



The effects of balanced biliteracy on Greek-German bilingual children's secondary discourse ability

Christiane Bongartz & Jacopo Torregrossa

To cite this article: Christiane Bongartz & Jacopo Torregrossa (2017): The effects of balanced biliteracy on Greek-German bilingual children's secondary discourse ability, *International Journal of Bilingual Education and Bilingualism*, DOI: [10.1080/13670050.2017.1355888](https://doi.org/10.1080/13670050.2017.1355888)

To link to this article: <http://dx.doi.org/10.1080/13670050.2017.1355888>



Published online: 16 Aug 2017.



Submit your article to this journal [↗](#)



Article views: 42



View related articles [↗](#)



View Crossmark data [↗](#)



The effects of balanced biliteracy on Greek-German bilingual children's secondary discourse ability

Christiane Bongartz^a and Jacopo Torregrossa^b

^aEnglish Department, University of Cologne, Cologne, Germany; ^bDepartment of Romance Languages, University of Hamburg, Hamburg, Germany

ABSTRACT

Several studies have revealed the cognitive and linguistic benefits of balanced bilingualism, but the research on balanced biliteracy is still in its initial stages (Marinis et al. submitted. "Biliteracy Education Impacts on Cognition Selectively."). This study investigates the positive effects of balanced biliteracy on the development of secondary discourse ability involved in narrative production. The analysis is based on narratives told by 30 Greek-German bilingual children (9–12 years) in both their languages. The children are divided in three groups, differing in language proficiency in each respective language and features of the educational setting (in terms of balance of teaching hours dedicated to each language). The results reveal that the features of the educational setting are the most reliable predictor of children's narrative skills. Moreover, balanced biliteracy has the effect of evening out imbalances in language proficiency. The study contributes to identifying the factors that affect secondary discourse ability and motivate variation in the development of literacy skills (Francis 2006. "The Development of Secondary Discourse Ability and Metalinguistic Awareness in Second Language Learners." *International Journal of Applied Linguistics* 16 (1): 37–60; Bongartz 2016. "Bilingual and Second Language Development and Literacy – Emerging Perspectives on an Intimate Relationship." Proceedings of the 21st International Symposium of Theoretical and Applied Linguistics 2013, Thessaloniki, April 5–7, 2013).

ARTICLE HISTORY

Received 13 May 2017
Accepted 18 June 2017

KEYWORDS

Narrative; secondary discourse; balanced biliteracy; bilingual proficiency; story grammar; referential strategies

1. Introduction

The ability to produce a coherent narrative discourse is considered as one of the best predictors of children's literacy skills and reading abilities (McCabe and Rollins 1994; Bliss, McCabe, and Miranda 1998; Walach 2008; Bongartz 2016; Gagarina 2016) as well as one of the main components of academic language proficiency (Cummins 1980; Shrubshall 1997; Miller et al. 2006). In particular, one has to distinguish between storytelling occurring in everyday conversation and 'school narratives' (Fang 2001) which are more closely associated with literacy practices. Everyday storytelling is an example of context-embedded language use and may result in an unstructured sequence of reported events. In fact, the connection between the different story constituents is not necessarily signaled by cohesive linguistic devices (such as sentential connectors and anaphoric relations), given that it can rely on shared background knowledge between the speaker and the interlocutor, as well as on paralinguistic cues, such as gestures and intonation (Nicolopoulou, McDowell, and Brockmeyer 2006). Adopting the terminology introduced in Francis (2006), this contextualized, everyday type of narrative discourse is supported by speakers' 'primary discourse ability', which

is achieved by everyone and is thus universal. On the contrary, the production of schooled narratives involves a more sophisticated, explicit knowledge of the requirements of narrative discourse (i.e. it involves 'translating knowing into telling', Francis 2006, 41). For instance, speakers know that they have to organize information in a narrative schema consisting of different components (e.g. setting, complicating actions and resolutions) and generate a coherent discourse by relying on lexical and grammatical means to indicate cohesive ties. Moreover, they have to build an explicit model of the interlocutor's mental state (Arnold, Brown-Schmidt, and Trueswell 2007), monitoring her attention and knowledge to avoid ambiguities or inconsistencies. This knowledge forms part of the speakers' 'secondary discourse ability' (Francis 2006; Bongartz 2016), which refers to a cluster of decontextualized linguistic practices involving 'explicitly construing, conveying and comprehending information' (Nicolopoulou, McDowell, and Brockmeyer 2006), as well as a 'retrospective and non-linear kind of processing', which is typically associated with discourse planning (Francis 2006, 46). The transition from primary to secondary discourse ability implies learning metalinguistic strategies and operations, which is the outcome of formal instruction and education. Therefore, the development of secondary discourse is more prone to variation than primary discourse, depending on the speakers' cognitive (e.g. processing speed and executive functions) and sociological profile (including features of the attended school system; cf. Gough and Bock 2001).

The analysis of bilingual narrative discourse affords a privileged viewpoint to identify possible sources of variation in secondary discourse ability. First, one can observe how secondary discourse is affected by proficiency in each of the two languages. Bilingual children with an imbalanced proficiency may lack the linguistic means necessary to translate 'knowing into telling' in their weaker language. Evidence in this direction comes from studies conducted on Spanish-English bilingual children (Gutiérrez-Clellen 2002; Montanari 2004; see the review in Section 2). Second, the study of bilingualism allow us to investigate to what extent variation in secondary discourse reflects differences in biliteracy.

Several studies have pointed out that bilingualism has a positive effect on cognitive and linguistic processing (cf. Bialystok 2001 for a review). For instance, it involves enhanced metalinguistic awareness in several domains (Cummins 1980; Ricciardelli 1992; Bialystok 1993, 2001 and Sanz 2012 for a review), including word awareness (cf. Cummins 1980), syntactic awareness (cf. Galambos and Goldin-Meadow 1990) and phonological awareness (cf. Bialystok, Majumder, and Martin 2003). However, these advantageous effects are more visible among fully functional, balanced bilinguals (Bialystok 2001). Extending this pattern to education, we expect that the benefits of biliteracy are enhanced in association with a bilingual educational program which is balanced in terms of hours of instruction in each language. Preliminary evidence in favor of this hypothesis is reported in Marinis, Tsimpli, and Bongartz (submitted). The authors analyze the cognitive effects of biliteracy in a group of bilingual children (8–12 years) of different language combinations and show that the more balanced the degree of biliteracy, the higher the scores in the executive function tasks.

In this paper, we will complement the study of Marinis et al. by identifying the impact of biliteracy on language, focusing on the development of secondary discourse ability. Based on the previous observations, we aim to understand which factors affect bilingual secondary discourse, focusing on the role of school setting (i.e. balanced vs. imbalanced biliteracy), language proficiency and cognitive development (i.e. executive functioning). In particular, following on Marinis, Tsimpli, and Bongartz (submitted), we will provide support for the hypothesis that among bilingual children, advanced secondary discourse is associated with balanced bilingual education. Furthermore, we aim to identify possible factors that contribute to variability in secondary discourse ability. This issue has relevant implications for the design of bilingual teaching programs which are successful in enhancing bilingual children's secondary discourse and, more in general, literacy skills.

