

## Source use and argumentation behavior in L1 and L2 writing: a within-writer comparison

Daphne van Weijen<sup>1</sup>  · Gert Rijlaarsdam<sup>1,2,3</sup> · Huub van den Bergh<sup>4</sup>

Published online: 12 April 2018  
© The Author(s) 2018

**Abstract** The aim of this study was to test whether Cummins' Linguistic Interdependence Hypothesis (LIH) might also apply to writing, by determining to what extent writers' text quality, source use and argumentation behavior are related in L1 and L2, how effective writers' behavior is and whether their L2 proficiency influenced the relations between them. To answer these questions, twenty students wrote four short argumentative source based essays each in L1 (Dutch) and four in L2 (English). A within-writer cross-linguistic comparison of their texts revealed that their L1 and L2 writing competencies appear to be related. Furthermore, writers' source use behavior differed to some extent between languages, but the strong positive correlations found between source use features suggest that in most cases this was more a person than a language effect. Similarly, for argumentation behavior, results showed some learner specific features (e.g. inclusion of titles and reference lists), but differences between languages for others (e.g. the inclusion of both arguments and counter-arguments). Effects of the different source use and argumentation features studied on text quality were limited and no clear effect of L2 proficiency on writers' behavior or their influence on text quality were found.

---

✉ Daphne van Weijen  
D.vanWeijen@uva.nl  
Gert Rijlaarsdam  
G.C.W.Rijlaarsdam@uva.nl  
Huub van den Bergh  
H.vandenBergh@uu.nl

<sup>1</sup> Research Institute of Child Development and Education, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands

<sup>2</sup> University of Antwerp, Antwerp, Belgium

<sup>3</sup> Umeå University, Umeå, Sweden

<sup>4</sup> Utrecht University, Utrecht, The Netherlands

Overall, in line with earlier research, these findings provide some additional support for Cummins' LIH and the idea that writers might have a common underlying source for writing related knowledge and practices which they can apply in multiple languages.

**Keywords** Argumentative writing · First language · Second language · Source use behavior · Argumentation behavior

## Introduction

In higher education, students are often required to write argumentative essays as part of their academic coursework using relevant sources. During argumentative essay writing writers must adhere to strict requirements related to argumentative structure (Coirier, Andriessen, & Chanquoy, 1999) and distribute their attention effectively between what they want to say, how they want to say it, and how to apply their reasoning skills (Kellogg, 2008, p. 2; see also Albrechtsen, 2004; Ferretti & Fan, 2016; Van Wijk, 1999). In addition, selecting and integrating external information effectively in a text, through direct quotations or by paraphrasing, can be a cognitively demanding skill to master (Cumming, Lai, & Cho, 2016; Pecorari & Shaw, 2012; Strobl, 2015), because it requires both advanced reading and writing skills and effective interaction between them (Mateos & Solé, 2009). Thus argumentative writing is clearly “*an intellectually challenging problem*” (Ferretti & Fan, 2016) and as a result it is not surprising that young or inexperienced writers find writing argumentative texts more difficult than writing narrative ones (McCutchen, 2011). An added complicating factor, is that specific characteristics of genres such as argumentative writing, can differ between cultures and languages (cf. Graham & Rijlaarsdam, 2016), although this depends to some degree on the extent to which languages are related.

Inappropriate source use is a common problem in source based writing, which tends to occur more frequently in L2 than L1 writing (Keck, 2014). However, why students use information from sources inappropriately, i.e. without correct attribution to the original author, intentionally or unintentionally, is often unclear. Plagiarism occurs when students intentionally attempt to present the words of other authors as their own (Pecorari & Shaw, 2012). But what if students' behavior is unintentional? Then they may be “patchwriting” rather than plagiarizing. Patchwriting occurs when student writers incorporate chunks of verbatim source text in their essays to support their writing, without rephrasing the content in their own words (Howard, 1995; Pecorari, 2003). Some consider this to be lazy or dishonest behavior, which should be considered plagiarism and penalized accordingly. But Howard (1995) proposed that novice writers might also use patchwriting unintentionally to cope with their lack of L2 proficiency or source use skill. If this is true, then teachers should support students through this developmental phase and help them to become better writers.

For argumentative source-based essays, it is clear that, in addition to general writing skills and sufficient language proficiency (particularly in L2), students must acquire two types of domain-specific practices: argumentation or reasoning practices and source use practices (see Ferretti & Fan, 2016; see also De La Paz, Ferretti, Wissinger, Yee, & MacArthur, 2012). Potential ways to help students master these practices include permitting students to translate texts from L1 to L2 (e.g. Wolfersberger, 2003), use their L1 during L2 writing (Van Weijen, Van den Bergh, Rijlaarsdam, & Sanders, 2009), and use patchwriting while learning about appropriate source use (Pecorari, 2003). However, another option worth investigating is whether it is possible to teach students argumentation and source use practices in one language (L1 or L2) in a way which might enable them to apply their new behaviour in other languages or school subjects as well. If so, this could save both students and teachers a lot of time and effort. Whether this option has potential will be investigated in this study, by testing whether Cummins' Linguistic Interdependence Hypothesis applies to writing as well.

### Cummins' Linguistic Interdependence Hypothesis (LIH)

To be able to succeed in education and read and write at an advanced level, students must develop their Cognitive Academic Language Proficiency (CALP; Cummins, 1980). Cummins (1980) proposed that all language learners acquire Basic Interpersonal Communication Skills (BICS) informally, while developing CALP requires determined effort and formalized schooling. A person's BICS can be seen, according to Cummins, as the tip of the iceberg, while one's CALP is the less visible submerged part. Subsequently, Cummins proposed that in multilingual learners, the L1 and L2 CALP areas overlap between languages to form a '*Common Underlying Proficiency*' (CUP, Cummins, 1980). In later studies he named this idea the linguistic interdependence hypothesis (cf. Cummins, 2016), which "*assumes that two languages are distinct but are supported by shared concepts and knowledge derived from learning, experience, and the cognitive and language abilities of learners*" (Chuang, Joshi, & Dixon, 2012, p. 98). Nearly four decades after first proposing his hypothesis, Cummins concluded that "*The common underlying proficiency makes possible transfer of concepts, skills, and learning strategies across languages*" (Cummins, 2016, p. 940). However, most studies which investigated the LIH in relation to L1 and L2 literacy, focused on reading rather than on writing (see Dressler & Kamil, 2006, cited in Cummins, 2016; see also Chuang et al., 2012; Van Gelderen, Schoonen, Stoel, De Glopper, & Hulstijn, 2007), although the link between reading and writing (e.g. Fitzgerald & Shanahan, 2000) and the possibility of transfer between languages for writing have been studied to some extent as well. Cumming, Rebuffot, and Ledwell (1989), for example, investigated whether writing expertise might, in part, be a language-independent cognitive ability, and concluded that writers appear to use "*fundamentally similar*" thinking processes when performing a summarizing task in L1 and L2 (Cumming et al., 1989, pp. 213–214). Berman (1994) found signs of transfer from Icelandic (L1) to English (L2). Furthermore, Schoonen et al. (2003) found that L1 writing proficiency was highly correlated with L2 writing proficiency, while Van

Weijen et al. (2009) concluded that L1 writing proficiency had an indirect influence on L2 text quality. Finally, Cumming et al. (2016), concluded that it was very difficult to make clear distinctions between L1 and L2 writing, i.e. they are similar in many ways. These findings offer some support for Cummins' LIH hypothesis and thus for a degree of interdependence between L1 and L2 writing. If so, then it might be possible for writers to access and apply writing related (metacognitive) knowledge and practices learnt in one language in other languages as well.

