

UC Berkeley

UC Berkeley Previously Published Works

Title

Lexical access and competition in bilingual children: The role of proficiency and the lexical similarity of the two languages

Permalink

<https://escholarship.org/uc/item/4f4021hj>

Authors

Persici, Valentina
Vihman, Marilyn
Burro, Roberto
[et al.](#)

Publication Date

2019-03-01

DOI

10.1016/j.jecp.2018.10.002

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Lexical access and competition in bilingual children

Lexical access and competition in bilingual children:

The role of proficiency and the lexical similarity of the two languages*

Valentina Persici^{a, b}, Marilyn Vihman^c, Roberto Burro^a, Marinella Majorano^a

^aDepartment of Human Sciences, University of Verona, Lungadige Porta Vittoria 17, 37129 Verona, Italy. Email: v.persici@campus.unimib.it; roberto.burro@univr.it; marinella.majorano@univr.it

^bDepartment of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126, Milan, Italy.

^cDepartment of Language and Linguistics Science, University of York, Heslington, York YO10 5DD, United Kingdom. Email: marilyn.vihman@york.ac.uk

*Acknowledgements

We would like to thank Elisa Muzzolon, Alice Daldosso, Valentina Amistadi, Marika Morelli, and Tamara Bastianello for their help with the data collection, Kremena Koleva and Márton Sóskuthy for their advice on the statistical analysis, and all the children who took part in the study, their families, and their teachers.

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

Declarations of interest: none

Word count: 10,335

Address for correspondence

Correspondence concerning this article should be addressed to Valentina Persici. Present address: Department of Psychology, University of Milano-Bicocca, Building U6, room 3170 b, Piazza dell'Ateneo Nuovo 1, 20126, Milan, Italy. Email: v.persici@campus.unimib.it.

Abstract

Using a picture-auditory-word recognition task we examine how early child bilinguals access their languages and how the languages affect one another. Accuracy and response times in false friends and semantically related words are compared to control conditions within and across languages and grades. Study 1 tests the performance of school-age children with balanced vs. unbalanced knowledge of L1 Italian and L2 German. Study 2 compares unbalanced bilingual children with L1 Italian and L2 French or German to investigate the effect of lexical similarity in the children's languages. The children were found to activate both languages upon receiving an auditory stimulus: Performance in each language was affected by proficiency in the other, degree of between-language similarity, and length of experience with each language. The BLINCS model is invoked as a plausible framework for conceptualizing the nature of bilingual phono-lexical representation and its effect on word recognition.

Keywords: child bilingual; Italian; German; French; false friends; cross-linguistic similarity

For over 20 years studies of the language processing of adult bilinguals have demonstrated that, far from accessing their languages independently when only one is required (Gerard & Scarborough, 1989), they consistently but unconsciously activate both in parallel (e.g., Kroll, Bobb, & Wodniecka, 2006; Thierry & Wu, 2007). Parallel activation is demonstrated in studies showing that the processing of cross-language neighbors – words that are similar in form, whether ‘cognates’, sharing both form and meaning (e.g., Italian *genio*, English *genius*), or ‘false friends’, with similar form but unrelated meanings (e.g., English *educated*, Italian *educato* ‘polite’) – differs from the processing of control words free of overlapping form (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000). The difference is usually evidenced in either facilitation of bilingual performance (with cognates, which benefit from the activation of the target word from both sources) or hindrance to it, when the words share form only. Studies of this kind provide strong evidence of non-selective language activation: Cross-linguistically shared word forms would not affect processing if bilinguals accessed their languages independently (Kroll, Gullifer, & Rossi, 2013).

The first studies involving cross-linguistic overlap in word form with or without shared meaning mainly used visual stimuli (Spivey & Marian, 1999). More recently, however, research using auditory processing has also strongly confirmed non-selective access through phonological priming effects (Lagrou, Hartsuiker, & Duyck, 2011), whether the participants are immersed in a first (L1; Lagrou et al., 2011) or a second language setting (L2; Marian & Spivey, 2003). Furthermore, experimental studies of early word learning have begun to ask whether bilingual children’s representational system is best conceptualized as separated by language or integrated, as in adults (DeAnda, Poulin-Dubois, Zesiger, & Friend, 2016). In a study of German-English-learning toddlers, for example, von Holzen and Mani (2012) found not only L2 phonological priming of the L1, but phono-lexical priming, with the prime mediated by translation (e.g., English *slide* primes German *kleid* ‘dress’). However, it remains unclear how bilingual development is affected by exposure to two languages from the start (‘simultaneous bilingualism’) as compared with L2 acquisition in pre-school (‘sequential bilingualism’).

