = .70$) and large ($d = 1.00$) with respect to the magnitude of an effect. However, we consider even small effects as noteworthy, since neither of the interventions focused on obtaining particular effects. (Compare to, for example, a small effect size for accurately forming the plural after explicit instruction on plural formation.)

Results

The final best-fit models are displayed in Table 2 and Table 3 and provide information on the effects of participation in the different settings on pretest and posttest.

¹ For each measure, we visually assessed that level-1 and level-2 residuals were normally distributed, the main statistical assumption of multilevel modeling.

Table 2 shows the estimates and standard errors for the eight measures of language development of the narrative writing task. To enhance transparency, statistically significant parameter estimates have been bolded.

Table 2: *Eight linguistic measures for the narrative writing task (Ice)*

	ComEff	Content	Syntactic complexity		Accuracy		Lexical diversity	Length
Measure	Hol_nar	Cont_nar	Ct_nar	Mlt_nar	Et_nar	Eft_nar	D_nar	Len_nar
Fixed Part								
Intercept	0.28 (0.23)	0.19 (0.17)	0.60 *** (0.17)	0.52 * (0.19)	-0.45 * (0.22)	0.40 (*) (0.22)	0.44 * (0.21)	0.25 (0.22)
Inside	-0.10 (0.29)	-0.48 (*) (0.26)	-0.86 *** (0.25)	-0.88 ** (0.28)	0.16 (0.29)	-0.29 (0.29)	-0.25 (0.27)	0.11 (0.28)
Outside	-0.24 (0.29)	-0.55 * (0.26)	-1.13 *** (0.25)	-0.99 *** (0.28)	0.22 (0.28)	-0.31 (0.28)	-0.61 * (0.26)	-0.28 (0.27)
Post-test	0.07 (0.10)	0.12 (0.10)	-0.17 (0.14)	-0.19 (0.13)	0.02 (0.11)	-0.06 (0.13)	0.02 (0.10)	-0.01 (0.10)
Post_inside	0.03 (0.15)	0.20 (0.15)	0.28 (0.21)	0.51 * (0.20)	-0.02 (0.16)	0.07 (0.19)	0.13 (0.15)	0.34 * (0.14)
Post_outside	0.13 (0.15)	0.26 (*) (0.15)	0.38 (*) (0.21)	0.45 * (0.20)	0.29 (*) (0.16)	-0.07 (0.19)	0.14 (0.15)	0.27 (*) (0.14)
Dutch_schooling	0.09 * (0.04)				-0.08 * (0.04)	0.09 * (0.04)	0.10 ** (0.04)	0.12 ** (0.04)
Age						-0.16 * (0.07)		
Boy	-0.37(*) (0.20)				0.45 * (0.20)	-0.31 (*) (0.19)	-0.35 (*) (0.19)	-0.46 * (0.19)
Random Part: Variances at pupil level								
Intercept variance	0.81 *** (0.13)	0.95 *** (0.15)	0.91 *** (0.14)	1.09 *** (0.17)	0.78 *** (0.12)	0.83 *** (0.13)	0.69 *** (0.11)	0.73 *** (0.11)
Covariance intercept & post-test	-0.11 * (0.06)	-0.18 ** (0.06)	-0.40 *** (0.09)	-0.52 *** (0.10)	-0.11 * (0.06)	-0.30 *** (0.08)	-0.09 * (0.05)	-0.08 * (0.05)
Post-test variance	0.32 *** (0.05)	0.32 *** (0.05)	0.62 *** (0.10)	0.57 *** (0.09)	0.36 *** (0.06)	0.53 *** (0.08)	0.34 *** (0.05)	0.29 *** (0.05)
-2*loglikelihood:	359.86	368.76	400.33	390.47	364.33	387.98	351.65	343.80
n units level 2	84	84	84	84	84	84	84	84
n units level 1	168	168	168	168	167	168	168	168

Note:

(a) Standardized MLM results for the following measures: holistically measured Communicative Effectiveness (Hol_nar), Aptness and completeness of Content (Con_nar), Syntactic complexity measured by subordination (Ct_nar) and by mean length of T-unit (Mlt_nar), Accuracy measured by Errors per T-unit (Et_nar) and by Errorfree production per T-unit (Eft_nar), Lexical diversity (D_nar) and Length as total words written (Len_nar).

(b) All scores are standardized-Z-scores, to allow for comparisons over different measures and tasks.

(c) Reference group = non-intervention group pretest scores.

(d) * Significant at the 0.05 level; ** Significant at the 0.01 level; *** Significant at the 0.001 level.

(e) (*) Non-significant trend towards significance at the 0.1 level;

(f) Standard errors are given in parentheses.

Table 3 displays the estimates and standard errors for each of the eight measures of language development on the persuasive task.

Table 3: Eight linguistic measures for the persuasive writing task (Folder)

	ComEff	Content	Syntactic complexity		Accuracy		Lexical diversity	Length
Measure	Hol_per	Cont_per	Ct_per	Mlt_per	Et_per	Eft_per	D_per	Len_per
Fixed Part								
Intercept	0.57 ** (0.16)	0.51 ** (0.16)	0.30 (0.18)	0.11 (0.18)	-0.44 ** (0.15)	0.48 ** (0.18)	0.33 (*) (0.17)	0.22 (0.17)
Inside	-1.13 *** (0.23)	-1.08 *** (0.24)	-0.31 (0.27)	-0.10 (0.27)	0.66 ** (0.23)	-0.77 ** (0.27)	-0.44 (*) (0.26)	-0.48 (*) (0.26)
Outside	-0.93 ** (0.23)	-1.12 *** (0.24)	-0.80 ** (0.27)	-0.43 (0.27)	0.46 * (0.23)	-0.63 * (0.27)	-0.79 ** (0.26)	-0.61 * (0.26)
Post-test	-0.08 (0.11)	0.05 (0.12)	-0.22 (0.20)	-0.26 (0.18)	0.12 (0.17)	-0.27 (0.18)	-0.22 (0.14)	-0.26 * (0.13)
Post_inside	0.53 *** (0.16)	0.55 ** (0.17)	0.32 (0.30)	0.50 (*) (0.26)	-0.20 (0.25)	0.47 (*) (0.26)	0.48 * (0.21)	0.99 *** (0.19)
Post_outside	0.20 (0.16)	0.41 * (0.17)	0.69 * (0.30)	0.68 ** (0.26)	0.41 (0.25)	0.06 (0.26)	0.60 ** (0.21)	0.62 *** (0.19)
Dutch_schooling						-0.12 (*) (0.07)	0.17 * (0.07)	0.14 * (0.07)
Age								
Boy								
Random Part: Variances at pupil level								
Intercept variance	0.78 *** (0.12)	0.81 *** (0.13)	1.07 *** (0.17)	1.01 *** (0.16)	0.74 *** (0.12)	1.01 *** (0.16)	0.95 *** (0.15)	0.92 *** (0.14)
Covariance intercept & post-test	-0.15 ** (0.06)	-0.25 *** (0.07)	-0.77 *** (0.15)	-0.55 *** (0.13)	-0.32 *** (0.10)	-0.61 *** (0.13)	-0.38 *** (0.09)	-0.29 *** (0.08)
Post-test variance	0.37 *** (0.06)	0.42 *** (0.07)	1.26 *** (0.19)	1.00 *** (0.15)	0.91 *** (0.14)	0.99 *** (0.15)	0.64 *** (0.10)	0.50 *** (0.08)
-2*loglikelihood:	363.80	370.16	454.00	447.51	430.50	437.85	410.80	394.75
n units level 2	84	84	84	84	84	84	84	84
n units level 1	168	168	168	168	168	168	168	168