It is important to mention, however, that earlier studies also suggested that language proficiency might influence the potential transfer of writing related practices from L1 to L2 or vice versa. Some stressed that writers can only apply their L1 knowledge in L2 if their L2 proficiency has risen above a certain threshold (Berman, 1994; Schoonen et al., 2003; Wolfersberger, 2003; see also Breuer, 2014, p. 70; Tillema, 2012, p. 79), while Cummins (1979) proposed that L1 knowledge can only be applied in L2 if it is sufficiently developed in the first language. Therefore, it is important to take writers' language proficiency levels into account as well.

### Source use and argumentation in academic writing

Earlier research on source based writing in L1 and L2 often focused on L2 writing only (e.g. Plakans & Gebriel, 2013; Weigle & Parker, 2012), or on comparing a group of L1 writers to a group of L2 writers (e.g. Keck, 2006, 2014; Shi, 2004), usually with single tasks per writer. However to determine whether source use and argumentation behavior are learner-specific practices, and thus transferrable between languages, requires a within-writer comparison of L1 and L2 writing across *multiple tasks*. In their recent synthesis on source based academic writing, Cumming et al. (2016) were surprised to find the number of studies based on within-writer comparisons of L1 and L2 source based writing was rather low, often included small student samples, and only one or two tasks at most (e.g. Kobayashi & Rinnert, 2008). They argued that within-writer cross-language comparisons are "... *needed to disentangle the differential effects of language proficiency and the ability to write from sources*" (Cumming et al., 2016, p. 53). This is necessary because research has shown that writing processes vary within writers, for example due to topic differences (cf. Van Weijen, 2009; Tillema, 2012), which means that to establish L1–L2 effects, multiple tasks must be collected per writer in each language.

Although earlier research has compared the way writers execute cognitive processes such as planning and formulating in L1 and L2 using multiple tasks per language (e.g. Van Weijen, 2009; Tillema, 2012), a comparison of source use and argumentation behavior in the same way has yet to be carried out. Such a comparison makes it possible to (a) compare variation in source use behavior and argumentation behavior not only between languages, but also *within writers* across multiple tasks, and (b) determine the effects of language proficiency and source use knowledge on source based argumentative writing in both languages. Weigle and Parker (2012), for example, found that "*only a small percentage of students borrowed extensively from source texts*" (p. 118), though "*a few students borrowed*

*substantially more than average, skewing these mean figures*” (p. 124). Similar results were found by Keck (2014). Perhaps, as these studies suggest, some participants show unique source use patterns, different from the rest of the group, because their patterns are to some extent learner-specific. The same might also hold for text structure and argumentation (see Sanders & Schilperoord, 2006).

## Aims and predictions

The aim of this study was to test whether Cummins’ Linguistic Interdependence Hypothesis (LIH) might also apply to writing, by answering the following questions:

1. To what extent is the quality of writers’ L1 and L2 texts related?
2. To what extent are writers’ source use and argumentation behavior, as reflected in their texts, related in L1 and L2?
3. To what extent are writers’ source use and argumentation behavior effective, i.e. related to text quality, in L1 and L2?
4. To what extent is the effectiveness of writers’ source use and argumentation behavior influenced or moderated by their L2 proficiency?

First, in line with Cummins’ linguistic interdependence hypothesis, and earlier results for reading (Chuang et al., 2012; Van Gelderen et al., 2007) and writing (Schoonen et al., 2003), we predict that we will find significant positive correlations between writers’ text quality scores in L1 and L2. Writers who are better writers in L1 are likely to be relatively better writers (compared to others in the group) in L2 as well.

Second, we predict that writers’ source use and argumentation behavior will be language independent, i.e. to a large extent learner specific (question 2). For source use, this means for example that some writers may have a clear preference for using larger quotes from the sources, while others may prefer to paraphrase source content in their texts in both languages. Regarding argumentation, we also expect to find similarities rather than differences between writers’ argumentation choices (opinion for, against or nuanced, alternating between arguments for or against) and text characteristics (e.g. providing a clear conclusion) between tasks and languages. This means we expect to find significant positive correlations between writers’ behavior for these aspects across languages, and/or a lack of significant differences between them. Thus, if we find no significant differences in writers’ source use and argumentation behavior between languages and tasks, that would suggest that there might be a common underlying source of writing related knowledge or practices available for writers to use in multiple languages.

Third, we predict that the effectiveness of writers’ source use and argumentation behavior (question 3) will be language independent. This means, for example, that if a writer’s source use behavior is similar in L1 and L2 then the quality of his or her texts is likely to be similar across languages as well (relative to that of others in the group). However, in line with earlier findings (e.g. Berman, 1994; Cummins, 1979; Schoonen et al., 2003) we predict that the effectiveness of writers’ source use and

argumentation behavior might be moderated by their language proficiency to some extent (question 4). In other words, if the effectiveness of writers' source use and argumentation behavior is found to differ between languages, this might be due to a lack of L2 proficiency. For example, less proficient L2 writers might be able to integrate sources correctly in their L1 texts, by paraphrasing source content in their own words, while in their L2 texts they may resort to patchwriting, i.e. copying chunks from the source texts without attribution, due to their lack of L2 proficiency. Thus, if we find significant differences in writers' behavior between languages, they might still have a common underlying source of knowledge and practices, but their potential to access to it might be moderated by their language proficiency.

## Method

### Data

Twenty students, all first year Bachelor English Majors ( $M = 18$  years old), each wrote four short argumentative essays in their L1 (Dutch) and four in their L2 (English). They all volunteered to take part and were given a small financial reward (50 Euros) for their efforts. The texts analyzed in this study were initially collected for another study (Van Weijen et al., 2009). The main focus of that study was an analysis of think-aloud data to determine the role of L1 use during L2 writing. However, students' source use and argumentation behavior have not yet been analyzed or reported on, which is why we considered it appropriate to analyze these texts in the current study.

Each assignment consisted of an issue, such as organ donation or camera surveillance, and six short sources. Participants had to provide their opinion on the issue, convince their readers of their point of view, support their opinion with arguments in a well-structured essay and include at least two of the sources provided in their text in a meaningful way (see Van Weijen et al., 2009 for an example of an assignment and an overview of the issues).

Each student wrote eight texts during four individual sessions, two texts per session, over a 6–8 week period, during their first semester at university. To avoid language and task order effects, the issues were evenly distributed across participants, languages, and sessions to ensure that each participant wrote only once on each issue, and that each task was completed as often in L1 and L2 and across sessions.

### Text quality

To answer the first research question, the quality of students' texts was assessed via two procedures, globally and analytically, by two different teams of five raters, one for each language. Texts were first assessed analytically, using a rating scheme which consisted of 18 criteria related to argumentation, content, text structure and conclusions. Raters were asked to rate the extent to which the text matched each of the criteria provided on a 5-point scale, ranging from not at all (1)

to meets the set criterion in every way (5). The scheme focused on text characteristics related to argumentation and structure, which might be similar across a writers' texts (cf. Sanders & Schilperoord, 2006).