Developmental studies have also asked how semantic connections are organized *within* each language. Semantic organization begins to emerge with the first words (DeAnda et al., 2016). Both within- and across-language semantic priming effects – that is, facilitation or inhibition in word processing after a semantically-related word – can be detected by the second year (Conboy & Mills, 2006). However, relative proficiency may limit these effects. In a study of 31-month-old simultaneous Mandarin-English bilinguals Singh (2014) reports within-language semantic priming in the dominant language only, suggesting that, in children at least, some minimum level of lexical knowledge is required before systematic interactions can occur. Thus we have evidence of interconnected phonological, lexical and semantic systems in bilingual children, with language dominance or proficiency modulating their interaction.

Up until now few studies have addressed these issues beyond the pre-school period, and those focusing on cross-language interference have mostly been based on reading (e.g. Brenders, van Hell, & Dijkstra, 2011; Sauval, Perre, Duncan, Marinus, & Casalis, 2017; Schröter & Schroeder, 2016). An exception is Poarch and van Hell (2012), who found clear evidence of bidirectional effects of cognates on *word production* in school-aged children. Thus, a study of the accuracy and speed with which bilingual children *process* false friends and semantically related words can shed additional light on the nature of bilingual lexical representation and its development and change with increasing L2 exposure over the early school years.

To our knowledge no study has yet addressed the presumed shift from relatively low-proficiency L2 learner ('unbalanced' bilingual, Grade I) to more fully proficient 'early (more balanced) bilingual' (Grade III) over the early school years, or compared such advances to those of balanced bilinguals. Furthermore, differences in L2 proficiency may affect lexical access and organization in the child's mental architecture; also, language interaction may be modulated by such factors as word-form overlap (Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011).

Proficiency. Advances in proficiency lead to changes in processing. Only speakers with higher proficiency have been found to make use of a common neural network for both their languages (Abutalebi, Cappa, & Perani, 2001). But studies differ as to whether cross-linguistic interference occurs only in the weaker language (Kupisch, 2012) or also in the L1 (e.g., Poarch & van Hell, 2012). The question of relative L2 proficiency arises here as well (Kroll, 2008): Adult studies provide evidence of L1 effects in both high- (Spivey & Marian, 1999) and low-proficiency L2 learners (Lagrou et al., 2011). Furthermore, proficiency has been shown to affect whether or not toddlers experience within-language semantic priming (Singh, 2014), but the issue has not been raised with school-aged bilinguals, who are consolidating their lexicon in both languages in parallel, along with other cognitive and developmental changes.

Lexical similarity between the two languages. Cross-language activation to the level of phonology is reported even when the bilinguals' languages are quite distinct (Kroll et al., 2006). However, it is unknown whether more closely related languages affect one another more than less similar languages (Costa, La Heij, & Navarrete, 2006). Theoretical support for this possibility comes from a recent computational model, BLINCS (Bilingual Language Interactive Network for Comprehension of Speech: Shook & Marian, 2013), designed to reflect the dynamics and deep interactivity of bilingual language processing.

BLINCS assumes that phonotactic differences – or differences in the sound sequences permitted in each language – support a distinction between bilinguals' linguistic representations, but also that conceptual or semantic and phonological representations are shared, together with phono-lexical or word-form representations. The self-organization modeled in BLINCS thus results in a mapping with separate lexical clusters for each language but with both cognates and false friends clustering together. This suggests that bilinguals might more readily distinguish languages with fewer false friends: 'Languages that have highly distinct, non-overlapping phonological inventories might show more separation at the phonological level' (p. 320); differences at the phono-lexical level would have the same effect.