2. Bilingual narratives: overview of previous literature and aims of the present study

Most studies on bilingual narratives have focused on the analysis of single aspects in an isolated way, such as reference production. Overall, these studies show that bilingual children use referential expressions differently from their monolingual peers. For this divergent pattern of use, several explanations have been proposed, including cross-linguistic influence between the referential systems of the bilinguals' two languages (Serratrice 2007; Chen and Lei 2012; Torregrossa et al. 2014), language proficiency (Andreou et al. 2015) and processing constraints (Torregrossa and Bongartz, *forthcoming*; Torregrossa, Bongartz, and Tsimpli, *submitted*). These accounts are not necessarily inconsistent with each other, since bilingual reference production may be the result of how all the above mentioned factors combine in an interactive manner (Bongartz 2016).

For the purposes of this study, the analysis of reference production may not be fully appropriate, given that the task of comparing bilingual performance in each respective language is complicated by the above mentioned variety of factors, as well as by cross-linguistic differences in the inventory of referential expressions. Rather, narratives will be analyzed in terms of their overall structure and content. In particular, we will assess the narratives' complexity on the basis of different measures, such as story length, features of the narrative schema, lexical diversity and interweaving patterns for tracking referents across different episodes. These measures abstract away from language-specific form-function mappings (e.g. use of referential expressions, verbal aspect and tense) and address directly higher-order, abstract knowledge of narrative structure that can be potentially shared across the two languages.

These 'abstract' measures have been taken into account in some studies comparing narrative skills in monolingual and bilingual children. Maviş, Tunçer, and Akyildiz (2011) and Maviş, Gagarina, Tunçer, Ünal, Yelegen, and Akyildiz (2012) – both studies are conference presentations and are reported in Maviş, Tunçer, and Gagarina (2016) – analyze story structure, story complexity and use of mental state terms in Turkish narratives produced by monolingual and bilingual children (Turkish-German and Turkish-Kurdish respectively), aged between 3 and 8 years in Maviş et al. (2011) and between 4 and 6 years in Maviş et al. (2012). The authors found no significant difference between bilinguals and monolinguals for all the measures considered, contrary to the pattern observed above concerning reference production. Other studies are based on the analysis of only bilingual performance (without reference to monolinguals) and aim to understand which factors affect bilingual narrative skills. By taking into account the performance of Turkish-German bilingual children (ranged in age between 3 and 8 years), Maviş, Tunçer, and Gagarina (2016) show that bilingual performance is affected by age (the older the children, the more complex the story structure) and elicitation task (telling vs. retelling): the quality of the story improved when the children first listened to a model story. Gutiérrez-Clellen (2002) analyzes narratives produced by English-Spanish bilingual children (7–8 years old) in both languages. The results reveal that in each language, bilingual narrative production is subject to great variation, which the author ascribes to differences in vocabulary and literacy experience, even if these factors are not controlled for in the statistical analysis. Furthermore, the author finds the same effect of task as in Maviş, Tunçer, and Gagarina (2016), since variation in bilingual performance is observed only in association with retellings. Montanari (2004) analyzes narratives elicited from three Spanish-English children (5–6 years old) in two different sessions over a span of six months and observes that narrative skills are significantly affected by both age and language proficiency, along the lines of Maviş, Tunçer, and Gagarina (2016) and Gutiérrez-Clellen (2002) respectively.

In this paper, we will introduce reference to the schooling system (in terms of hours dedicated to each language) as an additional factor that may contribute to variation of bilingual narrative performance, based on our assumption that education has a crucial role in the development of secondary discourse (see Section 1). In line with the reviewed literature, we will also analyze the impact of language proficiency, as well as its interaction with education.

On the contrary, we will not consider how the type of task (story-telling vs. retelling) affects performance. From the literature, it emerges that story-telling and retelling impose different types of cognitive demands. Retellings involve more attentional and memory resources, addressed to the comprehension and storage of the model narrative (Gibbons et al. 1986). Tellings require a greater effort in terms of picture decoding and discourse planning, since there is no prior script on which subjects can rely (Gagarina et al. 2012). However, both methodologies have been shown to be efficient for assessing children's secondary discourse ability (cf. the review in Maviş, Tunçer, and Gagarina 2016). In this paper, we rely on a story-retelling task. Apart from the advantages that story-retelling has in terms of avoiding misreading of the pictures and eliciting longer and more detailed narratives (Andreou et al. 2015), we believe that story-retelling approximates in a better manner the process of learning secondary discourse, consisting of 'grasping a model' (concerning how a narrative should be structured) and applying this model to the production of a new, original piece of discourse. Due to the nature of the task and its being dependent on auditory memory, in addition to educational setting and proficiency, the results of our study will be controlled for measures of working memory capacity.

3. The study

3.1. Participants

The study was conducted at three schools in the North Rhine Westphalia in Germany. In two of the schools, German is the main medium of instruction: in the school in Cologne, Greek is taught in afternoon classes intended for Greek heritage children with 4 h per week in the 3rd and 4th grades and 2 h in the 5th and 6th grades. In the comprehensive school in Krefeld, teaching in Greek is part of the curricular program starting from the 5th grade. Subjects that are taught in Greek are Language, Natural Sciences and Religion, amounting to 8 h of instruction in Greek per week. In the last school, which is in Düsseldorf, the number of hours of teaching in Greek amounts to 21 per week and exceeds the number of hours of teaching in German, which equals 10 per week. In terms of teaching balance, the school in Düsseldorf has the smallest difference between hours of teaching in each language (see Table 1).

For each school, we considered a sample of 10 Greek-German bilingual children (30 in total). The children ranged in age from 9.3 to 12.7 years (mean age: 11.3 years) and all attended the 5th or 6th grade. Since we are interested in understanding to what extent narrative skills are affected by the educational setting, the results will be presented by grouping children per school. We will thus distinguish three groups: Cologne (COL), Krefeld (KRE) and Düsseldorf (DUS).

We tested each child for non-verbal abilities by means of the Raven's colored matrices (Raven, Raven, and Court 1998), to control for non-verbal intelligence. A one-way ANOVA reveals that the three groups do not differ in non-verbal reasoning ($F(2) = 1.67$, $p = 0.20$) – the averages for each group are reported in Table 1 below.

Verbal abilities were tested in both languages by using the productive vocabulary tasks normed for monolinguals in Greek (Vogindroukas, Protopapas, and Sideridis 2009) and German (Petermann, Fröhlich, and Metz 2010), respectively. We consider the scores in the vocabulary tests as indicators of language proficiency. The Greek vocabulary task includes 50 pictures of nouns, while the German vocabulary task is made of 40 pictures meant to elicit 30 nouns and 10 verbs. In both languages, there is a significant effect of group on the vocabulary score (one-way ANOVA, $F(2) = 8.1$, $p = .03$ for Greek vocabulary and $F(2) = 9.1$, $p = .01$ for German vocabulary). To understand this effect of group, we ran post-hoc comparisons using the Fisher's Least Significant Difference (LSD). In Greek, the Fisher's test reveals that the mean scores of DUS are significantly greater than the ones of KRE and COL, which do not differ from each other. In German, the opposite picture emerges. DUS has the lowest mean vocabulary scores, and again KRE and COL do not differ from each other. In other terms, with respect to KRE and COL, DUS children are more proficient in

Table 1. Descriptive statistics for hours of teaching, proficiency, non-verbal intelligence and working memory measures across the three groups of children.