Note:

(a) Standardized MLM results for the following measures: Holistically measured Communicative Effectiveness (Hol_per), Aptness and completeness of Content (Cont_per), Syntactic complexity measured by subordination (Ct_per) and by mean length of T-unit (Mlt_per), Accuracy measured by Errors per T-unit (Et_per), Lexical diversity (D_fol) and by Errorfree production per T-unit (Eft_per) and Length as total words written (Len_per).

(b) All scores are standardized-Z-scores, to allow for comparisons over different measures and tasks.

(c) Reference group = non-intervention group, pretest scores

(d) * Significant at the 0.05 level; ** Significant at the 0.01 level; *** Significant at the 0.001 level.

(e) (*) Non-significant trend towards significance at the 0.1 level;

(f) Standard errors are given in parentheses.

Table 4 displays the effect sizes for the narrative task.

Table 4: Narrative writing task: Effect sizes of gain

	Hol_Na	Cont_Na	Ct_Na	Mlt_Na	Inv_Et_Na	Eft_Na	D_Na	Len_Na
	r	r	r	r	r	r	r	r
No intervention	0.07	0.12	-0.20	-0.24	-0.03	-0.07	0.02	-0.01
Inside	0.09	0.33	0.15	0.41	-0.01	-0.01	0.13	0.30
Outside	0.21	0.39	0.24	0.32	-0.32	-0.15	0.17	0.27

Note:

- (a) In the table we combined information on the size of the effects, and the direction: in case of negative development, we marked the effect size with a '- '.
- (b) The effect sizes corresponding to significant differences between pretest and posttest are marked in bold.

Table 5 displays the effect sizes for the persuasive task.

Table 5: Persuasive writing task: Effect sizes of gain

	Hol_Per	Cont_Per	CT_Per	Mlt_Per	ET_Per	EFT_Per	D_Per	Len_Per
No intervention	-0.09	0.06	-0.25	-0.28	-0.12	-0.30	-0.24	-0.28
Inside	0.50	0.70	0.09	0.24	0.08	0.24	0.27	0.75
Outside	0.13	0.53	0.52	0.44	-0.53	-0.23	0.42	0.40

Note:

- (a) In the table we combined information on the size of the effects, and the direction: in case of negative development, we marked the effect size with a '- '.
- (b) The effect sizes corresponding to significant differences between pretest and posttest are marked in bold.

First, we will look into the learner background variables “Dutch-schooling,” “Age” and “Gender.” In Table 2 and Table 3, we find that the amount of time spent in Dutch reception classes before the intervention (Dutch_school) significantly contributes to the models, especially for accuracy, lexical diversity and length measures. In addition, in the narrative task, gender contributes significantly to the model for a few measures, as does age, for one measure: girls

perform better on communicative effectiveness, accuracy, lexical diversity and length, and younger students produce more accurate narratives.

Subsequently, we will examine the results for the other fixed parameters. In our model, the results on the pretest in the non-intervention group was chosen as a reference, represented in the table as “Intercept”. Significant results for the intercept indicate that the pretest scores in the non-intervention group are significantly different from zero. The rows labelled “Inside” and “Outside” show the results on the pretest, for each intervention group; significant results in these rows indicate that pretest scores are significantly different from the intercept. The rows labelled “Post-test”, “Post_inside” and “Post_outside” show the results we are most interested in: the gains on the posttest, for each group.

Writing Development in the Intervention Groups

In contrast to the status-quo or negative gains of the non-intervention group, both intervention groups showed significant growth. A total of six out of 16 measures showed statistically significant development for the in-school group. For four of these measures, this development is empirically supported by effect sizes attesting for a small or a medium-sized effect. Also for the out-of-school group, six measures showed statistically significant development. For five of these measures, this development is empirically supported by sizes attesting for small effects.

In the in-school intervention. The parameter estimates for Post-inside indicate significantly greater positive development for the learners in the in-school intervention than for those in the non-intervention group, for six output measures in total. Regarding the narrative task, Table 2 shows greater positive development for syntactic complexity measured by length ($Mlt_nar = 0.51, p < .05$) and text length measured by word total ($Len_nar = 0.34, p < .05$).

However, an effect size for gain was noted only in the first measure, and it was considered to be small (Mlt_nar, $d=.41$).

In Table 3, statistically significant greater development can be seen for four measures on the persuasive writing task: for communicative effectiveness (Hol_per= 0.53, $p<.001$) and the measures of content (Cont_per= 0.55, $p<.01$), lexical diversity (D_per= 0.48, $p<.05$) and length (Len_per= 0.99, $p<.001$). In Table 5, we see that three of these statistically significant gains are supported by effect sizes, that range from small for communicative effectiveness (Hol_per, $d=.50$) to medium effects for content (Cont_per, $d=.70$) and length (Len_per, $d=.75$).

In the out-of-school intervention. The parameter estimates for Post-outside indicate significantly greater positive development for the learners in the out-of-school intervention than for those in the non-intervention group, for six output measures.

Table 2 shows significant growth for syntactic complexity on the narrative writing test in the outside-school intervention (Mlt_nar= 0.45, $p<.05$). Table 4, however, shows the effect sizes of gains on the narrative test to be below .40, attesting for no effect, for the significant complexity development, as well as for three trends towards significance.