Second, texts were assessed for global quality two weeks later, by comparing them to an average benchmark essay. For this procedure, raters were also provided with a brief description of the type of global text structure considered acceptable for this type of short essay (for more details on both assessment procedures, see Van Weijen, Van den Bergh, Rijlaarsdam, & Sanders, 2008; Van Weijen et al., 2009). The agreement between raters was satisfactory for the analytical (L1:  $\alpha = .88$ ; L2:  $\alpha = .93$ ) and global ratings (L1:  $\alpha = .82$ ; L2:  $\alpha = .83$ ). Furthermore, there were strong correlations between the two assessment methods in both L1 ( $r = .74$ ,  $r = .87$  when corrected for attenuation) and L2 ( $r = .69$ ;  $r = .79$  when corrected for attenuation). Therefore, we calculated a single quality score per essay in each language by adding up the global and analytical scores. Standardized scores ( $z$ -scores) per language were used to facilitate comparison across topics within each language. Overall, the average  $z$ -scores per topic across eight topics ranged from  $-.25$  to  $.30$  in L1 and  $-.32$  to  $.31$  in L2. Task order did not influence text quality (see Van Weijen et al., 2009). Subsequently, correlations between writers' L1 and L2 text quality scores were calculated as an indication of relatedness.

### Source use behavior

WCopFind, a plagiarism detection tool, was used to determine the extent to which students integrated information from the sources in their essays, and whether this varied between tasks and languages (Bloomfield, 2011; question 2). To find indications of copying from sources in the students' texts, the program filtered out all chunks of *four source words or more* copied verbatim from the sources. The chosen threshold was in line with earlier studies in which all meaningful chunks had at least four words (Shi, 2004; Weigle & Parker, 2012).

The program produced a comparison between each student's text and the corresponding assignment and sources. Two examples of such comparisons are presented in Fig. 1A, B. In both cases the student's text is shown on the left and the assignment text on the right. The chunks of at least four or more words which were found in both the student's text and the assignment are underlined in both panels. Figure 1A shows a text written by a student who used six short chunks in her text, while Fig. 1B shows a text written by another student on the same assignment who incorporated three longer extracts in her text.

Six variables were used to indicate the extent to which writers' source use varied between languages and tasks. First, we checked whether writers adhered to the source use requirement of at least 2 sources, by calculating the *number of unique sources* each writer included on average per text (see De La Paz et al., 2012). Second, we determined how heavily students relied on literal copying from the sources, by calculating two related measures: (a) the percentage of "*shared words*" (exact copies) from sources in each text (Keck, 2014), i.e. the percentage of the student's text which consists of source text words, taking text length into account; and (b) the *number of chunks* of four or more words writers copied from sources,



<p><b>A</b> Education in English</p> <p>Bilingual vwo is something new to us, it has been introduced a few years ago in The Netherlands. Not everybody is very enthusiastic about this method, but I think that education in English is a good idea, it helps youngsters to improve <b>their English skills. At the end of the course they will have reached 'near-native speaker' level and will be nearly as good as people for whom English is their mother tongue.</b> this is a big advantage of this method. The Kids can use these skills in further education, such as university where most books are in English and some <b>classes are even taught in English</b> by foreign professors. For example, <b>the Masters in Delft will be taught in English</b></p> <p>Another advantage of the bilingual method is the fact that <b>finding a job abroad</b> is easier, because of your language skills. Most young people have a goal of studying or working abroad, so the program opens doors for them. They also have an advantage to other foreign people who also want to find a job in a country were English is spoken.</p>	<p><b>Sources</b></p> <p>English is one of the most important languages of communication all over the world. Companies, universities and the cultural sector all ask for excellent skills in English. In the Netherlands more and more classes are being taught in English at university. Bilingual HAVO/VWO offers you the chance to lay the foundation for your <b>English skills. At the end of this course you will have reached 'near-native speaker' level. That means that you will be nearly as good as people for whom English is their mother tongue.</b> Source: <a href="http://www.meergonden.nl">www.meergonden.nl</a>, 2005.</p> <p>Experience has also taught us that students are less likely to enter into discussion during guest lectures in English than during classes which are taught in Dutch. Source: Louis Sicking, Hans Cools, e. a., <a href="http://www.leidennu.nl">www.leidennu.nl</a>, February 2003. Everyone can imagine that classes in English will be pre-prepared straightforward stories, and that interesting asides, quick replies and surprising, new examples will largely belong to the past. English will then not be an enrichment but will result in a deterioration of education, as many teachers of the University of Amsterdam have already indicated. Source: René Appel, Professor 'Acquisition and teaching of Dutch as a second language' University of Amsterdam, NRC-Handelsblad, June 18<sup>th</sup> 2003.</p> <p>The Oras-survey shows that students are moderately positive about English education. In February and March some five hundred students filled in the student council's survey on English-language education. A hot topic: from September most of <b>the Masters in Delft will be taught in English.</b> Most students appear 'moderately positive' towards English-language education. That's largely due to the fact that it increases the chance of <b>finding a job abroad.</b> Source: Joost Panhuysen, <a href="http://www.delta.tudelft.nl">www.delta.tudelft.nl</a>, April 25th 2002.</p> <p>Berlage Esprit is the first school in Amsterdam that offers a program in twwo (bilingual vwo). "In the first weeks of school it was hard to follow what was being said in English," Maria Janssen (13) says, "but if most of the class didn't understand something, the teacher said it again in Dutch. Now it's not hard anymore, we can understand things in English just as well as in Dutch." Source: Marc van Oostendorp, Onze Taal, <a href="http://www.vanoostendorp.nl">www.vanoostendorp.nl</a>, January 2000.</p> <p>Research was done in Delft on the effects of English-language education. It turned out that the number of students who failed a certain exam increased from 16 to 29% and the number of excellent results decreased from 25 to 13%. Source: Prof. dr. Cor van Bree, retired professor Dutch Linguistics at the University of Leiden, Leids Universiteits Blad, March 13th 2003.</p>
<p><b>B</b> Education in English shouldn't be questioned. English is at the moment the most important language in higher education, and students should be able to understand classes given in English. This is often not the case because English classes in secondary schools are often very poor. Dutch teachers of the English language often present their classes in Dutch rather than in English. This way they do not prepare their students enough for university. As said on the website <a href="http://www.meergonden.nl">www.meergonden.nl</a>: "<b>English is one of the most important languages of communication all over the world. Companies, universities and the cultural sector all ask for excellent skills in English. In the Netherlands more and more classes are being taught in English at university. Bilingual HAVO/VWO offers you the chance to lay the foundation for your English skills.</b>"</p> <p>Bilingual HAVO/VWO indeed lays a foundation for the English skills of the students, but not all secondary schools in the Netherlands are bilingual. This means that only students of bilingual secondary schools have that foundation. Therefore education in English should be encouraged at all the other secondary schools as well. A start would be to let Dutch teachers of the English language present their classes in English. This way the students can get used to the use of another, in this case the English, language, and it will prepare them for higher education and their career. <b>The Oras-survey shows that students are moderately positive about English education. In February and March some five hundred students filled in the student council's survey on English-language education. (...) Most students appear 'moderately positive' on English-language education. That's largely due to the fact that it increases the chance of finding a job abroad.</b> (source: Joost Panhuysen, <a href="http://www.delta.tudelft.nl">www.delta.tudelft.nl</a>, April 23th 2002)</p> <p>To conclude, English-language education in secondary school is highly important for students to be able to attend higher education.</p>	<p><b>Sources</b></p> <p>English is one of the most important languages of communication all over the world. Companies, universities and the cultural sector all ask for excellent skills in English. In the Netherlands more and more classes are being taught in English at university. Bilingual HAVO/VWO offers you the chance to lay the foundation for your <b>English skills.</b> At the end of this course you will have reached 'near-native speaker' level. That means that you will be nearly as good as people for whom English is their mother tongue. Source: <a href="http://www.meergonden.nl">www.meergonden.nl</a>, 2005.</p> <p>Experience has also taught us that students are less likely to enter into discussion during guest lectures in English than during classes which are taught in Dutch. Source: Louis Sicking, Hans Cools, e. a., <a href="http://www.leidennu.nl">www.leidennu.nl</a>, February 2003. Everyone can imagine that classes in English will be pre-prepared straightforward stories, and that interesting asides, quick replies and surprising, new examples will largely belong to the past. English will then not be an enrichment but will result in a deterioration of education, as many teachers of the University of Amsterdam have already indicated. Source: René Appel, Professor 'Acquisition and teaching of Dutch as a second language' University of Amsterdam, NRC-Handelsblad, June 18<sup>th</sup> 2003.</p> <p><b>The Oras-survey shows that students are moderately positive about English education. In February and March some five hundred students filled in the student council's survey on English-language education.</b> A hot topic: from September most of <b>the Masters in Delft will be taught in English.</b> Most students appear 'moderately positive' towards English-language education. <b>That's largely due to the fact that it increases the chance of finding a job abroad.</b> Source: Joost Panhuysen, <a href="http://www.delta.tudelft.nl">www.delta.tudelft.nl</a>, April 25th 2002.</p> <p>Berlage Esprit is the first school in Amsterdam that offers a program in twwo (bilingual vwo). "In the first weeks of school it was hard to follow what was being said in English," Maria Janssen (13) says, "but if most of the class didn't understand something, the teacher said it again in Dutch. Now it's not hard anymore, we can understand things in English just as well as in Dutch." Source: Marc van Oostendorp, Onze Taal, <a href="http://www.vanoostendorp.nl">www.vanoostendorp.nl</a>, January 2000.</p> <p>Research was done in Delft on the effects of English-language education. It turned out that the number of students who failed a certain exam increased from 16 to 29% and the number of excellent results decreased from 25 to 13%. Source: Prof. dr. Cor van Bree, retired professor Dutch Linguistics at the University of Leiden, Leids Universiteits Blad, March 13th 2003.</p>