Thus despite extensive study of adult bilingualism, many developmental questions remain: (i) How do differing levels of proficiency and 'balance' between a child's languages affect the extent to which she experiences cross-linguistic interaction and semantic priming? (ii) How do the nature and extent of these effects change with greater exposure and use? (iii) Does the phono-lexical similarity of the languages affect cross-linguistic interference?

The present study

We investigate how languages are accessed and organized and affect one another in bilingual children, with particular reference to language proficiency, experience (language exposure and use), and cross-linguistic lexical similarity. In Study 1 we tested two groups of school-age children with the same L1 (Italian) but different levels of exposure to and proficiency in what we will refer to as their L2, German. Participants in the first group were simultaneous bilingual learners of German and Italian with on-going exposure to both languages, while

those in the second group acquired their L2 after their L1 and experienced less everyday exposure to it.

In Study 2, to assess the role of between-language lexical similarity, we compared the group of German L2 learners to sequential bilinguals with French, a Romance language lexically much closer to Italian than is German, as their L2. Participants in both studies were living in a bilingual area in which both Italian and German or French are official languages and in which bilingualism is wide-spread and well supported. Most children living in these areas are first exposed to the L2 in preschool; all children study the L2 and/or other subjects in that language in school. In these areas – and unlike the rest of Italy – German or French have precedence over English, which is learned only as the L3.

In each study half the children were in first grade, half in third grade at time of testing; these sequential bilinguals were thus bracketing the shift from L2 learner (in Grade I) to ‘early bilingual’ (in Grade III). Relative L2 proficiency will thus differ across groups, making it possible to address the complex dynamics of bilingualism over the course of development in the early school years.

We compared the children’s responses to word forms presented auditorily in combination with pictured meanings in four conditions: (i) Match and (ii) Mismatch between image and heard word, (iii) False-friends and (iv) Semantic-relation (mismatched but semantically related words).

The study was designed to test the following predictions:

- 1) Language balance or dominance will affect performance (that is, both accuracy and speed of response): Balanced bilingual children will represent their two languages equally well, so performance will not differ by language, whereas unbalanced bilingual children will perform better in their stronger language, Italian.
- 2) If, like adults, bilingual children activate their languages simultaneously even when in ‘monolingual mode’ (Grosjean, 1998), the presence of two languages will affect their performance in both, resulting in more errors and slower responses in False-friends than in either control condition, Match or Mismatch. However, we ask
 - a) is only the L2 affected (Singh, 2014) or are effects bidirectional (Marian & Spivey, 2003)?
 - b) does relative L2 proficiency modulate the effect (Poarch & van Hell, 2012)?
- 3) Mismatches will be difficult to process, resulting in interference, but especially when the paired words are semantically related (Rosinski, 1977). Moreover, since higher proficiency leads to stronger semantic connections (DeAnda et al., 2016), the balanced bilingual children will experience less interference than the unbalanced group in L2 Semantic-relation.
- 4) Balanced and unbalanced bilingual children can be expected to differ in their language development. Accordingly, only the unbalanced bilingual children will perform better in the L2 in Grade III than in Grade I, consolidating their L2 lexical knowledge with the added years of experience – an advance unnecessary for the balanced bilingual children in our task.
- 5) If a bilingual’s languages are organized as proposed in BLINCS, for example, more lexically similar languages should be more difficult to keep apart – a complement to the finding that greater phonological overlap facilitates early word learning (in 18-month-olds learning Spanish and Catalan: Bosch & Ramon-Casas, 2014; in 2-year-olds learning English and another language: Floccia et al., 2018). Accordingly, we expect children learning L2 French alongside L1 Italian to be less accurate and slower in responding to ambiguous words than children learning L2 German.

Study 1

Two groups of child bilinguals were tested on word recognition for evidence of cross-linguistic phonological and within-language semantic interference and how these are affected by L2 proficiency and experience.