Table 3 shows statistically significant growth on the persuasive task for a total of five measures, that is, for all analytic measures, except for accuracy. There is a significant development for content (Cont_per= 0.41, $p<.05$), both measures of syntactic complexity (Mlt_per = 0.68, $p<.01$ and Ct_per= 0.69, $p<.05$) and lexical diversity (D_per= 0.60, $p<.01$) and text length (Len_per= 0.62, $p<.001$). Table 5 shows that the corresponding effect sizes are all above .40, but below .70, indicative of small effects.

Differential Effect of the In-school and Out-of-school interventions. To consider the differential effects of participation in the two interventions, we compared learner growth in both groups. We used Wald tests to compare the fixed estimates of "Post_inside" and "Post_outside," within the model. Table 6 shows the output of the chi-square tests, on the narrative task, followed by the persuasive task.

Table 6: Differential Language Gains

	gain inside	gain outside	χ^2	p-value		gain inside	gain outside	χ^2	p-value
Hol_nar	0.10	0.20	0.03	0.86	Hol_per	0.45	0.12	0.26	0.61
Cont_nar	0.32	0.37	0.00	0.97	Cont_per	0.60	0.46	0.60	0.44
Ct_nar	0.11	0.21	0.56	0.45	Ct_per	0.09	0.46	0.24	0.62
Mlt_nar	0.32	0.26	0.67	0.41	Mlt_per	0.24	0.42	0.33	0.57
Et_nar	0.00	0.31	1.88	0.17	Et_per	-0.08	0.53	2.22	0.14
Eft_nar	0.01	-0.13	0.37	0.55	Eft_per	0.21	-0.21	1.23	0.27
D_nar	0.14	0.15	1.80	0.18	D_per	0.26	0.38	0.85	0.36
Len_nar	0.33	0.21	3.13	0.08 (*)	Len_per	0.73	0.37	3.68	0.05 (*)

Note:

(a) *Results of the Wald tests comparing the growth in the task-based in-school group to the growth in the out-of-school group, for the eight measures, on the two tasks*

(b) (*) *Non-significant trend towards significance at the 0.1 level;*

(c) *N= 52*

Table 6 shows no significant differences in writing development between the in-school and out-of-school group. For no single measure, on either of the writing tasks, the difference in gain is statistically significant. Nevertheless, there is a trend towards a differential impact for text length, with greater gains for the in-school group apparent in both tasks (Len_nar, $\chi^2(1) = 3.13$, $p = 0.08$; Len_per, $\chi^2(1) = 3.68$, $p = 0.05$). The marginal significance for the narrative task is consistent with the absence of size effect for gains in both interventions ($d < .30$, for both interventions, Table 4). For the development of text length in the persuasive task, the trend

towards significance for differential growth is also reflected in the effect sizes reported earlier: small for the out-of-school ($d = .40$), but medium for the in-school intervention ($d = .75$).

Discussion

The main goal of the present study was to assess the differential impact of the integration of multimodal composition in an in-school, task-based intervention and in an out-of-school leisure program, on L2 learners' academic writing development. This study addresses a gap in the existing research by investigating measurable effects of multimodal composition on early L2 learning. In particular, this study is among the first to do so for out-of-school programs. Secondly, the range of aspects of L2 writing proficiency that was studied went beyond vocabulary acquisition, and also included the complexity and accuracy of output, text length and a measure of communicative effectiveness.

The results of the study show that the integration of multimodal composition positively affected adolescents' L2 writing skills. A confound of pre-test/post-test practice effects is unlikely, as no statistically significant growth was observed for the non-intervention group. Learners in both settings – in-school and out-of-school - showed significant growth in L2 writing compared to the non-intervention group. Moreover, significant positive growth is shown in both interventions for multiple measures of L2 writing development (i.e., communicative success, content, syntactic complexity, lexical diversity and text length).

Still, the reported gains were dependent on (1) the output measures and (2) the sort of tasks used to assess writing development. Firstly, there was considerable variation between the different measures of writing development. The greatest gains were observed for text length and syntactic complexity (measured by mean length of T-unit). As postulated, both greater text length and complexity gains might be linked with learners' enhanced motivation for writing in the L2 through multimodal composition, combined with the collaborative processes in which they were

involved. On the one hand, the positive experiences with multimodal writing may have promoted willingness to write and writing confidence in learners, resulting in longer texts. The possibility to learn from peers on the other hand, may have resulted in increased complexity. Previous studies have indicated the positive impact of motivation and authenticity on text length (Lo & Hyland, 2007; Mak & Coniam, 2008) and on syntactic complexity (Mak & Coniam, 2008). Neither of the interventions led to measurable accuracy development, even though accuracy gains are typically associated with language instruction. Possibly, the intervention was too short to lead to measurable accuracy gains. Learners' beginner proficiency levels or the focus of the interventions also could account for this. Accuracy development has indeed been linked to proficiency, as more accurate grammatical output only gradually develops in intermediate L2 stages (Ellis, 2009). It is very likely that the beginner learners had not yet managed to reach the intermediate L2 stage in this intervention. Alternatively, learners and teachers/coaches may have disregarded accuracy for the most part, as a strong emphasis on accuracy might have undermined positive features of multimodal composition, such as learner motivation (Lo & Hyland, 2007).

Secondly, learner development was task-dependent. Greater writing development and larger effect sizes were observed for the tests that matched the content of the activities in which learners were actively involved. More specifically, for communicative effectiveness, content and lexical diversity, significant growth was only shown on the test tasks that resembled the writing tasks the learners performed during the intervention. This is not a surprising finding, as limitations on transfer of learning have been demonstrated in numerous studies (James, 2009; Larsen-Freeman, 2013; Perkins & Salomon, 1994). This study, then, confirms that the similarity of content between a test and the tasks performed during an intervention influences research results.

The lack of differences between in-school and out-of-school interventions is somewhat surprising, as writing is traditionally associated with school and large meta-analyses point out the importance of instructional techniques such as strategy instruction, prewriting activities and process writing, among others (Graham & Perrin, 2007). Despite the clear differences in the design of both interventions, we only found trends towards significant differential development. This trend towards greater text length, on both tests, for learners in the in-school intervention might result from teachers pushing learners to write longer texts during the intervention (Mills, 2010b). One explanation for the lack of further differential effects might be that the characteristics of multimodal composition compensate for one another and the limitations of one setting, such as fewer instances of explicit instruction or scaffolding practices outside the classroom, may have been balanced by the higher levels of autonomy or engagement.