Group	N	Hours of teaching in German	Hours of teaching in Greek		Mean	SD
DUS (Düsseldorf)	10	10 (32.26%)	21 (67.74%)	<i>Raven's colored matrixes</i>	27	2.98
				<i>Mr. X</i>	14.8	5.22
				<i>Vocabulary score Greek</i>	.85	.60
				<i>Vocabulary score German</i>	-.84	1.27
KRE (Krefeld)	10	20 (71.43%)	8 (28.57%)	<i>Raven's colored matrixes</i>	29.6	3.06
				<i>Mr. X</i>	15.6	3.2
				<i>Vocabulary score Greek</i>	-.34	.79
				<i>Vocabulary score German</i>	.19	.42
COL (Cologne)	10	28 (93.33%)	2 (6.67%)	<i>Raven's colored matrixes</i>	28.3	3.46
				<i>Mr. X</i>	14.1	4.99
				<i>Vocabulary score Greek</i>	-.05	1.01
				<i>Vocabulary score German</i>	.65	.38

Greek, but less proficient in German. Table 1 reports the mean scores and standard deviations in the vocabulary tests in both languages, based on the z-score values for comparability reasons, since the vocabulary tests in the two languages are based on different scales.

We also tested children's Working Memory capacity (WM, henceforth) by using the Mr. X task (which taps, in particular, visuo-spatial WM in the central executive, cf. Alloway 2007). Mr. X was administered by means of PowerPoint slides on a computer screen. The child had to look at two figures (one with a blue hat and one with a yellow hat) holding a ball in their hands, and indicate whether they held the ball in the same or a different hand. After each trial, the child had to recall, in sequence, the position at which the figure with the blue hat held the ball, by pointing to one of eight points on the screen. The trials are of increasing complexity, ranging from a block containing 6 trials with only one Mr. X to a block containing 6 trials with 7 Mr. Xs. There is no significant difference in WM as tested by Mr. X across the three groups ($F(2) = 0.27$, $p = 0.76$). Table 1 summarizes the information described up to this point.

In order to control for language dominance outside the school setting, each child was administered a questionnaire by means of a one-to-one interview. The questionnaire consisted of four main modules targeting each of the child's two languages: i) *home language history*, which refers to the amount of language exposure before the age of six (i.e. before attending school); ii) *early literacy preparedness*, concerning literacy input received prior to schooling; iii) *current literacy*, which is related to literacy practices outside schooling contexts (e.g. writing e-mails or letters and reading books, comics or newspapers); iv) *current language use*, which involves language input and output with family members and friends (cf. Andreou et al. 2015; Bongartz 2016; Dosi and Papadopoulou, forthcoming). For all four modules, we assigned each of the two languages a score, which was the sum of the scores from the individual answers. For answers stating that both languages were used in equal proportion, we split the associated scores between the two languages. We subtracted one language total (German) from the other (Greek): a positive score indicates dominance in Greek, while a negative score reflects dominance in German. The closer the score to zero, the more balanced the child in the corresponding module. Table 2 reports the means and standard deviations of the z-scored differences in the four modules across the three groups of children.

There is a significant effect of group on the difference between the scores in the two languages in all four modules (*home language history*: $F(2) = 3.38$, $p = .05$; *early literacy preparedness*: $F(2) = 6.92$, $p < .05$; *current literacy*: $F(2) = 22.37$, $p < .05$; *current language use*: $F(2) = 18.24$, $p < .05$). LSD post-hoc comparisons reveal that in home language history, current literacy and current language use, the DUS children differ significantly from the KRE and COL children, who do not differ from each other. In particular, the DUS children are Greek dominant, while the KRE and COL children are German dominant (Table 2). The picture is slightly different when early literacy preparedness is taken into account, in that DUS and KRE children are Greek dominant and differ significantly from COL children, who are German dominant (Table 2).

Table 2. Descriptive statistics for the four modules in the questionnaires across the three groups of children.

Module	Group	Mean	SD
<i>Difference home language history</i>	DUS	.76	.89
	KRE	−.29	.91
	COL	−.24	.96
<i>Difference early literacy preparedness</i>	DUS	.58	.48
	KRE	.25	.89
	COL	−.70	1.06
<i>Difference current literacy</i>	DUS	1.33	.79
	KRE	−.36	.47
	COL	−.57	.61
<i>Difference current language use</i>	DUS	1.23	.81
	KRE	−.23	.43
	COL	−.63	.76

Note: Positive values indicate dominance in Greek, while negative values dominance in German.

To conclude, the profiling data obtained from the questionnaires match the language dominance profile emerging from the analysis of the schooling contexts (hours of teaching), as well as the measures of proficiency (assessed on the basis of vocabulary tests). For the children in DUS, Greek is the dominant language in and outside schooling contexts. On the contrary, language usage among KRE and COL children is German dominant, with the only exception of early literacy exposure among KRE children. In this paper, we will investigate how the described dominance profile affects narrative production. However, before proceeding to the description of the study, it should be emphasized once again that the extent to which one or the other language is the main medium of instruction is different across the three schools: the DUS school is the most balanced in terms of hours of teaching dedicated to each language.

3.2. Materials

The materials consists of 60 bilingual narratives, 30 in Greek and 30 in German. The narratives were retellings elicited using the Edmonton Narrative Norms Instrument (ENNI) designed by Schneider, Dubé, and Hayward (2005). ENNI includes six stories, divided into three groups of increasing complexity. For our task, we used the two most complex stories, one for eliciting the retelling in Greek and one in German. These two stories have been designed to be structurally equivalent. Each consists of 13 pictures (with no text) representing a series of events involving two major characters (an elephant girl and a giraffe boy in one story and a dog girl and a rabbit boy in the other) and two minor ones (of different gender, too).

The task was administered as a sequence of Power Point slides on a computer screen. First, participants had to choose one of three envelopes. All the envelopes contained one of the two target stories, but children were told that the envelopes contained different stories (Serratrice 2007). Then, the participants looked at the story pictures two by two, while listening to the model story on the headphones. Finally, once the thirteen picture synopsis had appeared on the screen, they had to tell the story to the investigator, who feigned ignorance of the plot. The stories were audio-recorded and then transcribed into CHAT format (MacWhinney 2000) by a Greek and a German native speaker, respectively. We refer to Andreou et al. (2015) for further details concerning the methodology and the procedure of the experiment.

3.3. Data analysis

As pointed out in Section 2, our analysis of narratives is based on ‘abstract’ measures related to the narratives’ formal structure and content. The next sections will provide a list of the measures that we

will take into account, together with an explanation of the reasoning behind our choice of these measures. Furthermore, in line with our assumption that story-retelling involves conforming to a narrative model in a way that resembles the process of learning secondary discourse (Section 2), the results of the analysis of bilingual narratives will be interpreted with respect to the features of the model narrative in each language.