**Fig. 1** A Example of a student's text (left) including six shorter chunks copied from the task's sources (right) marked in bold and underlined. B Example of a student's text (left) including several long chunks copied from the task's sources (right) marked in bold and underlined

identified by WCopyFind, a measure used in previous research as well (Shi, 2004; Weigle & Parker, 2012, see also Keck, 2014). Third, we determined to what extent writers integrated the copied words in their texts, operationalized as writers' *average chunk length* (number of words per chunk) per text, calculated by dividing the number of copied words by the number of chunks. The longer the average chunk



length, the higher the chance that writers used only a few longer chunks/direct quotes. Finally, we determined whether the chunks writers copied were mainly exact chunks or examples of patchwriting. Therefore, we calculated (a) the *number of exact copied chunks* per text, defined as the number of chunks containing full sentences or clauses (see Fig. 1A); and (b) the *number of patchwriting chunks* per text, defined as the number of chunks containing incomplete clauses or short phrases (see Fig. 1B). The exact copied and patchwriting chunks were coded manually from the WCopyFind data. For each of these six measures we calculated the mean per student for L1 and L2, as well as the standard deviation across four tasks within L1 and L2, to determine the variation between tasks in each language (question 2).

### Argumentation behavior

Students' argumentation behavior was compared across tasks and languages by examining a number of features of argumentative essays for each text, to determine whether they were stable within writers between languages (question 3). We chose to focus on two types of text features (1) four basic text features included in the analytical rating scheme, and (2) three features related to writers' use of sources containing arguments for or against the issue at hand.

First, we determined whether each essay contained four basic text features: a title, the writer's opinion, a clear conclusion, and a reference list, and checked whether writers consistently included these elements in their eight different texts. In addition, we checked whether the opinion and conclusion provided were clearly for the statement, against it, or contained a nuanced view (weighing both sides).

Second, we focused on how writers used the six different sources provided in each assignment to support their argumentation in their texts. There were three types of sources: sources containing only arguments for the issue, sources with only arguments against the issue, or mixed sources containing arguments which could be used to support both standpoints. Thus based on the sources a writer chose to incorporate in his or her text we could determine whether the direction of the sources used supported the writers' opinion provided in each text. We did this by checking whether (a) the direction of the first source a writer used in the text matched his opinion, (b) the direction of the last source used in the text matched the writer's opinion, and (c) how many times the direction of the sources used switched in the text. For example, one text might include only three sources containing arguments for the issue, which means there would be no switches (for–for–for), while another text might start with a source for the issue, followed by a source against the issue, and conclude with another source containing an argument for the issue, which would be counted as 2 switches (for–against–for). Thus the number of switches in source argument direction can be seen as an indication that writers used both arguments and counter-arguments in their texts, and thus as an indication of argumentation complexity. The higher the number of switches in source direction in a text, the higher the chance that the text contains arguments and counter-arguments, and thus a more complex argumentation structure.

Finally, we wanted to determine how writers incorporated the content of each source they used in their texts. As mentioned in the section on source use above, the

WCOPYFIND data were coded to determine how many exact copies or patchwriting chunks writers used per text, regardless of the source(s) they came from. However, the extent to which writers incorporated each source's content in their own words, using paraphrasing, was unclear. To check this, we manually coded for each source whether writers included its content in their texts using exact copies, patchwriting, paraphrasing or a mix of these three.

All source use and argumentation text features were coded by the first author, while a second trained coder coded all texts for one of the eight assignments in L1 and L2 as well ( $n = 20$ ). The intercoder reliability was sufficient for all these measures (Cohen's Kappa median = .67, mode = 1.00, range = .41–1.00), except for determining source use language for the first source writers included in the text for which Cohen's Kappa was lower (.38) but percentage agreement was sufficient (.60). The intraclass correlation between coders was also found to be satisfactory for the number of unique sources included in each text ( $icc = .77$ ). The first and second coders discussed the differences in their coding choices, after which the first coder coded the features of all the remaining texts in the same way.

## L2 proficiency

Students' L2 proficiency was measured using three different instruments, to determine whether their proficiency influenced the effectiveness of their source use and argumentation behavior (question 4). First, students were given a timed English vocabulary test, which consisted of 64 English sentences, ranked in order of difficulty, in which a missing word, provided in Dutch, had to be filled in (see Van Weijen et al., 2009 for more details). The test's internal consistency was satisfactory ( $\alpha = .81$ ). On average students answered 77% of the items correctly ( $M = 46.25$  points,  $SD = 6.77$ ).

The second instrument, was an L2 revision test, in which students were asked to improve the quality of a section of text in a flyer aimed at recruiting high school girls for team sports. The text contained 20 spelling and grammatical errors and points were awarded per error spotted and per successful correction. The task was based on a revision task developed by Hayes, Flower, Schriver, Stratman, and Carey (1985) and adapted for use in Dutch by Broekkamp and Van den Bergh (1996). On average students successfully spotted and corrected 72.13% of the errors in the text (mean score = 28.85,  $SD = 6.23$ ).