Methods

Participants. Seventy-seven children living in the bilingual area of South Tyrol, in northeast Italy, participated. Thirty-six are native speakers of Italian and German who use both languages daily at school and with family and peers (balanced bilingual children): 18 in first grade (Grade I; $M_{\text{age}} = 6.8$ years), 18 in third grade (Grade III; $M_{\text{age}} = 8.6$). The unbalanced bilingual children were 41 Italian L2 German learners living in the same area: 21 in first grade ($M_{\text{age}} = 6.7$), 20 in third grade ($M_{\text{age}} = 8.8$). The groups did not differ significantly in child age (within grade) or parental education (see Table 1 for participants' characteristics). All participants had normal or corrected-to-normal vision and normal hearing and were free of learning difficulties or disabilities.

Materials. Children were grouped based on age of first L2 acquisition (before or after age 3: McLaughlin, 1984) and on whether or not they had had continuous exposure to that language and used both languages daily, based on the Utrecht Bilingual Language Exposure Calculator (UBiLEC, Unsworth, 2013) administered to the parents, and the children's expressive and receptive abilities, based on the questionnaire and the Peabody Picture Vocabulary Test-Revised (PPVT-R, Dunn & Dunn, 1981; Stella, Pizzoli, & Tressoldi, 2000), administered in both languages (Table 1).

Table 1
Balanced and unbalanced Italian-German bilingual groups.

	Balanced ^a (N = 36)		Unbalanced ^{a, b} (N = 41)		Balanced vs. Unbalanced					
	Grade I (N = 18)	Grade III (N = 18)	Grade I (N = 21)	Grade III (N = 20)	<i>t</i> (75)	<i>p</i>				
Sex (male:female)	8:10		8:13							
Age (years)	6.8 (0.3)	8.6 (0.4)	6.7 (0.4)	8.8 (0.3)	0.32	.747				
Age at first L2 exposure (years)	1.7 (1.3)	1.4 (1.3)	4.3 (1.7)	3.7 (1.6)	7.22	<.001				
Parental education										
Mothers	13.83	2.55	13.67	2.66	13.19	1.83	14	2.25	0.31	.756
Fathers	13.39	2.72	12.33	2.74	13.67	2.50	12.70	2.58	0.55	.582
<i>Italian</i>										
Q ^c : Comprehension	4.00 (0.91)	4.25 (0.68)	4.67 (0.48)	4.35 (0.67)	2.74	.008				
Q: Production	3.67 (0.91)	4.00 (0.63)	4.19 (0.81)	4.20 (0.52)	2	.050				
PPVT	89.28 (17.96)	102.31 (7.20)	97.86 (9.88)	106.65 (9.82)	1.93	.057				

	Balanced ^a (N = 36)		Unbalanced ^{a, b} (N = 41)		Balanced vs. Unbalanced					
<i>German</i>										
Q:										
Comprehension	3.94	(0.73)	4.06	(0.57)	1.52	(0.68)	2.45	(0.89)	-10.88	<.001
Q:										
Production	3.39	(1.29)	3.63	(0.72)	1.52	(0.60)	2.30	(0.86)	-7.35	<.001
PPVT	92.72	(16.66)	106.31	(6.45)	47.00	(17.98)	70.35	(34.20)	-7.68	<.001
<i>Italian vs. German</i>										
	Balanced		Unbalanced							
Q:										
Comprehension	$t(33) = 0.85$	$p = .40$	$t(40) = 13.52$	$p < .001$						
Q:										
Production	$t(33) = 1.68$	$p = .10$	$t(40) = 14.67$	$p < .001$						
PPVT	$t(35) = 1.32$	$p = .19$	$t(40) = 9.83$	$p < .001$						

Note. Q = questionnaire; PPVT = Peabody Picture Vocabulary Test. (Standard deviations from the mean).

^a All of the children are learning to read and write in both languages at primary school, are learning English as L3 and are reported as using each language about half of the time.

^b These children were generally first exposed to L2 at preschool; they prefer to use Italian.

^c Scale from 0 (no skills) to 5 (excellent skills).

Experiment. Children's accuracy and speed in making lexical judgments were tested in a computer-based experiment. Each child participated in two sessions, one in Italian, one in German; 60 picture-word pairs were presented per language session (see Appendix), in one of four conditions:

1. *Match*: image and auditory word refer to same object, action, or quality, e.g., image and word, *sedia* 'chair';
2. *Mismatch*: image and word are unrelated, e.g., image, *ballare* 'to dance', word, *dormire* 'to sleep';
3. *False-friends*: image and word are mismatched, but the word is phonologically similar to the relevant word in the other language, e.g., image, *suora* 'nun', word, *nonna* 'grandmother' (cf. German *Nonne* 'nun');
4. *Semantic-relation*: image and word are mismatched but semantically related, e.g., image, *bere* 'to drink', word, *mangiare* 'to eat'.