3.3.1. Story grammar

Each narrative was scored following the procedure indicated in Schneider, Dubé, and Hayward (2005). The authors identify seven components of a ‘good’ narrative, i.e. setting, initiating event, internal response (i.e. a response to an initiating event), internal plan (i.e. how characters intend to deal with the initiating event), attempt, outcome and reaction (i.e. a response to an outcome). Among these components, initiating event, attempt and outcome are considered as ‘core’ elements. The same component may occur several times within the narrative. For instance, each character (major or minor) may have his/her own internal plan or reaction to the events. To score the stories, presence of the core components is indicated by assignment of two points, while the other components are assigned one point. (1) exemplifies the scoring procedure, based on some extracts from one of the two model narratives in German.

(1)	EXTRACT	COMPONENT	SCORE
	Eines Tages entschlossen sich ein verspieltes Hundemädchen und ein fröhlicher Hasenjunge, die Freunde waren, im Wald spazieren zu gehen (...) [One day, a playful female dog and a happy rabbit boy, who were friends, decided to go for a walk in the wood (...)]	SETTING	1
	Aber aus Versehen flog der Ballon Haso dabei aus der Händen (...) [But suddenly the balloon flew away from Haso's hands (...)].	INITIATING EVENT	2
	Während Hundina hoch sprang, um ihn zu fassen, rief sie: 'Oh nein, mein Lieblingsballon fliegt davon!' (...) [While Hundina jumped high, to take it, she cried: 'Oh no! My favorite balloon flies away!']	INTERNAL RESPONSE	1
	Er dachte, dass die einzige Möglichkeit war, seine Freundin wieder glücklich zu machen, ihr dort einen neuen Ballon zu holen. [He thought that the only possibility to make his female friend happy again was to take there a new balloon for her]	INTERNAL PLAN	1
	Er lief zu dem alten Hasen so schnell er konnte, und fragte ihn nach dem schönsten Luftballon, den er hatte, um ihn seiner traurigen Freundin zu schenken. [He went to the old rabbit as fast as he could and asked him the most beautiful balloon that he had, to give it to his sad female friend (...)]	ATTEMPT	2
	(...) fand aber kein Geld (...) [(...) but he found no money (...)]	OUTCOME	2
	Da war er traurig (...) [Then he was sad (...)]	REACTION	1

The results related to the bilingual narratives will be reported as the difference between the scores of the model narratives and the scores of the respective child narratives.

3.3.1.1. Mental state words. The use of mental state words is one of the principal indicators of advanced story-telling: the complexity of the narrative is enhanced by ‘confrontations between intentional action and obstacles (...) and characters’ interacting plans’ (Francis 2006, 42). We aim to understand whether high scores in story grammar (Section 3.3.1) correlate with the expression of plans, reactions, intentions and feelings through the use of mental state words. To this purpose, we counted the types of mental state terms occurring in each narrative. These include adjectives of emotion (e.g. happy, sad, angry, impatient) and verbs of volition (e.g. want, decide), affect (admire, feel, fear, to be jealous, to be worried, agree, like) and cognition (e.g. think, know). The results are reported as the difference between the model stories and the child stories in number of mental state word types (normalized for the square root of the number of content word types).

3.3.2. Story length and syntactic complexity

Across the literature on child narratives, story length and syntactic complexity are considered as reliable indicators of the development of narrative skills (a.o., Gillam and Johnston 1992; Newman and McGregor 2006; Altman et al. 2016). We divided each narrative into clauses, defined by the occurrence of a verbal form. Story length corresponds to the number of clauses. As a second step, for each clause, we indicated whether it was a main or a subordinate clause. For subordinate clauses, we did not consider their depth of embedding (e.g. first vs. second level of embedding) or their finiteness (finite vs. non-finite). Syntactic complexity is measured as the ratio between the number of subordinate clauses and the total number of clauses. As was the case of story grammar, the results concerning story length and syntactic complexity will be based on the difference between the scores in the model stories and the child stories.

3.3.2.1. Complex DPs. Ravid and Berman (2010) use the production of heavy, complex noun phrases (here DPs) as a diagnostic of literacy levels. We counted the occurrence of DPs containing qualitative adjectives (e.g. *nice, interesting, clever, old*) – excluding numeral, demonstrative and possessive adjectives (e.g. *three, this, other, his*) – and/or followed by a relative clause. We normalized the number of complex DPs for the square number of definite and indefinite DPs occurring in the narrative. Finally, we calculated the difference between the model stories and the children's stories in the occurrence of complex DPs.

3.3.3. Lexical diversity

In many studies, it has been pointed out that the level of lexical diversity correlates positively with the perception of the speaker's linguistic competence and effectiveness of the message (cf. Bradac, Davies, and Courtright 1977). Therefore, lexical diversity is usually taken as an index of the quality of a narrative (e.g. Berman and Bracha 2007). In our paper, lexical diversity is calculated as the ratio between the number of content word types (nouns, verbs, adjectives, adverbs) and the square root of tokens (Guiraud 1960). The presentation of the results is based on the difference between the scores associated with the model stories and the ones associated with the child stories.

3.3.4. Shifts in argument role

Stories involving several characters prove challenging in early development: children follow a thematic strategy and organize the narrative around a single protagonist (Hickmann and Hendriks 1999). Mentioning different characters and switching between them are associated with later stages in the development of secondary discourse (Bongartz 2016). Here we argue that the quality of the narrative depends on how these switches are realized. For instance, the narrative extract in (2) contains several switches from one character (the female dog) to the other (the rabbit), but the story unfolds in an additive manner: the units seem to be juxtaposed to each other. On the contrary, the narrative extract in (3) exploits a more 'dynamic' pattern of switches (from subject to non-subject and vice versa), which renders the narrative more cohesive.

- (2) **Die Hündin** hatte eine Karre und da hatte sie einen Ballon. **Der Haso** hat das gesehen und der wollte mit dem spielen. **Die Hündin** hat gesagt: 'Wir müssen das erstmal raus # also binden'. Danach hat **der Haso** versucht ...

[The female dog had a cart and there she had a balloon. The male rabbit saw that and he wanted to play with it. The female dog said: 'We have first to away # uhm untie it'. Then the rabbit tried ...] – DUS18; age: 10.6

- (3) Und **der Hase** hat gesehen, dass **seine Freundin** eine Schubkarre mit hatte, an den ein Ballon angefestet war. Und da wollte **der Hase** den abmachen, um mit **seiner Freundin** zu spielen. Und **die Hündin** hat **zu ihm** gesagt, **er** muss erstmal ausbinden, damit **er** den Ballon kriegt.

[And the dog saw that his female friend had a cart with her, to which a balloon was bound. And then the rabbit wanted to untie it, to play with his female friend. And the female dog told him that he has to untie it first, so that he takes the balloon ...] – DUS20; age: 10.0

The more complex, dynamic interweaving pattern in (3) reflects more advanced literacy skills. To assess the complexity of the weaving pattern for reference tracking, we counted, for all story

characters, the switches from subject role to non-subject, as well as the switches from non-subject to subject. We then normalized the results for the total number of mentions of all characters in the story. The statistical analysis was based on the difference values between the model narratives and the respective child narratives.

3.3.5. Correlational analyses

The results of the analysis of the measures considered in this section were correlated, for each of the two bilinguals' languages, with hours of teaching (in each respective language), measures of language proficiency and working memory. Finally, we investigated to what extent the quality of the performance in the weaker language at school (German for DUS and Greek for KRE and COL) correlates with the amount of instruction in that language. This allows us to identify the role of balanced biliteracy in the development of secondary discourse.