The third and final instrument used to measure students' L2 proficiency was their score on the Written English examination they had to take at the end of the semester. The exam consisted of 120 multiple choice questions on the following topics: Writing (Spelling, Style, Punctuation, Syntagmatics; 30 questions), Argumentation (30 questions), Grammar (30 questions) and Vocabulary (90 questions). On average, students answered two-thirds of the questions correctly (mean score = 6.77 on a 10 point scale,  $SD = 1.85$ ).

A reliability analysis confirmed that there were sufficient grounds for combining students' scores for these three instruments into a single score which represented their L2 proficiency ( $\alpha = .64$ ). Therefore the mean score across these three

instruments was calculated for each student and then converted to standardized scores for use in further analyses (range: 1.50–2.01).

## Analyses

Each writer in this study wrote eight tasks, four in L1 and four in L2, which means the tasks were nested within writers. Therefore, to answer research questions 2, 3 and 4, we applied multilevel analyses via SPSS mixed models where possible, with topics and subjects as random factors. To assess relations between writers' source use and argumentation behavior, we implemented a model with language (L1, L2) as fixed factor. For nominal variables we ran Chi square tests to determine whether they differed between languages or not. To answer the third and fourth research questions we set up models with text quality as dependent variable, two factors—language and the explanatory source or argumentation variable—and L2-proficiency as covariate, and included the two way interaction between language and the explanatory variable, and the three way interaction between language, L2-proficiency and the explanatory variable.

## Results

### Text quality and length

To answer our first question, we calculated the mean z-scores per topic in each language and correlations between writers' mean text quality scores and text length within and between languages. Results revealed a positive correlation between text quality in L1 and L2 ( $r = .73$ ). Second, the variation in text quality across four tasks ( $SD$  of mean text quality per person) was calculated as a measure of stability of text quality per language. A cross-linguistic comparison revealed a negative correlation between writers' average L1 text quality and the variation in their L2 text quality scores ( $r = -.64$ ). Third, an analysis of text length revealed a strong positive correlation between writers' average text length in both languages ( $r = .90$ ), while further analysis confirmed that students' texts did not differ in length, on average, between languages ( $F(1, 157) = .089, p = .77$ ; L1:  $M = 292.39$  words,  $SD = 79.03$ ; L2:  $M = 296.27$  words,  $SD = 84.54$ ). Finally, multilevel analysis revealed longer texts were generally rated with higher scores ( $F(1, 147.60) = 41.28, p < .001$ ), but there was no significant difference between languages ( $F(1, 137.10) = 1.88, p = .17$ ).

### Source use behavior in L1 and L2

The second question was to what extent writers' source use and argumentation behavior were related between languages. The WCopyFind output was analyzed to determine the variation within and across languages for six source use measures (see Table 1). First, we carried out multilevel analyses to determine whether the extent

**Table 1** Overview of mean occurrence of source use related text features in L1 and L2 and effects of language

Source use related text features (mean of 4 tasks)	Language		Difference between languages?
	L1 (Dutch) Mean (SE) (N = 20)	L2 (English) Mean (SE) (N = 20)	
Number of unique sources included per text	2.38 (.117)	2.15 (.117)	$F(1, 132.381) = 4.311, p = .04$
Percentage words in text copied from sources	.17 (.02)	.20 (.02)	$F(1, 131.885) = 5.767, p = .018$
Number of chunks copied from sources	5.24 (.53)	6.44 (.53)	$F(1, 131.714) = 9.277, p = .003$
Average chunk length per text	10.33 (1.27)	8.93 (1.27)	$F(1, 131.951) = 2.933, p = .089$
Number of exact copied chunks per text	.90 (.21)	1.08 (.21)	$F(1, 131.737) = 1.332, p = .251$
Number of patchwriting chunks per text	3.51 (.45)	4.30 (.45)	$F(1, 131.657) = 4.495, p = .036$

to which writers used these features in their texts differed in L1 and L2 (see Table 1).

Results revealed that the number of sources included, the percentage of words and chunks copied from sources, and the number of patchwriting chunks per text, differed between languages, while writers' average chunk length, and the number of exact copied chunks included per text did not (see Table 1). On average, writers included slightly more sources in L1 than in L2, although they included at least two source extracts per text in both languages (range: 0–5), in line with the minimum required by the assignment. Furthermore, writers included a higher percentage of copied words, more copied chunks, and more patchwriting chunks on average in their texts in L2 than in L1.

Second, to gain insight in the relations between features within and between languages and the extent to which these relations were stable across languages, correlations for these measures within and between languages were generated to check for relations between them (see Appendix Table A1 for an overview). Within languages, there appear to be stronger relations between features in L2 than in L1. In L1, each source use feature correlated positively with at least two other features ( $r$  range = .59–.81). The only negative correlation in L1 was found between the number of unique sources writers used and mean chunk length ( $r = -.53$ ); texts with fewer unique sources tend to contain longer chunks. We also found positive correlations in L1 between the number of unique sources per text, the number of chunks, and the number of patchwriting chunks per text; texts with more unique sources are likely to include more patchwriting chunks. In L2, each feature was positively correlated with at least four others ( $r$  range = .51–.86), except for the number of unique sources used, which correlated only with mean number of chunks and mean number of patchwriting chunks, as was the case in L1.

Finally, a cross-language comparison revealed that all of the features were related to their counterparts across languages ( $r$  range = .48–.85), except for the number of unique sources used and patchwriting chunks, for which no significant L1–L2 correlations were found.

## Argumentation behavior in L1 and L2

Results of the argumentation behavior analysis revealed that writers' use of general text features such as titles and reference lists did not differ between languages (see Table 2). On average, writers included titles in the majority of their texts, but were much less likely to provide reference lists. However, a significant difference between languages was found for the inclusion of clear conclusions in texts, which appeared more often on average in L2 texts than in those written in L1.

Next, we checked whether each text contained the writer's opinion, and whether the opinion and conclusion provided were clearly for the statement, against it, or provided a nuanced view (weighing both sides). An overview of choices for text features related to writers' opinions in L1 and L2 is provided in Table 3.

Results of the comparison between languages for these features revealed no significant differences in writers' chosen opinion, or their choices for source direction between languages, i.e. only for, only against or a mix of both. Furthermore, we investigated whether the direction of the sources used supported the writers' opinion provided in each text, and found that writers' L1 texts were more likely to contain switches in source direction than their L2 texts (see Table 4).