Another explanation lies in the methodological limitations of our study. Arguably, the instruments and the selection of measures or analyses that were chosen to accommodate for the heterogeneity and low proficiency levels of the group, may not have been suited to detect the differential effects of the two settings within as short a time span as two weeks. More fine-grained measures, or longer intervention periods might have done so.

Finally, even though both interventions were carefully prepared, they may have not been implemented according to plan. The beginner proficiency levels of the learners, for example, might in fact have led to less learner autonomy and more scaffolding practices in the out-of-school setting and/or less explicit instruction and metalanguage in the in-school setting. A cursory inspection of the video data collected for both interventions revealed that this might have been the case for some groups in our study. Further examination of the manner of implementation may shed greater light on this issue. All in all, this study supports Graham and

Perin's (2007) call to look further into innovative techniques for promoting writing skills, such as multimodal composition and to extend this research to contexts beyond the classroom.

Conclusion and Suggestions for Further Research

This study complements earlier studies on multimodal composition by (1) empirically investigating its effects on L2 writing development, (2) opting for multiple indices of language proficiency to measure progress, (3) assessing writing development with the use of more than one task and (4) by doing so with learners of beginner proficiency levels.

The findings indicate that integrating multimodal composition in settings in and out of school can promote multiple aspects of L2 writing development, even for low-proficiency learners. This seems especially the case for writing tasks that match the tasks performed during interventions in terms of content. Even though multimodal compositions may be approached differently in in-school and out-of-school settings, this does not necessarily lead to significant differences in L2 writing development.

Since this is the initial attempt to examine the effect of multimodal composition in different settings on writing development, the generalization of findings to other in-school and out-of-school settings and other learner populations would clearly be unwarranted and premature, especially since the group that served as a baseline (i.e., the non-intervention group) did not receive any writing practice. More research, involving other learners, of different proficiency levels, working with different teachers and coaches on different tasks and contrasted to other writing interventions, is needed. Furthermore, due to the comparison of groups with L2 practice (i.e., in in-school and out-of-school settings) to those without (i.e., in the non-intervention group), advantages of multimodal composition may have been artificially boosted. An additional

comparison to groups experiencing traditional writing instruction would therefore add to our understanding of multimodal composition's added value for writing development.

To measure the differential effects of interventions involving multimodal composition more fine-grained instruments may need to be developed. In addition, the length of the intervention needs to be varied in follow-up studies. It is also important to gain insight into the role that the coaches and teachers played in fostering the learners' L2 writing development.

A final limitation and possible area for future research lies in the choice to measure effects in terms of writing development. Learners nowadays are required to develop 21st century skills and competences, which for example also include the development of digital literacies or the ability to combine multiple modes effectively in a single text. Additionally, other data sources (e.g., students' perceptions and writing processes) may be used to further examine the effects of multimodal writing in both settings. Further research is necessary to provide more insight into the full potential of multimodal composition in terms of promoting L2 acquisition and writing.

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. <http://doi.org/10.1016/j.jml.2012.11.001>
- Black, R. W. (2005). Access and affiliation: The literacy and composition practices of English-language learners in an online fanfiction community. *Journal of Adolescent & Adult Literacy*, 49, 118–128. <http://doi.org/10.1598/JAAL.49.2.4>
- Bloch, J. (2007). Abdullah's blogging: A Generation 1.5 student enters the blogosphere. *Language Learning & Technology*, 11(2), 128–141.
- Bouwer, R., Béguin, A., Sanders, T., & van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing*, 32, 83-100. <https://doi.org/10.1177/0265532214542994>
- Bruce, D. L. (2009). Writing with visual images: Examining the video composition processes of high school students. *Research in the Teaching of English*, 43, 426–450.
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (Vol. 32, pp. 21–46). Amsterdam: John Benjamins.
- Chapelle, C. A. (2005). Interactionist SLA theory in CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 53–64). Mahwah, NJ: Lawrence Erlbaum.
- Chávez, V., & Soep, E. (2005). Youth radio and the pedagogy of collegiality. *Harvard Educational Review*, 75, 409–434.
- Cope, B., & Kalantzis, M. (2000). *Multiliteracies: Literacy learning and the design of social futures*. London: Routledge.

- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press. Retrieved from <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680459f97>
- Dal, M. (2010). Digital video production and task based language learning. *Ráðstefnurit Netlu – Menntakvika 2010*. Retrieved from <http://skemman.is/item/view/1946/7848>
- Early, M., & Marshall, S. (2008). Adolescent ESL students' interpretation and appreciation of literary texts: A case study of multimodality. *Canadian Modern Language Review*, 64, 377–397.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33, 209–224. <http://doi.org/10.1016/j.system.2004.12.006>
- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19, 221–246.
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. New York, NY: Routledge.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323. <http://doi.org/10.1017/S0272263100015047>
- Fuente, M. J. (2003). Is SLA interactionist theory relevant to CALL? A study on the effects of computer-mediated interaction in L2 vocabulary acquisition. *Computer Assisted Language Learning*, 16, 47–81. <http://doi.org/10.1076/call.16.1.47.15526>
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Portsmouth, NH: Heineman.

- Gilje, Ø. (2010). Multimodal redesign in filmmaking practices: An inquiry of young filmmakers' deployment of semiotic tools in their filmmaking practice. *Written Communication*, 27, 494–522. <http://doi.org/10.1177/0741088310377874>
- González-Lloret, M., & Ortega, L. (2014). *Technology-mediated TBLT: Researching technology and tasks*. Amsterdam: John Benjamins.
- Graham, S., & Perin, D. (2007). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading*, 11, 313–335.
<http://doi.org/10.1080/10888430701530664>
- Hafner, C. A., Chik, A., & Jones, R. H. (2015). Digital literacies and language learning. *Language Learning and Technology*, 19(3), 1-7.
- Halverson, E. R. (2010). Film as identity exploration: A multimodal analysis of youth-produced films. *Teachers College Record*, 112, 2352–2378.
- Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, 39, 101–117.
<http://doi.org/10.3758/BF03192848>
- Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. *Review of Educational Research*, 71, 575–611.
<http://doi.org/10.3102/00346543071004575>
- Ito, M., Horst, H., Bittanti, M., Boyd, D., Herr-Stephenson, B., Lange, P. G., ... Robinson, L. (2008). *Living and learning with new media: Summary of findings from the digital youth project* (ED536072). Retrieved from <http://eric.ed.gov/?id=ED536072>
- James, M. A. (2009). “Far” transfer of learning outcomes from an ESL writing course: Can the gap be bridged? *Journal of Second Language Writing*, 18, 69–84.
<https://doi.org/10.1016/j.jslw.2009.01.001>