4. Results

Table 3 presents an overview of the results for each of the (main) five modules considered in the preceding section and reports the average scores and the standard deviations for each group of children.

4.1. Story grammar

In the Greek narratives, the three groups of children differ significantly from each other in the values of the difference between the story grammar score in the model stories and the respective child stories (one-way ANOVA: $F(2) = 6.76, p < .05$). Post-hoc tests reveal that the difference values associated with DUS children are lower than the ones of KRE and COL children (Fischer's LSD: DUS-KRE, $p = .04$; DUS-COL, $p = .001$). In other terms, DUS children's narratives are the closest to the story model. KRE children tend to perform better than COL children, but the two groups do not differ from each other (Fischer's LSD: KRE-COL, $p = .14$). A significant negative strong correlation was found between the difference scores and hours of teaching ($r = -.57, p = .001$): the greater the number of hours of teaching in Greek, the closer the child's narrative to the model story. The correlation between the difference scores and the vocabulary scores appears to be weaker ($r = -.42, p = .02$). On the contrary, we found no significant correlation between difference scores and WM.

The picture emerging from the analysis of the German narratives is completely different. The three groups of children do not differ significantly from each other ($F(2) = .19, p > .05$) and there is no

Table 3. Descriptive statistics for each group of children in each of the five measures of narrative complexity.

Measure		Greek		German	
		Mean	SD	Mean	SD
Story grammar	DUS	7.4	2.8	9.3	.86
	KRE	10.7	2.3	9.4	1.1
	COL	13.0	4.7	8.6	.99
Story length	DUS	30.1	8.0	30.9	6.3
	KRE	34.7	6.0	30.0	4.3
	COL	41.7	9.5	31.2	4.2
Syntactic complexity	DUS	2.85	.60	3.1	.58
	KRE	2.95	.45	3.1	.39
	COL	3.48	.58	3.1	.43
Lexical diversity	DUS	2.0	.48	3.38	.74
	KRE	3.0	.59	2.37	.56
	COL	3.1	.65	2.18	.41
Shifts in argument role	DUS	.32	.28	.76	.60
	KRE	.69	.41	1.03	.69
	COL	.67	.43	.78	.66

Note: The results are divided for Greek and German.

significant correlation between difference scores, on the one hand, and hours of teaching, vocabulary score and WM on the other.

Also, the three groups differ with respect to the performance in story grammar in the non-dominant language at school ($F(2) = 3.01, p = .06$). DUS children in German perform better than COL children in Greek (Fischer's LSD: DUS-COL, $p = .02$). KRE children (in Greek) exhibit an intermediate behavior between DUS and COL, and do not differ significantly from either of the two groups (Fischer's LSD: DUS-KRE, $p = .37$; KRE-COL, $p = .14$). Performance in the weaker language at school (expressed again in terms of difference scores) correlates negatively with hours of teaching in that language ($r = -.42, p = .02$).

4.1.1. Mental state words

The results related to the use of mental state words pattern similarly to the ones concerning story grammar. In Greek, the difference between the model stories and the children's narratives is less pronounced among DUS children ($M = .40, SD = .22$) than KRE ($M = .66, SD = .17$) and COL children ($M = .74, SD = .15$). A one-way ANOVA reveals a significant group effect ($F(2) = 8.51, p = .002$). DUS children differ significantly from both KRE and COL children, who do not differ from each other (Fischer's LSD: DUS-KRE, $p = .007$; DUS-COL, $p = .001$, KRE-COL = .40). Furthermore, the difference scores in the use of mental state words correlate negatively with hours of teaching in Greek ($r = -.65, p < .001$).

In German, the three groups of children do not differ from each other ($F(2) = .12, p > .05$ – for DUS: $M = .56, SD = .30$; for KRE: $M = .62, SD = .37$; for COL: $M = .63, SD = .19$), and no significant correlation is found with hours of teaching, vocabulary score and WM.

The analysis related to the performance in the non-dominant language at school reveals no significant group effect ($F(2) = 2.21, p = .13$). However, separate t -tests show that the difference between DUS and COL children approaches significance ($t(18) = -1.92, p = .06$). On the contrary, KRE children do not differ from either group (t -tests: DUS-KRE, $t(18) = -1.18, p = .26$; COL-KRE, $t(18) = -.99, p = .33$). The correlation between performance in the weaker language and hours of teaching in that language is nearly significant ($r = -.36, p = .07$).

4.2. Story length and syntactic complexity

When considering story length in Greek narratives, DUS children are closer to the story model, as compared to KRE and COL children ($F(2) = 5.38, p = .01$). Post-hoc tests reveal that, while DUS and KRE children do not differ from each other, they significantly differ from COL children (Fischer's LSD: DUS-KRE, $p = .21$; DUS-COL, $p = .003$, KRE-COL, $p = .05$). The correlational analyses reveal a significant relationship between difference scores in story length and hours of teaching ($r = -.51, p = .004$), but no association with vocabulary score or WM.

In German, the three groups of children do not differ from each other ($F(2) = .15, p > .05$). No correlation is found between the difference scores in story length and hours of teaching, vocabulary score or WM.

The comparison of the performance in the non-dominant language at school across the three groups shows a significant group effect ($F(2) = 5.46, p = .01$). DUS and KRE children differ from COL children (Fischer's LSD: DUS-COL, $p = .003$; KRE-COL, $p = .04$), but do not differ significantly from each other (Fischer's LSD DUS-KRE, $p = .26$). Furthermore, the scores in the non-dominant language strongly correlate with hours of teaching in that language ($r = -.53, p = .003$).

The analysis of syntactic complexity in Greek reveals a significant group effect ($F(2) = 3.8, p = .03$). In particular, DUS and KRE children differ from COL children (Fischer's LSD: DUS-COL, $p = .02$; KRE-COL, $p = .03$), while they do not differ from each other (Fischer's LSD DUS-KRE, $p = .79$). There is a moderate correlation between the difference values in syntactic complexity and hours of teaching ($r = -.38, p = .04$), but no correlation with vocabulary score or WM.

In German, the three groups do not differ from each other. Furthermore, no significant correlation is found between syntactic complexity, hours of teaching, vocabulary score and WM.

The children's performance in syntactic complexity in the weaker language at school does not differ across the three groups ($F(2) = 2.6, p = .10$). However, separate t -tests show that COL children differ significantly from DUS and KRE children (t -tests: DUS-COL: $t(18) = -1.9, p = .05$; KRE-COL: $t(18) = -2.3, p = .04$), producing narratives of lower syntactic complexity. On the contrary, DUS and KRE children do not differ significantly from each other (t -test: DUS-KRE: $t(18) = .56, p = .58$). Moreover, we found a moderate correlation between syntactic complexity in the weaker language at school and hours of teaching in that language ($r = -.37, p = .05$).

4.2.1. Complex DPs

In Greek, the analysis concerning the occurrence of complex DPs reveals a significant group effect ($F(2) = 3.32, p = .05$). Post-doc tests show that DUS children ($M = 1.5, SD = .20$) differ significantly from COL ($M = 1.77, SD = .24$) and KRE children ($M = 1.74, SD = .21$), who do not differ from each other (Fischer's LSD: DUS-KRE, $p = .02$; DUS-COL, $p = .04$; KRE-COL, $p = .83$). We found that the occurrence of complex DPs in Greek correlates with hours of teaching ($r = -.42, p = .03$).