There were equal numbers of nouns, verbs, and adjectives. Stimuli were recorded using Praat (Boersma & Weenink, 2017)¹. Words presented auditorily had similar length and frequency across languages and conditions in both Study 1 and Study 2 and were of the same grammatical class as the correct response word. Of all the words used as stimuli in each language experiment, 88% and 90%, for Italian and German respectively, had a similar

¹ The researcher speaks standard Italian, without regional accent. She began learning French and German at age 14. For the L2 experiments, stimuli were recorded multiple times; the most native-sounding tokens were selected.

accentual pattern². The familiarity, concreteness, visual complexity, and goodness of depiction of the pictures used in the two studies did not differ significantly between conditions or language (see Table A1 of the Appendix for stimuli characteristics). Words were considered phonologically similar cross-linguistically if they shared at least three phonemes (or sound segments).

All words presented in the L2 False-friends condition of both studies were also presented to ten Italian adults lacking knowledge of German or French ($M_{\text{age}} = 28$ years), to ensure that these critical words are perceived as phonologically similar to the Italian word for the pictured referent. Eighteen German and French cognates were included as fillers. After each foreign word was read aloud, the participant said the Italian word sounding most like it. The adults chose the expected word in 84% of the critical cases; each false friend was chosen significantly more often than not ($\chi^2_{(1)} = 85.92, p < .001$).

All Italian stimuli were further tested with 46 Italian monolingual children, 25 in first grade (12 males, $M_{\text{age}} = 7.0$), 21 in third grade (11 males, $M_{\text{age}} = 8.8$). No significant difference was expected between False-friends and the control conditions here, as participants knew no French or German. A Mixed Model ANOVA on percent correct responses (GLMM Poisson familiarity with log-link function), with Grade and Condition as fixed effects and Subject as random effect, showed, as expected, no significant differences between control conditions and False-friends.

Procedure. Each child carried out all tests in separate 30-minute Italian and German sessions on different mornings during school hours. Half the children were tested in Italian first, the rest in German first. Both sessions were preceded by a brief conversation in the test language, to increase language activation and provide time to adjust to the researcher's accent. Instructions were given in the language of the session; the PPVT was always administered before the experiment. In the experiment each child was told they would see a picture and hear a word. They were asked to indicate as quickly and accurately as possible whether they thought the word described the picture by pressing a key labeled with a smiley (yes) or a sad face (no).

Stimuli were presented in random order using stand-alone software (programming language: Actionscript 3 for Adobe AIR runtime environments) on an Acer Aspire 5750G laptop. All pictures had the same size and resolution and appeared at the center of a 15.6in 1,366 × 768 screen. Each picture was preceded by an attention-getter, which stayed on until the child pressed the spacebar to indicate s/he was ready and which was followed, after 500 milliseconds (ms), by the paired word. All words were played at the same intensity level, using surround-sound audio technology. After the child pressed a key, the attention-getter appeared again. The next picture followed only after the child pressed the spacebar, giving participants time to rest between stimuli if necessary. Accuracy scores and response times (RTs, recorded to nearest ms from onset of spoken word) were later analyzed and compared.

Data cleaning and analysis

Three Italian picture-word pairs, one German pair, and two French pairs in Mismatch and False-friends, were excluded as being possibly semantically related or otherwise problematic³.

² 'Similar' accentual word patterns differ by no more than one syllable in length and in stress placement.

³ We administered (i) a semantic association test to 20 adults. After each picture name was read aloud participants were to say the first three words of the same category that came to

The remaining items were analyzed using the Linear (for RTs) and Generalized Linear (for accuracy) Mixed-Effects Models (LMMs/GLMMs) in R (lmer/glmer functions in package lme4; Bates, Mächler, Bolker & Walker, 2015; R Development Core Team, 2016). On lmer/glmer outputs we performed Mixed Model ANOVA Tables via likelihood ratio tests (afex package; Barr, 2013; Barr, Levy, Scheepers & Tily, 2013; Bates, Kliegl, Vasishth & Baayen, 2015). In the case of significant effects or interactions ($p < .05$) we conducted post-hoc tests using the Bonferroni correction.