- Jenkins, H. (2006). *Confronting the challenges of participatory culture: Media education for the 21st Century*. An occasional paper on digital media and learning (ED536086). Retrieved from <http://eric.ed.gov/?id=ED536086>
- Jewitt, C. (2006). *Technology, literacy and learning: A multimodal approach*. London, UK: Routledge.
- Jiang, L., & Luk, J. (2016). Multimodal composing as a learning activity in English classrooms: Inquiring into the sources of its motivational capacity. *System*, 59, 1-11.
- Karlsson, A. (2015). *How can blogging in foreign language education improve pupils' writing skills? A research synthesis* (Master's thesis). Retrieved from <https://dspace.mah.se/handle/2043/18709>
- Kelly, C. (2012). The cafeteria as contact zone: Developing a multicultural perspective through multilingual and multimodal literacies. *Journal of Adolescent & Adult Literacy*, 56, 301–310. <http://doi.org/10.1002/JAAL.00143>
- Kern, R., & Schultz, J. M. (2005). Beyond orality: Investigating literacy and the literary in second and foreign language instruction. *The Modern Language Journal*, 89, 381–392.
- Kessler, G. (2007). Formal and informal CALL preparation and teacher attitude toward technology. *Computer Assisted Language Learning*, 20, 173–188. <http://doi.org/10.1080/09588220701331394>
- Lammers, J. C., Magnifico, A. M., & Curwood, J. S. (2014). Exploring tools, places, and ways of being: Audience matters for developing writers. In K. E. Pytash & R. E. Ferdig (Eds.), *Exploring technology for writing and writing instruction* (pp. 186–201). Hershey, PA: IGI Global.

- Larsen-Freeman, D. (2013). Transfer of learning transformed. *Language Learning*, 63, 107–129.
<http://doi.org/10.1111/j.1467-9922.2012.00740.x>
- Lenhart, A., Arafeh, S., & Smith, A. (2008). *Writing, technology and teens*. Pew Internet & American Life Project (ED524313). Retrieved from <https://eric.ed.gov/?id=ED524313>
- Li, M. (2013). Individual novices and collective experts: Collective scaffolding in wiki-based small group writing. *System*, 41, 752–769. <https://doi.org/10.1016/j.system.2013.07.021>
- Lo, J., & Hyland, F. (2007). Enhancing students' engagement and motivation in writing: The case of primary students in Hong Kong. *Journal of Second Language Writing*, 16, 219–237.
- Lotherington, H., Neville-Verardi, D., & Ronda, N. S. (2009). English in cyberspace: Negotiating digital literacies in a climate of educational accountability. In L. B. Abraham & L. Williams (Eds.), *Electronic discourse in language learning and language teaching* (pp. 11–42). Amsterdam: John Benjamins.
- MacWhinney, B. (2000a). *The CHILDES project: Tools for analyzing talk. Part 1: The CHAT transcription format* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. (2000b). *The CHILDES project: Tools for analyzing talk. Part 2: The CLAN programs* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Magnifico, A. M. (2010). Writing for whom? Cognition, motivation, and a writer's audience. *Educational Psychologist*, 45, 167–184. <http://doi.org/10.1080/00461520.2010.493470>
- Mak, B., & Coniam, D. (2008). Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System*, 36, 437–455.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19, 85–104.
<http://doi.org/10.1191/0265532202lt221oa>

- Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, 45, 241–259.
- Miller, S. M. (2013). A research metasynthesis on digital video composing in classrooms An evidence-based framework toward a pedagogy for embodied learning. *Journal of Literacy Research*, 45, 386–430. <http://doi.org/10.1177/1086296X13504867>
- Mills, K. A. (2010a). A review of the “digital turn” in the New Literacy Studies. *Review of Educational Research*, 80, 246–271. <http://doi.org/10.3102/0034654310364401>
- Mills, K. A. (2010b). Shrek meets Vygotsky: Rethinking adolescents' multimodal literacy practices in schools. *Journal of Adolescent & Adult Literacy*, 54, 35-45.
- Murray, D. E. (2005). Technologies for second language literacy. *Annual Review of Applied Linguistics*, 25, 188–201. <http://doi.org/10.1017/S0267190505000103>
- Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, 14, 397–419.
- Niiya, M., Warschauer, M., & Zheng, B. (2013). Emerging literacies in digital media and L2 secondary writing. In L. C. de Oliveira & T. Silva (Eds.), *L2 writing in secondary classrooms: Student experiences, academic issues, and teacher education* (pp. 104-116. New York, NY: Routledge.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
<http://doi.org/10.1111/0023-8333.00136>

- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578.
<http://doi.org/10.1093/applin/amp044>
- OECD (2011), *School Evaluation in the Flemish Community of Belgium 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing.
<http://dx.doi.org/10.1787/9789264116726-en>
- Ohler, J. B. (2013). *Digital storytelling in the classroom: New media pathways to literacy, learning, and creativity* (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Oskoz, A., & Elola, I. (2014). Integrating digital stories in the writing class: Toward a 21st-century literacy. In J. P. Guikema & L. Williams (Eds.), *Digital literacies in foreign and second language education* (pp. 179–200). San Marcos, TX: CALICO.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30, 590–601. <http://doi.org/10.1093/applin/amp045>
- Perkins, D. N., & Salomon, G. (1994). Transfer of learning. In T. Husen & T. N. Postlethwait (Eds.), *International encyclopedia of education* (2nd ed., Vol. 11, pp. 6452–6457). Oxford: Pergamon Press.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.
- Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 91–116). Mahwah, NJ: Lawrence Erlbaum.

Purcell-Gates, V., Anderson, J., Gagne, M., Jang, K., Lenters, K. A., & McTavish, M. (2012).

Measuring situated literacy activity: Challenges and promises. *Journal of Literacy*

Research, 44, 396-425. <http://doi.org/10.1177/1086296X12457167>

Rasbash, J., Charlton, C., Jones, K., & Pillinger, R. (2014). *Manual supplement for MLwiN*

Version 2.31. University of Bristol. Retrieved from

<http://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/2-33/supplement-print.pdf>

Sadik, A. (2008). Digital storytelling: A meaningful technology-integrated approach for engaged student learning. *Educational Technology Research and Development*, 56, 487–506.