In German, the occurrence of complex DPs does not differ across the three groups ($F(2) = 2.37, p = .11$). DUS: $M = .36, SD = .46$; KRE: $M = .64, SD = .40$; COL: $M = .71, SD = .25$). No significant correlation is found with hours of teaching, vocabulary score and WM.

The occurrence of complex DPs in the weaker language at school differs across the three groups ($F(2) = 6.17, p = .001$). In particular, DUS children produce in German more complex DPs than KRE and COL children in Greek (Fischer's LSD: DUS-KRE, $p < .001$; DUS-COL, $p < .001$; KRE-COL, $p = .88$). Finally, the occurrence of complex DPs in the weaker language correlates strongly with hours of teaching in that language ($r = -.56, p = .001$).

4.3. Lexical diversity

In Greek, we found a significant effect of group on lexical diversity ($F(2) = 11.68, p < .001$). Post-hoc tests reveal that DUS children outperform KRE and COL children, who do not differ from each other (Fischer's LSD: DUS-KRE, $p < .001$; DUS-COL, $p < .001$; KRE-COL, $p = .81$). We found a strong correlation of lexical diversity difference scores with both vocabulary score ($r = -.70$ and $p < .01$) and hours of teaching ($r = -.66, p < .01$). On the contrary, there is no significant correlation with WM.

The analysis of German narratives reveals an opposite trend. As is the case of Greek narratives, we found a significant effect of group on lexical diversity ($F(2) = 12.21, p < .001$). However, DUS children lag behind KRE and COL children, who do not differ from each other (Fischer's LSD: DUS-KRE, $p = .001$; DUS-COL, $p < .001$; KRE-COL, $p = .48$). Furthermore, lexical diversity difference scores are strongly correlated with both vocabulary score ($r = -.66, p < .001$) and hours of teaching in German ($r = -.68, p < .001$), but no correlation with WM is observed.

Contrary to the pattern observed with story grammar and story length, the performance in lexical diversity in the weaker language at school does not differ across the three groups ($F(2) = .55, p > .05$), nor there is any correlation with hours of teaching in that language.

4.4. Shifts in argument role

In Greek, a one-way ANOVA with difference values in number of shifts as dependent variable reveals an effect of group that approaches significance ($F(2) = 3.02, p = .06$). In particular, DUS children are the closest to the model story (Fischer's LSD: DUS-KRE, $p = .03$; DUS-COL, $p = 0.5$; KRE-COL, $p = .87$). The difference values correlate negatively with hours of teaching ($r = -.57, p = .002$) and, to a lesser extent, with vocabulary score ($r = -.40, p = .03$).

The analysis of the German narratives offers a different picture. We found no group effect in the difference values nor any significant correlation with vocabulary score, hours of teaching and WM.

Comparing the realization of argument shifts in the weaker language at school across the three groups shows that there is a significant group effect ($F(2) = 5.89, p = .008$), in that DUS children

outperform KRE and COL children, who do not differ from each other (Fischer's LSD: DUS-KRE, $p = .004$; DUS-COL, $p = .009$, KRE-COL, $p = .75$). Moreover, there is a weak correlation between the difference values in argument shifts and hours of teaching, which is nearly significant ($r = -.31$, $p = .07$).

5. Summary of the results and discussion

The aim of this study was to examine the benefits of balanced biliteracy for the development of secondary discourse ability and, more in general, to understand which factors contribute to variation in narrative skills. In particular, we considered the effects related to the amount of literacy instruction (hours of teaching) and language proficiency (vocabulary experience), controlling for executive functions. The study was based on the analysis of narrative skills of three groups of Greek-German bilingual children, who are educated in three different school systems (varying in the amount of hours dedicated to Greek and German respectively) and differ in their degree of proficiency in each respective language. Table 4 offers an overview of the results of the three groups of children in the main narrative measures considered in this study. In each box, we report the three groups ordered according to the quality of their narrative production and the factors affecting their performance, ordered from the best to the worst predictor (only significant correlations are reported).

Let us start our discussion considering the two groups of children that when compared to each other in terms of profiling data, exhibit the most divergent results. On the one hand, the vocabulary testing and the demographic and history information extracted from the questionnaires show that DUS children are Greek dominant. Moreover, the same group of children attends a school in which Greek is the main medium of instruction. On the other hand, COL children are German dominant (according to the same measures) and are also exposed to a German dominant education. If secondary discourse ability were affected only by language proficiency and dominance effects, one would expect DUS children to have better narrative skills than COL children in Greek, and vice versa, COL children to do better in German. However, the results of our study show that this hypothesis is only partially correct. While it is true that in Greek DUS children outperform their COL peers for all the indicators of secondary discourse introduced in this study, their performance in German does not differ from the one of COL children (with the only exception of lexical diversity, fn. 2).

In several parts of this paper, we pointed out that the schooling systems in DUS and COL differ in their degree of balanced biliteracy: the difference in the amount of instruction in each language is stronger in COL than in DUS. Accordingly, we interpret the divergence of the results related to DUS and COL children as showing that in narrative production balanced biliteracy has the effect of counterbalancing language proficiency and dominance effects, as well as enhancing secondary discourse also in the weaker language. Crucial evidence in favor of this hypothesis comes from the analysis of the children's performance in the weaker language (i.e. non-dominant in terms of proficiency and education). For all the indicators of secondary discourse (with exception of lexical diversity – fn. 2), DUS children in German outperform COL children in Greek. Furthermore, we showed that the

Table 4. Summary of the results.

	GREEK	GERMAN	WEAKER LANGUAGE AT SCHOOL
<i>Story grammar</i>	DUS > KRE = COL hours of teaching > vocabulary	DUS = KRE = COL	DUS > (=) KRE > (=) COL hours of teaching
<i>Story length</i>	DUS = KRE > COL hours of teaching	DUS = KRE = COL	DUS = KRE > COL hours of teaching
<i>Syntactic complexity</i>	DUS = KRE > COL hours of teaching	DUS = KRE = COL	DUS = KRE > COL hours of teaching
<i>Lexical diversity</i>	DUS > KRE = COL vocabulary > hours of teaching	KRE = COL > DUS hours of teaching > vocabulary	DUS = KRE = COL
<i>Shifts</i>	DUS > KRE = COL hours of teaching > vocabulary	DUS = KRE = COL	DUS > KRE = COL hours of teaching

Note: In each box, the groups are ordered for quality of performance and the factors for efficacy of prediction.

quality of the performance in the weaker language always correlates with the amount of hours dedicated to that language at school.