**Table 2** Overview of general text features in L1 and L2 (Proportions) and effects of language

Features	Language		Effect of language used?
	L1 (Dutch) Mean (SE)	L2 (English) Mean (SE)	
Texts with title	.86 (.06)	.87 (.06)	$F(1, 131.414) = .041, p = .841$
Texts with conclusion	.85 (.04)	.94 (.04)	$F(1, 131.497) = 4.147, p = .044$
Texts with reference list	.10 (.06)	.10 (.06)	$F(1, .001) = .001, p = 1.000$

**Table 3** Overview of argumentation related text features in L1 and L2 (percentages across four tasks in each language)

Feature	Opinion					
	Language	Unclear (%)	Against (%)	Nuanced/ mixed (%)	For (%)	Difference? Chi square ( $\chi^2$ )
Writer's opinion	L1	7.5	31.3	22.5	38.8	$\chi^2 = 5.82, df = 3, p = .12$
	L2	3.8	25.3	13.9	57.0	
Direction of sources in text	L1	0	26.6	45.6	27.8	$\chi^2 = 2.68, df = 2, p = .26$
	L2	0	22.1	37.7	40.3	

**Table 4** Overview of argumentation related text features in L1 and L2 (percentages across four tasks in each language)

Features	Language		
	L1 (Dutch) Mean (SE)	L2 (English) Mean (SE)	Effect of language used?
First source used matches opinion	.19 (.06)	.15 (.05)	$F(1, 93.43) = 0.311, p = 0.578$
Last source used matches opinion	.38 (.07)	.27 (.07)	$F(1, 123.40) = 2.331, p = 0.129$
Number of switches in source direction	.65 (0.12)	.39 (0.12)	$F(1, 120.46) = 6.292, p = 0.013$

**Table 5** Overview of chosen method for incorporating source content in texts (percentages) in each language

Language	Chosen method for incorporating source content in text			
	Mainly exact copy (%)	Mainly patchwriting (%)	Mainly paraphrased (%)	Mixed (%)
L1	22.8	7.6	20.3	49.4
L2	29.9	15.6	22.1	32.5
Overall	26.3	11.5	21.2	41

Finally, we checked how writers chose to incorporate the content of each source in their texts: using exact copies, patchwriting, paraphrasing or a mix of these three, and found no significant differences between languages ( $\chi^2 = 5.68, df = 3, p = .13$ ; see Table 5).

## L2 proficiency, source use behavior and text quality

In the last step of the analysis we investigated to what extent writers' source use and argumentation behavior was related to text quality (question 3) and to what extent these relations were moderated by L2 proficiency (question 4). Multilevel analyses were carried out to determine the influence of source use text features and language (L1 or L2) on text quality. Results of the analysis are summarized in Table 6 below.

**Table 6** Overview of relations between source use variables, text quality, language used and L2 proficiency (significant values in bold)

Variable	Effect on text quality	Interaction with language?	Influence of L2 proficiency on interaction?
Number of unique sources included per text	$p = .310$	$p = .816$	$p = .071$
Percentage of words copied per text	$p = .991$	$p = .404$	$p = .129$
Average number of chunks per text	$p = .399$	$p = .896$	$p = .452$
Average chunk length	$p = .017$	$p = .506$	$p = .977$
Average number of exact chunks	$p = .004$	$p = .889$	$p = .481$
Average number of patchwriting chunks	$p = .447$	$p = .426$	$p = .674$

Results of the analysis revealed that average chunk length is positively related to text quality ( $F(1, 153.331) = 5.596, p = .017$ ), as is the number of exact copied chunks ( $F(1, 152.515) = 8.629, p = .004$ ), although these relations did not differ significantly between languages. Thus, writers who used longer and more exact chunks were likely to write texts of better quality, regardless of the language written in. No significant differences between languages were found for the other features we studied (see Table 6). Subsequently, multilevel analyses revealed that writers' L2 proficiency level did not influence or moderate the relations between writers' source use behavior and text quality for any of the variables studied, as no significant interactions were found (see Table 6).

## L2 proficiency, argumentation behavior and text quality

In line with the source use analysis we used multilevel analyses to determine whether writers' use of these argumentative text features was related to the quality of students' texts across languages (see Table 7).

Positive effects were found for providing a title ( $F(1, 154.357) = 6.125, p = .014$ ), a clear conclusion ( $F(1, 143.133) = 11.301, p < .001$ ) and providing a reference list ( $F(1, 136.721) = 4.784.931, p < .001$ ) on text quality, but this effect only differed between languages for the last feature. Providing a reference list was positively related to text quality in L1 ( $r = .34, p = .002$ ), but not in L2 ( $r = .02, p = .84$ ). No significant relations were found between the direction of sources used and text quality and writers' choices for integrating sources in their texts and text quality between languages.

Finally we checked whether writers' L2 proficiency level influenced these relations in any way. Unfortunately, besides an effect of L2 proficiency on the relations between writers' method for source content integration in their texts ( $F(1, 133.647) = 6.319, p = .013$ ), no other influence of proficiency on writers' behavior was found.

**Table 7** Overview of relations between argumentation variables, text quality, language used and L2 proficiency (significant values in bold)

Variable	Effect on text quality	Interaction with language?	Influence of L2 proficiency?
Providing a title	$p = .014$	$p = .170$	$p = .232$
Providing a clear conclusion	$p < .001$	$p = .789$	$p = .286$
Providing a reference list	$p = .863$	$p < .030$	$p = .185$
Direction of sources (mainly for, mainly against, or mixed)	$p = .706$	$p = .918$	$p = .485$
First source used matches opinion	$p = .784$	$p = .248$	$p = .951$
Last source used matches opinion	$p = .615$	$p = .322$	$p = .418$
Number of switches in direction of sources	$p = .166$	$p = .782$	$p = .632$
Method for incorporating source content (exact, patch, paraphrase or mixed)	$p = .803$	$p = .495$	$p = .013$



## Conclusion and discussion

The aim of this study was to test whether Cummins' Linguistic Interdependence Hypothesis (LIH) might also apply to writing, by examining the extent to which writers' text quality, source use and argumentation behavior differed between languages and whether their L2 proficiency influenced the relations between them.

First, in line with Cummins' LIH, and earlier results for reading (Chuang et al., 2012; Van Gelderen et al., 2007) and writing (Schoonen et al., 2003), we found positive correlations between text quality and text length between languages as well. As these findings are based on four tasks per writer per language, this enables us to draw conclusions about their writing expertise in both languages. Thus better L1 writers seem to be better writers in L2 as well, and those who tend to write longer texts in their L1 are likely to write longer texts in their L2. However, there are also indications that writers' L1 writing competency influences their L2 writing expertise to some extent: writers' L2 text quality appears to be less stable between tasks for weaker L1 writers than for better ones.

Second, we predicted that writers' source use and argumentation behavior would be language independent, i.e. to a large extent learner specific. Multilevel analyses revealed that writers source use behavior differed significantly between languages for four of the six variables studied. On average, they included slightly more unique sources, a higher percentage of copied words, more copied chunks, and more patchwriting chunks in their texts in L2 than in L1, which suggests that they relied more strongly on source content in their L2 writing than in L1.

Further analyses revealed all of the source use features were strongly related to their counterparts across languages, except for the number of unique sources used and patchwriting chunks, for which no significant L1–L2 correlations were found. Writers who tend to copy more words from sources, in longer and more exact chunks in L1, are also likely to do so in L2, which is in line with findings from earlier between-writer comparisons (Keck, 2014; Weigle & Parker, 2012). Furthermore, because writers' use of patchwriting chunks was not correlated between languages, this suggests it might be influenced by L2 proficiency to some extent. Thus, although we found a language effect for four out of six variables, the strong positive correlations found between source use features across languages suggest that in most cases this was more a person than a task effect.