<http://doi.org/10.1007/s11423-008-9091-8>

Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers.

Language Learning, 54, 525–582. <http://doi.org/10.1111/j.0023-8333.2004.00264.x>

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268–286.

<http://doi.org/10.1017/S0267190500003585>

Smith, B. E. (2014). Beyond words: A review of research on adolescents and multimodal

composition. In R. E. Ferdig & K. E. Pytash (Eds.), *Exploring multimodal composition and digital writing* (pp. 1–19). Hershey, PA: IGI Global.

Steeves, N. (2015). *Fairytales in secondary English classrooms? Exploring multimodality, visual literacy and student resistance* (Doctoral dissertation). Retrieved from

<http://dspace.library.uvic.ca:8080/handle/1828/6705>

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

<http://doi.org/10.1093/applin/16.3.371>

- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (99-118). Longman, Pearson Education. Retrieved from: <http://hdl.handle.net/10315/2820>.
- Sylvester, R., & Greenidge, W. (2009). Digital storytelling: Extending the potential for struggling writers. *The Reading Teacher*, 63, 284–295. <http://doi.org/10.1598/RT.63.4.3>
- Tan, J. P.-L., & McWilliam, E. (2009). From literacy to multiliteracies: Diverse learners and pedagogical practice. *Pedagogies: An International Journal*, 4, 213–225. <http://doi.org/10.1080/15544800903076119>
- Thorne, S. L., & Black, R. W. (2007). Language and literacy development in computer-mediated contexts and communities. *Annual Review of Applied Linguistics*, 27, 133–160. <http://doi.org/10.1017/S0267190508070074>
- Thorne, S. L., Fischer, I., & Lu, X. (2012). The semiotic ecology and linguistic complexity of an online game world. *ReCALL*, 24, 279–301. <http://doi.org/10.1017/S0958344012000158>
- Tonkyn, A. (2012). Measuring and perceiving changes in oral complexity, accuracy and fluency: Examining instructed learners' short-term gains. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 221–245). Amsterdam: John Benjamins.
- van den Bergh, H., De Maeyer, S., van Weijen, D., & Tillema, M. (2012). Generalizability of text quality scores. In E. van Steendam, M. Tillema, G. Rijlaarsdam, & H. van den Bergh (Eds.), *Measuring writing: Recent insights into theory, methodology and practices*, (pp. 23–32). Leiden: Brill.

- Van den Branden, K., Bygate, M., & Norris, J. M. (Eds.). (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamins.
- VanKooten, C. (2016). "The video was what did it for me": Developing meta-awareness about composition across media. *College English*, 79, 57-80.
- Vercellotti, M. L. (2012). *Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other* (Unpublished doctoral dissertation). University of Pittsburgh. Retrieved from http://d-scholarship.pitt.edu/12071/1/Vercellotti_CAF_v3.pdf
- Verheyden, L., Van den Branden, K., Rijlaarsdam, G., van den Bergh, H., & De Maeyer, S. (2010). Written narrations by 8- to 10-year-old Turkish pupils in Flemish primary education: A follow-up of seven text features. *Journal of Research in Reading*, 33, 20–38. <http://doi.org/10.1111/j.1467-9817.2009.01430.x>
- Ware, P., & Hellmich, E. (2014). CALL in the K-12 context: Language learning outcomes and opportunities. *CALICO Journal*, 31, 140-157. <http://dx.doi.org/10.11139/cj.31.2.140-157>
- Warschauer, M. (1996). Motivational aspects of using computers for writing and communication. In M. Warschauer (Ed.), *Telecollaboration in foreign language learning: Proceedings of the Hawai'i symposium* (pp. 29–46). Second Language Teaching & Curriculum Centre, University of Hawai'i at Mānoa.
- Warschauer, M. (2009). Digital literacy studies: Progress and prospects. In M. Baynham & M. Prinsloo (Eds.), *The future of literacy studies* (pp. 123–40). New York, NY: Palgrave Macmillan.

- Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education*, 34, 179–225. <http://doi.org/10.3102/0091732X09349791>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. National Foreign Language Resource Center, University of Hawai‘i at Mānoa.
- Yang, Y.-T. C., & Wu, W.-C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & Education*, 59, 339–352. <http://doi.org/10.1016/j.compedu.2011.12.012>

Table 1: Background variables of participants

	Task-based instruction	Out-of-school project	Non-intervention group	Total
Number of participants	26	26	32	84
Country of origin	Afghanistan, Brazil, Bulgaria, Cameroon, Ecuador, Ethiopia, Freetown, Ghana, Guinea, Iraq, Kosovo, Morocco, Pakistan, Senegal, Sierra Leone, Somalia, Sri Lanka, Tibet	Afghanistan, Brazil, Burkina Faso, Cameroon, China, Djibouti, Guinea, Guatemala, India, Iraq, Iran, Liberia, Madagascar, Mongolia, Morocco, Nepal, Pakistan, Senegal, Sri Lanka, Togo, Turkey	Afghanistan, China, Ghana, Guinea, Hungary, Iraq, Morocco, Nepal, Pakistan, Portugal, Romania, Rwanda, Somalia, Spain, Sri Lanka, Turkey, Ukraine	30 countries in intervention groups, 6 additional countries in non-intervention group
Gender	F: 8, M: 18	F: 12, M: 14	F: 11, M: 21	F: 31, M: 53
Age	14.7 (1.04)	14.7 (1.10)	15.1 (1.60)	14.9 (1.30)
Dutch reception classes in months	3.75 (1.84)	4.15 (2.34)	8.13 (2.95)	5.54 (2.46)

Note: Standard deviations are given in parentheses.