The results concerning the narrative production by KRE children is fully consistent with the outlined picture. This group of children does not differ from COL children in language proficiency and in the results of three of the four modules in the questionnaires (home language history, current literacy and current language use). However, their performance in Greek is not comparable to the performance of their COL peers, since KRE children perform similarly to DUS children in some measures of secondary discourse (e.g. story length and syntactic complexity) and to COL children in other measures (e.g. story grammar and shifts in argument role). This intermediate behavior is not surprising, given that their degree of biliteracy is much more balanced than COL children, but (slightly) less balanced than DUS children.¹

To summarize, our study shows that the combination of (relatively) low proficiency with imbalanced biliteracy leads to a reduced performance in secondary discourse ability. On the contrary, balanced literacy has the effect of evening out the negative effects of low proficiency and facilitating the 'sharing' of secondary discourse between the two bilinguals' languages. The extent to which different skills can be shared and the factors involved in this sharing will be investigated in future work, by directly comparing the bilingual performance in each respective language (cf. Knopp 2017).

With respect to the other issue raised in this paper – concerning the factors involved in variation of secondary discourse – the analysis of the Greek narratives and of the narratives produced in the weaker language at school is quite revealing. The amount of instruction in Greek or in the weaker language (in terms of hours of teaching) is often the best predictor of the children's performance.² This supports the idea that secondary discourse is the outcome of education and formal instruction (Francis 2006; Bongartz 2016). The same correlation between children's performance and amount of instruction was not found when analyzing children's narrative skills in German. Again, this is due to the fact that, among the children in DUS, the lower amount of hours in German is compensated by the benefits of balanced biliteracy. This result suggests that quantitative data are not sufficient to account for children's literacy preparedness. Qualitative data involving an in-depth analysis of the different educational contexts play an important role, too.

Notes

1. The analysis of the questionnaires shows that KRE children are Greek dominant in early literacy. It is thus likely that early literacy interacts with balanced biliteracy in enhancing secondary discourse. In particular, we showed that in all measures of syntactic complexity (story length and use of subordinate clauses) with the exception of the production of complex DPs, KRE children perform in Greek similarly to their peers in DUS. This suggests that early literacy has a positive correlation with syntactic complexity abilities, in line with what has been proposed in the literature (cf. Scarborough et al. 1991; Walker et al. 1994).
2. The only exception to this generalization is lexical diversity, for which language proficiency plays a key role. We propose that the reason for this is twofold. First, in this study language proficiency is assessed based on vocabulary knowledge. Second, among the indicators of secondary discourse, lexical diversity may be the less 'abstract'. In other terms, even if the child knows that a good narrative involves variation in the lexical choices, his performance in lexical diversity is constrained by proficiency to a greater extent than with the other indicators of narrative skills.

Acknowledgements

The data presented in this paper have been collected and analyzed within the CoLiBi project – Cognition, Literacy and Bilingualism in Greek-German speaking children (principal investigators: Christiane Bongartz and Ianthi Tsimpli), jointly founded by IKY (Greek State Scholarship Foundation) and DAAD (German Academic Exchange Service). We thank Maria Andreou, Eva Knopp, Ianthi Tsimpli, our anonymous reviewers and the audience of the BALED workshop (23–24 April 2015, University of Thessaloniki) for their precious feedback. All remaining errors are ours.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Deutscher Akademischer Austauschdienst; BALED: [grant number MIS377313]; IKY (Greek Scholarships Foundation).

Notes on contributors

Christiane Bongartz is Professor of English Linguistics at the University of Cologne. Her research covers a range of theoretical and applied issues in language acquisition in various contexts (L1, 2L1 and L2), with particular reference to the interaction between language-internal factors and acquisitional context(s).

Jacopo Torregrossa is Professor of Romance Linguistics at the University of Hamburg. His research interests focus on monolingual and bilingual acquisition of linguistic phenomena at the syntax-semantics and syntax-phonology interface and the interaction between language and cognitive abilities. Christiane Bongartz and Jacopo Torregrossa are Principal Investigators of the project "Reference management in bilingual narratives" of the SFB 1252 "Prominence in Language" at the University of Cologne.

References

- Alloway, T. P. 2007. *Automated Working Memory Assessment*. London: Harcourt Assessment.
- Altman, C., S. Armon-Lotem, S. Fichman, and J. Walters. 2016. "Macrostructure, Microstructure, and Mental State Terms in the Narratives of English-Hebrew Bilingual Preschool Children with and Without Specific Language Impairment." *Applied Psycholinguistics* 37 (1): 165–193.
- Andreou, M., E. Knopp, C. Bongartz, I. M. Tsimpli. 2015. "Character Reference in Greek-German Bilingual Children's Narratives." In *EUROSLA Yearbook 2015*, edited by L. Roberts, K. McManus, N. Vanek and D. Trenkic, 1–40. Amsterdam: John Benjamins.
- Arnold, J., S. Brown-Schmidt, and J. C. Trueswell. 2007. "Children's Use of Gender and Order-of-Mention During Pronoun Comprehension." *Language and Cognitive Processes* 22: 527–565.
- Berman, R., and N. Bracha. 2007. "Comparing Narrative and Expository Text Construction Across Adolescence: A Developmental Paradox." *Discourse Processes* 43 (2): 79–120.
- Bialystok, E. 1993. "Metalinguistic Awareness: The Development of Children's Representations of Language." In *System of Representation in Children: Development and Use*, edited by C. Pratt and A. Garton, 211–233. London: Wiley & Sons.
- Bialystok, E. 2001. *Bilingualism in Development: Language, Literacy and Cognition*. New York: Cambridge University Press.
- Bialystok, E., S. Majumder, and M. M. Martin. 2003. "Developing Phonological Awareness: is There a Bilingual Advantage?" *Applied Psycholinguistics* 24: 27–44.
- Bliss, L., A. McCabe, and E. Miranda. 1998. "Narrative Assessment Profile: Discourse Analysis for School-Age Children." *Journal of Communication Disorders* 31 (4): 347–363.
- Bongartz, C. 2016. "Bilingual and Second Language Development and Literacy – Emerging Perspectives on an Intimate Relationship." Proceedings of the 21st international symposium of theoretical and applied linguistics 2013, Thessaloniki, April 5–7, 2013.
- Bradac, J. J., R. A. Davies, and J. A. Courtright. 1977. "The Role of Prior Message Context in Evaluative Judgments of High- and Low-Diversity Messages." *Language and Speech* 20: 295–307.
- Chen, L., and J. Lei. 2012. "The Production of Referring Expressions in Oral Narratives of Chinese-English Bilingual Speakers and Monolingual Peers." *Child Language Teaching and Therapy* 29 (1): 41–55.
- Cummins, J. 1980. "The Cross-Linguistic Dimensions of Language Proficiency: Implications for Bilingual Education and the Optimal Age Issue." *Tesol Quarterly* 14 (2): 40–52.
- Dosi, I., and D. Papadopoulou. forthcoming. "The Role of Educational Setting in the Acquisition of Aspect and Cognitive Development: Evidence from Greek-German Bilingual Children." *International Journal of Bilingual Education and Bilingualism*.
- Fang, Z. 2001. "The Development of Schooled Narrative Competence among Second Graders." *Reading Psychology* 22: 205–223.
- Francis, N. 2006. "The Development of Secondary Discourse Ability and Metalinguistic Awareness in Second Language Learners." *International Journal of Applied Linguistics* 16 (1): 37–60.
- Gagarina, N. 2016. "Narratives of Russian-German Preschool and Primary School Bilinguals: *Rasskaz* and *Erzaehlung*." *Applied Psycholinguistics* 37 (1): 91–122.