A cross-language comparison of writers' argumentation behavior revealed that whether writers chose to include titles or reference lists in their texts was similar between tasks across languages and thus largely learner specific. For example, only two of the writers in our study included a rudimentary reference list in every single text they wrote, while the others did not. But writers tended to include clearer conclusions in their texts on average in L2 than in L1, which suggests they might be inclined to draw more straightforward or less nuanced conclusions in L2. Moreover, we found that writers were more likely to incorporate both sources for and against the issue at hand and switch between the two points of view in their texts in L1 than in L2. This suggests that when writing in L2, writers were more likely to revert to a more straightforward canonical argumentation structure, in which they only provide

arguments from sources which clearly support their chosen opinion, rather than including and refuting counter-arguments as well (cf. Sanders & Schilperoord, 2006). Overall, these results provide a first indication that writers' source use and argumentation behavior might be quite similar between languages and that, as predicted by Cummins' LIH, there might be a common underlying source of writing related knowledge or practices available for writers to use in multiple languages (Cummins, 1980, 2016).

Third, we predicted that the effectiveness of writers' source use and argumentation behavior (question 3) would be language independent as well. Results of the analysis revealed that only writers' average chunk length and the number of exact copied chunks used were positively related to text quality and these relations did not differ significantly between languages. This suggests that writers' chunk length and use of exact chunks were effective in both L1 and L2. This might be due in part to the positive correlation between text length and text quality, as including longer exact chunks often results in longer texts. No direct relations or differences between languages were found between the other source use features and text quality. We were somewhat surprised to find no effect of the number of unique sources use on text quality, as earlier research by De La Paz et al. (2012) found that better writers tended to use more sources in their texts, although this might be due to the fact that students in their study wrote argumentative historical essays for which source use can be far more complex.

Similarly, for argumentation behavior, we only found positive effects on text quality for the inclusion of titles and conclusions, which is in line with De La Paz et al. (2012) who also found a small positive effect for the inclusion of conclusions. Texts with titles and clear conclusions were judged to be higher in quality than those lacking these features, regardless of the language written in, while inclusion of a reference list only had a positive effect on text quality in L1, not in L2. Surprisingly, the inclusion of sources including both arguments and counter-arguments did not appear to influence text quality directly, which might be due to the influence of L2 proficiency on these relations, as proposed in earlier studies (e.g. Berman, 1994; Cummins, 1979; Schoonen et al., 2003).

Therefore, in the final step of the analysis we checked whether writers' L2 proficiency level influenced their source use and argumentation behavior and their effectiveness in any way. For source use, we found that L2 proficiency did not influence or moderate any of the relations between source use behavior and text quality studied. However, there was a trend visible for the number of unique sources included, which suggests further research on the influence of L2 proficiency on source use behavior might be warranted. Finally, results for argumentation behavior revealed no clear effect of writers' L2 proficiency on any of the source use or argumentation behavior variables studied, except on the method for integrating source content.

When taken together, all these findings seem to provide additional support for Cummins' LIH hypothesis. Although they are somewhat tentative, our findings, based on a within-writer study in L1 and L2 using multiple tasks, appear to confirm those of earlier studies (Berman, 1994; Cumming et al., 2016; Cummins, 1979; Keck, 2014; Kobayashi & Rinnert, 2008; Schoonen et al., 2003; Weigle & Parker, 2012) and thus provide some additional support for the idea that writers might have a common underlying source of writing related knowledge and practices, such as source use and

argumentation behavior, which they can apply in multiple languages. However, the role of L2 proficiency in this process remains unclear and requires further investigation.

### Limitations and suggestions for further research

A clear strength of the current study was the fact that we carried out a within-writer cross-language comparison, as proposed by Cumming et al. (2016) with eight texts per writer, four in each language. However, as in any study, the present study also has a number of limitations. First of all, the absence of an L1 proficiency measure, besides L1 text quality, made it impossible to investigate the influence of L1 proficiency on differences between writers' behavior across languages to a greater extent.

Second, while the sample size in our study was reasonable and the within-writer comparison we carried out was novel, the fact that we used eight very similar tasks was perhaps less well chosen: our findings are bound to the specific type of argumentative task we used. If we had chosen to use two different types of argumentative tasks, or tasks with two different contexts, but had still found similarities in writers' behavior between tasks, that would have provided stronger support for the validity of Cummins' LIH.

Third, the fact that we chose to compare two closely related languages, Dutch and English, which share the same orthography and similar basic argumentative structures, is also a possible limitation of our study. If we had chosen to compare two very different languages, such as English and Chinese, then it is likely our results would have been somewhat different, although we would be surprised to find huge differences, because earlier studies for reading (Chuang et al., 2012) and writing (Kobayashi & Rinnert, 2008) found strong relations between languages with different orthographies as well.

Fourth, writers' argumentation behavior is multifaceted and we were not able to cover all its aspects in this study. Therefore, in a future study, we would like to investigate writers' opinions, the number and type of arguments they integrated in their texts and whether these differed between languages in more detail. Finally, although our conclusions are somewhat tentative, we do still think that it might be possible for students to learn writing related practices in one language in such a way that they can apply them in a second language as well. Therefore, future research might focus on intervention studies comparing students' writing behavior in bilingual versus monolingual schools, which could generate pedagogical implications to help students improve the effectiveness of their source based writing in multiple languages at the same time.

**Acknowledgements** We wish to thank Mujgan Buyuktas Kara for the suggestion to use WCopyFind to analyze students' source use behavior (Buyuktas Kara, Rijlaarsdam, & Van Steendam 2015). Furthermore we also wish to thank Martijn Koek for contributing to the data coding process.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### Appendix A

**Table A1** Correlations between source use related text features in L1 and L2 (mean values across four tasks in each language)

Language	L1	L2										
		Text features	1	2	3	4	5	6				
L1	1. Mean # of unique sources	1										
	2. Mean % words copied	.03	1									
	3. Mean # of chunks	.64**	.59**	1								
	4. Mean chunk length	-.53*	.63**	-.08	1							
	5. Mean # of exact chunks	-.32	.62**	.11	.67**	1						
	6. Mean # of patch chunks	.70**	.23	.81**	-.37	-.42	1					
L2	1. Mean # of unique sources	.26	-.12	.01	-.12	-.07	0	1				
	2. Mean % words copied	-.07	.71**	.37	.54*	.57**	.08	.29	1			
	3. Mean # of chunks	.25	.37	.48*	.13	.30	.27	.51*	.77**	1		
	4. Mean chunk length	-.35	.66**	.13	.85**	.71**	-.16	-.08	.73**	.34	1	
	5. Mean # of exact chunks	-.12	.70**	.41	.56*	.52*	.19	-.07	.74**	.52*	.67**	1
	6. Mean # of patch chunks	.28	.02	.27	-.17	-.05	.28	.59**	.52*	.86**	.05	.15