Table 2: *Eight linguistic measures for the narrative writing task (Ice)*

	ComEff	Content	Syntactic complexity		Accuracy		Lexical diversity	Length
Measure	Hol_nar	Cont_nar	Ct_nar	Mlt_nar	Et_nar	Eft_nar	D_nar	Len_nar
Fixed Part								
Intercept	0.28 (0.23)	0.19 (0.17)	0.60 *** (0.17)	0.52 * (0.19)	-0.45 * (0.22)	0.40 (*) (0.22)	0.44 * (0.21)	0.25 (0.22)
Inside	-0.10 (0.29)	-0.48 (*) (0.26)	-0.86 *** (0.25)	-0.88 ** (0.28)	0.16 (0.29)	-0.29 (0.29)	-0.25 (0.27)	0.11 (0.28)
Outside	-0.24 (0.29)	-0.55 * (0.26)	-1.13 *** (0.25)	-0.99 *** (0.28)	0.22 (0.28)	-0.31 (0.28)	-0.61 * (0.26)	-0.28 (0.27)
Post-test	0.07 (0.10)	0.12 (0.10)	-0.17 (0.14)	-0.19 (0.13)	0.02 (0.11)	-0.06 (0.13)	0.02 (0.10)	-0.01 (0.10)
Post_inside	0.03 (0.15)	0.20 (0.15)	0.28 (0.21)	0.51 * (0.20)	-0.02 (0.16)	0.07 (0.19)	0.13 (0.15)	0.34 * (0.14)
Post_outside	0.13 (0.15)	0.26 (*) (0.15)	0.38 (*) (0.21)	0.45 * (0.20)	0.29 (*) (0.16)	-0.07 (0.19)	0.14 (0.15)	0.27 (*) (0.14)
Dutch_schooling	0.09 * (0.04)				-0.08 * (0.04)	0.09 * (0.04)	0.10 ** (0.04)	0.12 ** (0.04)
Age						-0.16 * (0.07)		
Boy	-0.37(*) (0.20)				0.45 * (0.20)	-0.31 (*) (0.19)	-0.35 (*) (0.19)	-0.46 * (0.19)
Random Part: Variances at pupil level								
Intercept variance	0.81 *** (0.13)	0.95 *** (0.15)	0.91 *** (0.14)	1.09 *** (0.17)	0.78 *** (0.12)	0.83 *** (0.13)	0.69 *** (0.11)	0.73 *** (0.11)
Covariance intercept & post-test	-0.11 * (0.06)	-0.18 ** (0.06)	-0.40 *** (0.09)	-0.52 *** (0.10)	-0.11 * (0.06)	-0.30 *** (0.08)	-0.09 * (0.05)	-0.08 * (0.05)
Post-test variance	0.32 *** (0.05)	0.32 *** (0.05)	0.62 *** (0.10)	0.57 *** (0.09)	0.36 *** (0.06)	0.53 *** (0.08)	0.34 *** (0.05)	0.29 *** (0.05)
-2*loglikelihood:	359.86	368.76	400.33	390.47	364.33	387.98	351.65	343.80
n units level 2	84	84	84	84	84	84	84	84
n units level 1	168	168	168	168	167	168	168	168

Note:

(a) Standardized MLM results for the following measures: holistically measured Communicative Effectiveness (Hol_nar), Aptness and completeness of Content (Con_nar), Syntactic complexity measured by subordination (Ct_nar) and by mean length of T-unit (Mlt_nar), Accuracy measured by Errors per T-unit (Et_nar) and by Errorfree production per T-unit (Eft_nar), Lexical diversity (D_nar) and Length as total words written (Len_nar).

(b) All scores are standardized-Z-scores, to allow for comparisons over different measures and tasks.

(c) Reference group = non-intervention group pretest scores.

(d) * Significant at the 0.05 level; ** Significant at the 0.01 level; *** Significant at the 0.001 level.

(e) (*) Non-significant trend towards significance at the 0.1 level.

(f) Standard errors are given in parentheses.

Table 3: Eight linguistic measures for the persuasive writing task (Folder)

	ComEff	Content	Syntactic complexity		Accuracy		Lexical diversity	Length
Measure	Hol_per	Cont_per	Ct_per	Mlt_per	Et_per	Eft_per	D_per	Len_per
Fixed Part								
Intercept	0.57 ** (0.16)	0.51 ** (0.16)	0.30 (0.18)	0.11 (0.18)	-0.44 ** (0.15)	0.48 ** (0.18)	0.33 (*) (0.17)	0.22 (0.17)
Inside	-1.13 *** (0.23)	-1.08 *** (0.24)	-0.31 (0.27)	-0.10 (0.27)	0.66 ** (0.23)	-0.77 ** (0.27)	-0.44 (*) (0.26)	-0.48 (*) (0.26)
Outside	-0.93 ** (0.23)	-1.12 *** (0.24)	-0.80 ** (0.27)	-0.43 (0.27)	0.46 * (0.23)	-0.63 * (0.27)	-0.79 ** (0.26)	-0.61 * (0.26)
Post-test	-0.08 (0.11)	0.05 (0.12)	-0.22 (0.20)	-0.26 (0.18)	0.12 (0.17)	-0.27 (0.18)	-0.22 (0.14)	-0.26 * (0.13)
Post_inside	0.53 *** (0.16)	0.55 ** (0.17)	0.32 (0.30)	0.50 (*) (0.26)	-0.20 (0.25)	0.47 (*) (0.26)	0.48 * (0.21)	0.99 *** (0.19)
Post_outside	0.20 (0.16)	0.41 * (0.17)	0.69 * (0.30)	0.68 ** (0.26)	0.41 (0.25)	0.06 (0.26)	0.60 ** (0.21)	0.62 *** (0.19)
Dutch_schooling						-0.12 (*) (0.07)	0.17 * (0.07)	0.14 * (0.07)
Age								
Boy								
Random Part: Variances at pupil level								
Intercept variance	0.78 *** (0.12)	0.81 *** (0.13)	1.07 *** (0.17)	1.01 *** (0.16)	0.74 *** (0.12)	1.01 *** (0.16)	0.95 *** (0.15)	0.92 *** (0.14)
Covariance intercept & post-test	-0.15 ** (0.06)	-0.25 *** (0.07)	-0.77 *** (0.15)	-0.55 *** (0.13)	-0.32 *** (0.10)	-0.61 *** (0.13)	-0.38 *** (0.09)	-0.29 *** (0.08)
Post-test variance	0.37 *** (0.06)	0.42 *** (0.07)	1.26 *** (0.19)	1.00 *** (0.15)	0.91 *** (0.14)	0.99 *** (0.15)	0.64 *** (0.10)	0.50 *** (0.08)
-2*loglikelihood:	363.80	370.16	454.00	447.51	430.50	437.85	410.80	394.75
n units level 2	84	84	84	84	84	84	84	84
n units level 1	168	168	168	168	168	168	168	168

Note:

(a) Standardized MLM results for the following measures: Holistically measured Communicative Effectiveness (Hol_per), Aptness and completeness of Content (Cont_per), Syntactic complexity measured by subordination (Ct_per) and by mean length of T-unit (Mlt_per), Accuracy measured by Errors per T-unit (Et_per), Lexical diversity (D_fol) and by Errorfree production per T-unit (Eft_per) and Length as total words written (Len_per).