- Gagarina, N., D. Klop, S. Kunnari, K. Tantele, T. Välimaa, I. Balčiūnienė, U. Bohnacker, and J. Walters. 2012. *MAIN: Multilingual Assessment Instrument for Narratives*. ZAS Papers in Linguistics 56. Berlin: ZAS.
- Galambos, S. J., and S. Goldin-Meadow. 1990. "The Effects of Learning Two Languages on Levels of Metalinguistic Awareness." *Cognition* 34: 1–56.
- Gibbons, J., D. R. Anderson, R. Smith, D. E. Field, and C. Fischer. 1986. "Young Children's Recall and Reconstruction of Audio and Audiovisual Narratives." *Child Development* 57 (4): 1014–1023.
- Gillam, R. B., and J. R. Johnston. 1992. "Spoken and Written Language Relationship in Language/Learning-Impaired and Normally Achieving School-Age Children." *Journal of Speech Language and Hearing Research* 35 (6): 1303–1315.
- Gough, D., and Z. Bock. 2001. "Alternative Perspectives on Orality, Literacy and Education: A View from South Africa." *Journal of Multilingual and Multicultural Development* 22: 95–111.
- Guiraud, P. 1960. *Problèmes et méthodes de la statistique linguistique*. Paris: Presses Universitaires de France.
- Gutiérrez-Clellen, V. 2002. "Narratives in Two Languages: Assessing Performance of Bilingual Children." *Linguistics and Education* 13: 175–197.
- Hickmann, M., and H. Hendriks. 1999. "Cohesion and Anaphora in Children's Narratives: A Comparison of English, French, German and Chinese." *Journal of Child Language* 26: 419–452.
- Knopp, E. 2017. From bilingual to biliterate: Narrative discourse abilities in Greek-German and Greek-English bilinguals from Germany and the U.S. Doctoral diss., University of Cologne.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marinis, T., I. M. Tsimpli, and C. Bongartz. submitted. "Biliteracy Education Impacts on Cognition Selectively."
- Mavis, I., N. Gagarina, M. Tunçer, O. Ünal, D. Yelegen, and D. Akyildiz. 2012. "Administering the telling (Baby birds/Baby goats) & re-telling (Cat/Dog) stories to Turkish-German bilingual children living in Berlin." Paper presented at COST ACTION IS0804 Sixth MC & WGs Meeting, Berlin.
- Mavis, I., M. Tunçer, and D. Akyildiz. 2011. "Administering the Baby Birds to bilingual children speaking Turkish-Kurdish (TR/KR) and monolingual children speaking Turkish." Poster session presented at COST ACTION IS0804 Fifth MC & WGs Meeting, Malta.
- Maviş, I., M. Tunçer, and N. Gagarina. 2016. "Macrostructure Components in Narrations of Turkish-German Bilingual Children." *Applied Psycholinguistics* 37 (1): 69–89.
- McCabe, A., and P. Rosenthal Rollins. 1994. "Assessment of Preschool Narrative Skills." *American Journal of Speech-Language Pathology* 3: 45–56.
- Miller, J. F., J. Heilmann, A. Nockerts, A. Iglesias, L. Fabiano, and D. J. Francis. 2006. "Oral Language and Reading in Bilingual Children." *Learning Disabilities Research and Practice* 21 (1): 30–43.
- Montanari, S. 2004. "The Development of Narrative Competence in the L1 and L2 of Spanish-English Bilingual Children." *International Journal of Bilingualism* 8: 449–497.
- Newman, R. M., and K. K. McGregor. 2006. "Teachers and Laypersons Discern Quality Differences Between Narratives Produced by Children With or Without SLI." *Journal of Speech Language and Hearing Research* 49: 1022–1036.
- Nicolopoulou, A., J. McDowell, and C. Brockmeyer. 2006. "Narrative Play and Emergent Literacy: Storytelling and Story-Acting Meets Journal Writing." In *Play = Learning*, edited by D. Singer, R. Golinkoff, and K. Hirsh-Pasek, 124–144. New York: Oxford University Press.
- Petermann, F., L. P. Fröhlich, and D. Metz. 2010. SET 5-10. Sprachstandserhebung für Kinder im Alter von 5-10 Jahren. Göttingen: Hochgrefe.
- Raven, J., J. C. Raven, and J. H. Court. 1998. *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices*. San Antonio, TX: Harcourt Assessment.
- Ravid, D., and R. A. Berman. 2010. "Developing Noun Phrase Complexity at School Age: A Text-Embedded Cross-Linguistic Analysis." *First Language* 30: 3–26.
- Ricciardelli, L. A. 1992. "Bilingualism and Cognitive Development in Relation to Threshold Theory." *Journal of Psycholinguistic Research* 21: 301–316.
- Sanz, C. 2012. "Multilingualism and Metalinguistic Awareness." In *The Encyclopedia of Applied Linguistics*, edited by C. A. Chapelle, 3933–3942. Wiley: Wiley-Blackwell.
- Scarborough, H. S., L. Rescorla, H. Tager-Flusberg, A. Fowler, and V. Sudhalter. 1991. "The Relation of Utterance Length to Grammatical Complexity in Normal and Language-Disordered Groups." *Applied Psycholinguistics* 12: 23–45.
- Schneider, P., V. R. Dubé, and D. Hayward. 2005. "The Edmonton Narrative Norms Instrument." Accessed [03 October, 2015] from University of Alberta Faculty of Rehabilitation Medicine website: <http://www.rehabresearch.ualberta.ca/enni>.
- Serratrice, L. 2007. "Referential Cohesion in the Narratives of Bilingual English-Italian Children and Monolingual Peers." *Journal of Pragmatics* 39 (6): 1058–1087.
- Shrubshall, P. 1997. "Narrative, Argument and Literacy: A Comparative Study of the Narrative Discourse Development of Monolingual and Bilingual 5-10-Year-Old Learners." *Journal of Multilingual and Multicultural Development* 18: 402–421.
- Torregrossa, J., and C. Bongartz. forthcoming. "Activation of Referents in the Bilingual Mind." In *Individual Difference in Anaphora Resolution: Language and Cognitive Effects*, edited by G. Fotiadou and I. M. Tsimpli. Amsterdam: John Benjamins.

- Torregrossa, J., C. Bongartz, and I. Tsimpli. [submitted](#). "Bilingual reference production: A cognitive-computational account." *Linguistic Approaches to Bilingualism*.
- Torregrossa, J., C. Bongartz, E. Knopp, M. Andreou, I. M. Tsimpli, and M. Kaltsa. [2014](#). "The Role of Accessibility in Bilingual Referential Cohesion." Paper presented at the workshop "acquisition in different circumstances", *Congresso Internazionale della Società di Linguistica Italiana*, University of Udine, September 25–27.
- Vogindroukas, I., A. Protopapas, and G. Sideridis. [2009](#). *Experiment on the Expressive Vocabulary*. Chania, Crete: Glafki.
- Walach, G. P. [2008](#). *Language Intervention for School-Age Students: Setting Goals for Academic Success*. St. Louis, MO: Mosby.
- Walker, D., C. Greenwood, B. Hart, and J. Carta. [1994](#). "Prediction of School Outcomes Based on Early Language Production and Socioeconomic Factors." *Child Development* 65: 606–621.