\* $p < 0.05$ ; \*\* $p < 0.01$

## References

- Albrechtsen, D. (2004). Writing in two languages—comparing processing in relation to two writing tasks in a “within subjects” design: attention to argumentation in on-line processing. In L. Allal & J. Dolz (Eds.), *Proceedings writing 2004 [CD]*. Geneva: Adcom Productions.
- Berman, R. (1994). Learners’ transfer of writing skills between languages. *TESL Canada Journal/Revue TESL du Canada*, 12(1), 29–46. <https://doi.org/10.18806/tesl.v12i1.642>.
- Bloomfield, L. (2011). WCopyFind. Plagiarism detection tool (Version 4.1.5) [Software]. Available from <http://plagiarism.bloomfieldmedia.com/wordpress/software/wcopyfind/>.
- Breuer, E. O. (2014). *First language versus foreign language: Fluency, errors and revision processes in first language and foreign language writing*. Frankfurt/Main: Peter Lang.
- Broekkamp, H., & Van den Bergh, H. (1996). Attention strategies in revising a foreign language text. In G. Rijlaarsdam & E. Espéret (Series Eds.), G. Rijlaarsdam, H. Van den Bergh & M. Couzijn (Vol Eds.), *Studies in writing Vol. 2, theories, models and methodology in writing research* (pp. 170–181). Amsterdam: Amsterdam University Press.
- Buyuktas Kara, M., Rijlaarsdam, G. C. W., & Van Steendam, E. (2015). Effects of video-based peer observation on summarizing from multiple sources. In *Paper presented at the 16th biennial EARLI conference*, Cyprus, August 25–29, 2015.
- Chuang, H. K., Joshi, R. M., & Dixon, L. Q. (2012). Cross-language transfer of reading ability: Evidence from Taiwanese ninth-grade adolescents. *Journal of Literacy Research*, 44(1), 97–119. <https://doi.org/10.1177/1086296X111431157>.
- Coirier, P., Andriessen, J., & Chanquoy, L. (1999) From planning to translating: the specificity of argumentative writing. In G. Rijlaarsdam & E. Espéret (Series Eds.) & J. Andriessen & P. Coirier (Vol. Eds.), *Studies in writing: Vol. 5. Foundations of argumentative text processing* (pp. 1–28). Amsterdam: Amsterdam University Press.
- Cumming, A., Lai, C., & Cho, H. (2016). Students’ writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes*, 23(2016), 47–58. <https://doi.org/10.1016/j.jeap.2016.06.002>.
- Cumming, A., Rebuffot, J., & Ledwell, M. (1989). Reading and summarizing challenging texts in first and second languages. *Reading and Writing: An Interdisciplinary Journal*, 2, 201–219. <https://doi.org/10.1007/BF00377643>.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222–251. <https://doi.org/10.3102/00346543049002222>.
- Cummins, J. (1980). The construct of language proficiency in bilingual education. In J. E. Atlantis (Ed.), *Current issues in bilingual education. Georgetown University round table on languages and linguistics (GURT) 1980* (pp. 81–103). Washington, DC: Georgetown University Press.
- Cummins, J. (2016). Reflections on Cummins (1980), “The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue”. *TESOL Quarterly*, 50(4), 940–944. <https://doi.org/10.1002/tesq.339>.
- De La Paz, S., Ferretti, R. P., Wissinger, D., Yee, L., & MacArthur, C. A. (2012). Adolescents’ disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication*, 29(4), 412–454. <https://doi.org/10.1177/0741088312461591>.
- Dressler, C., & Kamil, M. (2006). First- and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners*. Report of the National Literacy Panel on Language-Minority Children and Youth (pp. 197–238). Mahwah, NJ: Lawrence Erlbaum.
- Ferretti, R. P., & Fan, Y. (2016). Argumentative writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 301–315). New York, NY: Guilford.
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35(1), 39–50. [https://doi.org/10.1207/S15326985EP3501\\_5](https://doi.org/10.1207/S15326985EP3501_5).
- Graham, S., & Rijlaarsdam, G. (2016). Writing education around the globe: Introduction and call for a new global analysis. *Reading and Writing*, 29(5), 781–792. <https://doi.org/10.1007/s11145-016-9640-1>.
- Hayes, J. R., Flower, L., Schriver, L., Stratman, J., & Carey, L. (1985). *Cognitive processes in revision* (technical report). ERIC document, ED 267 396.
- Howard, R. M. (1995). Plagiarisms, authorships, and the academic death penalty. *College English*, 57, 788–806. <https://doi.org/10.2307/378403>.

- Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, 15, 261–278. <https://doi.org/10.1016/j.jslw.2006.09.006>.
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, 25, 4–22. <https://doi.org/10.1016/j.jslw.2014.05.005>.
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1, 1–26. <https://doi.org/10.17239/jowr-2008.01.01.1>.
- Kobayashi, H., & Rinnert, C. (2008). Task response and text construction across L1 and L2 writing. *Journal of Second Language Writing*, 17, 7–29. <https://doi.org/10.1016/j.jslw.2007.08.004>.
- Mateos, M., & Solé, I. (2009). Synthesizing information from various texts: A study of procedures and products at different educational levels. *European Journal of Psychology of Education*, 24, 435–451. <https://doi.org/10.1007/BF03178760>.
- McCutchen, D. (2011). From novice to expert: Implications of language skills and writing-relevant knowledge for memory during the development of writing skill. *Journal of Writing Research*, 3, 51–68. <https://doi.org/10.17239/jowr-2011.03.01.3>.
- Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12, 317–345. <https://doi.org/10.1016/j.jslw.2003.08.004>.
- Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing*, 21, 149–164. <https://doi.org/10.1016/j.jslw.2012.03.006>.
- Plakans, L., & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, 22, 217–230. <https://doi.org/10.1016/j.jslw.2013.02.003>.
- Sanders, T., & Schilperoord, J. (2006). Text structure as a window on the cognition of writing: How text analysis provides insights in writing products and writing processes. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *The handbook of writing research* (pp. 386–402). New York, NY: Guilford Publications.
- Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., et al. (2003). First language and second language writing: the role of linguistic knowledge, speed of processing and metacognitive knowledge. *Language Learning*, 53(1), 165–202. <https://doi.org/10.1111/1467-9922.00213>.
- Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication*, 2, 171–200. <https://doi.org/10.1177/0741088303262846>.
- Strobl, C. (2015). Learning to think and write together: Collaborative synthesis writing, supported by a script and a video-based model. In G. Rijlaarsdam (Series Ed.) & M. Deane & T. Guasch (Vol. Eds.), *Studies in writing: Vol. 29, Learning and teaching writing online*, (pp. 67–93). Leiden: Brill.
- Tillema, M. (2012). *Writing in first and second language: Empirical studies on text quality and writing processes*. Utrecht: LOT.
- Van Gelderen, A., Schoonen, R., Stoel, R. D., De Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, 99(3), 477–491. <https://doi.org/10.1037/0022-0663.99.3.477>.
- Van Weijen, D. (2009). *Writing processes, text quality, and task effects: Empirical studies in first and second language writing*. Utrecht: LOT Dissertation Series.
- Van Weijen, D., Van den Bergh, B., Rijlaarsdam, G., & Sanders, T. (2008). Differences in process and process-product relations in L2 writing. *ITL Applied Linguistics*, 156, 203–226. <https://doi.org/10.2143/ITL.156.0.2034433>.
- Van Weijen, D., Van den Bergh, H., Rijlaarsdam, G. C. W., & Sanders, T. J. M. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing*, 18(4), 235–250. <https://doi.org/10.1016/j.jslw.2009.06.003>.
- Van Wijk, C. (1999). Conceptual processes in argumentation: A developmental perspective. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 31–50). Amsterdam: Amsterdam University Press.
- Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing*, 21, 118–133. <https://doi.org/10.1016/j.jslw.2012.03.004>.
- Woltersberger, M. (2003). L1 to L2 writing process and strategy transfer: A look at lower proficiency writers. *TESL-EJ: Teaching English as a Second or Foreign Language*, 7(2), A6 1–15.