(b) All scores are standardized-Z-scores, to allow for comparisons over different measures and tasks.

(c) Reference group = non-intervention group, pretest scores

(d) * Significant at the 0.05 level; ** Significant at the 0.01 level; *** Significant at the 0.001 level.

(e) (*) Non-significant trend towards significance at the 0.1 level.

(f) Standard errors are given in parentheses.

Table 4: *Narrative writing task: Effect sizes of gain*

	Hol_Nar	Cont_Nar	Ct_Nar	Mlt_Nar	Inv_Et_Nar	Eft_Nar	D_Nar	Len_Nar
No intervention	0.07	0.12	-0.20	-0.24	-0.03	-0.07	0.02	-0.01
Inside	0.09	0.33	0.15	0.41	-0.01	-0.01	0.13	0.30
Outside	0.21	0.39	0.24	0.32	-0.32	-0.15	0.17	0.27

Note:

(a) *In the table we combined information on the size of the effects, and the direction: in case of negative development, we marked the effect size with a ‘-’.*

(b) *The effect sizes corresponding to significant differences between pretest and posttest are marked in bold.*

Table 5: Persuasive writing task: Effect sizes of gain

	Hol_Per	Cont_Per	CT_Per	Mlt_Per	ET_Per	EFT_Per	D_Per	Len_Per
No intervention	-0.09	0.06	-0.25	-0.28	-0.12	-0.30	-0.24	-0.28
Inside	0.50	0.70	0.09	0.24	0.08	0.24	0.27	0.75
Outside	0.13	0.53	0.52	0.44	-0.53	-0.23	0.42	0.40

Note:

(a) In the table we combined information on the size of the effects, and the direction: in case of negative development, we marked the effect size with a '-'.
(b) The effect sizes corresponding to significant differences between pretest and posttest are marked in bold.

Table 6: Differential Language Gains

	gain inside	gain outside	χ^2	p-value		gain inside	gain outside	χ^2	p-value
Hol_nar	0.10	0.20	0.03	0.86	Hol_per	0.45	0.12	0.26	0.61
Cont_nar	0.32	0.37	0.00	0.97	Cont_per	0.60	0.46	0.60	0.44
Ct_nar	0.11	0.21	0.56	0.45	Ct_per	0.09	0.46	0.24	0.62
Mlt_nar	0.32	0.26	0.67	0.41	Mlt_per	0.24	0.42	0.33	0.57
Et_nar	0.00	0.31	1.88	0.17	Et_per	-0.08	0.53	2.22	0.14
Eft_nar	0.01	-0.13	0.37	0.55	Eft_per	0.21	-0.21	1.23	0.27
D_nar	0.14	0.15	1.80	0.18	D_per	0.26	0.38	0.85	0.36
Len_nar	0.33	0.21	3.13	0.08 (*)	Len_per	0.73	0.37	3.68	0.05 (*)

Note:
(a) Results of the Wald tests comparing the growth in the task-based in-school group to the growth in the out-of-school group, for the eight measures, on the two tasks
(b) (*) Non-significant trend towards significance at the 0.1 level;
(c) N= 52

Author's cv

Goedele Vandommele, University of Leuven, goedele.vandommele@kuleuven.be

Goedele Vandommele manages research projects at the Centre for Language and Education at University Leuven. Her main research interest lies in second language acquisition and task-based language learning. Before obtaining her PhD, Goedele worked on various projects related to materials development, in-service teacher training and research, all on promoting language learning and adolescent literacy. Through this work, she developed an interest in second language acquisition, particularly with adolescents.

Kris Van den Branden, University of Leuven, kris.vandenbranden@kuleuven.be

Kris Van den Branden is a professor of linguistics and teacher educator at University of Leuven. He also is director of the Centre for Language and Education. This centre combines research activities with in-service-training. His main research interests have to do with the teacher's role in language teaching, task-based language teaching, and equity and excellence in education.

Koen Van Gorp, Michigan State University, vangorpk@msu.edu.

Koen Van Gorp is Head of Foreign Language Assessment at Center for Language Teaching Advancement, Michigan State University. Previously, he managed research projects at the Centre for Language and Education at University Leuven. In this centre, he coordinated the Certificate for Dutch as a Foreign language (CNAVt), a standardized test to measure the Dutch language ability of non-native speakers. His research is focused on second language acquisition and testing, task-based language teaching, L2 knowledge construction.

Sven De Maeyer, University of Antwerp, sven.demaeyer@uantwerpen.be

Sven De Maeyer is a professor in Educational Sciences, with expertise in Research Methodology and Statistics at University of Antwerp. His research is focused on measurement problems in educational sciences, in particular on validity and reliability in (school)effectiveness research.

In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch

Highlights

- Collaborative Multimodal Writing resulted in L2 academic writing development.
- Learners gained in content, communicative success, syntax, vocabulary, text length.
- No accuracy effects were found.
- Different contexts (in-school and out-of-school settings) did not produce differential gains.

STARTPAGINA!

DIT IS DE WEBSITE!

HEY, DIT BEN IK.

FOTO'S!!

WAT VINDEN DE BELGEN BELANGRIJK?

MORE...

De leerlingen stellen de website voor, op HUN manier!

Hieronder vind je links naar Youtube filmpjes waarin de leerlingen de website voorstellen. Bekijk ze allemaal!

Alles wat je moet weten over de website:

Toneel op de speelplaats.

Zomerklap! De waterklas stelt de website voor. episode

Zomerklap! De waterklas stelt de website voor. episode

Het nieuwsinterview

Zomerklap! De waterklas stelt de website voor. episode

De wereldbekende rap.

Zomerklap! De waterklas stelt de website voor. episode

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Vertel het verhaal uit de strip aan iemand die de prenten NIET ziet.

Wat gebeurt er op de tekeningen? Schrijf op!

1

2

3

4

5

6

7

8

1 MUSEUM AAN DE STROOM (MAS)

2

Deze folder werd geschreven door:

Antwerpen is de grootste stad van Vlaanderen. Het is de stad aan de rivier de Schelde, de stad van de koffiekoeken, de stad van vele culturen... Een fijne stad om in te wonen, maar ook om te bezoeken!