Results

Figure 1 reports the results for accuracy and RTs by language (see Table A2 for more detail). From this figure it is evident that in Grade III both groups of children were close to ceiling in accuracy in Match and Mismatch in Italian. This provides a baseline for evaluating the effects of both proficiency (more/less balanced) and condition (control conditions, Match and Mismatch, vs. experimental conditions, False-friends and Semantic-relation).

Figure 1. Accuracy and response times, by condition, in Italian and German for our Italian-German balanced and unbalanced bilingual children, Grades I and III.

The results were analyzed using two Mixed Model ANOVAs, one on percentage of correct responses (GLMM Poisson family with log link-function), the other on correct RTs (LMM Gaussian family with identity link-function), both with Group (Balanced, Unbalanced), Grade (I, III), Language (Italian, German), and Condition (1, 2, 3, and 4) as fixed effects, and Subject as random effect. For the sake of clarity we present the results as responses to questions corresponding to our predictions.

1) *Do balanced bilingual children show equally strong word-concept links in the two languages, whereas unbalanced bilingual children respond better in Italian?*

Analysis showed significant interactions between Group, Language, and Grade ($\chi^2_{(1)} = 5.26, p = .022$), and Group, Language, and Condition ($\chi^2_{(3)} = 13.56, p = .004$) in accuracy, and a significant Group \times Grade \times Language \times Condition interaction ($\chi^2_{(3)} = 9.64, p = .022$) in RTs. Pairwise comparisons revealed no significant difference in Italian vs. German for the balanced bilingual children in either accuracy (even considering condition or grade separately) or RTs (even within condition and grade; see Tables 2 for accuracy and 3 for RTs).

For the unbalanced bilingual children, in contrast, pairwise comparisons showed significantly higher accuracy in Italian than in German within grade or within condition, in Match and False-friends (Table 2). However, the unbalanced bilingual children's response times in Italian and German were similar within condition and grade (Table 3).

Table 2

mind. We excluded from subsequent analysis pairs whose auditory-stimulus word was among the first three words cited by 10 or more participants. (ii) We estimated the semantic distance between image and auditory word, using 'Word-Embeddings Italian Semantic Spaces' (WEISSs; Marelli, 2017), 1 and 2. Items with a score more than 2.5 SD below the mean were removed. Finally, we excluded item pairs on which at least half the monolingual children had an accuracy score of 0, indicating that these stimuli were too difficult (given that those participants knew neither French nor German).

Accuracy: Pairwise comparisons for the significant main effects and interactions. Only the pairwise comparisons discussed in text are reported

Main effect or interaction	Group	Grade	Language	Condition	z	p	Effect-size (Odds-ratio)
Grade		I vs. III			-4.93	<.001	0.85
Condition				Match vs. False-friends	8.65	<.001	1.51
				Mismatch vs. False-friends	9.31	<.001	1.57
				Match vs. Mismatch	-0.82	1.000	0.96
				Mismatch vs. Semantic-relation	1.41	.956	1.06
Group × Language	Balanced		Italian vs.		1.11	1.000	1.05
	Unbalanced		German		9.35	<.001	1.54
Language × Condition			German	Match vs. False-friends	9.55	<.001	1.98
				Mismatch vs. False-friends	10.10	<.001	2.07
			Italian	Match vs. False-friends	2.28	.630	1.16
				Mismatch vs. False-friends	2.67	.212	1.19
Group × Language × Condition	Balanced		German	Match vs. False-friends	4.94	<.001	1.59
				Mismatch vs. False-friends	4.72	<.001	1.57
			Italian	Match vs. False-friends	1.99	1.000	1.20
				Mismatch vs. False-friends	2.45	1.000	1.26
			Italian vs. German	Match	-0.05	1.000	0.99
		Mismatch		0.59	1.000	1.05	
		False-friends		2.74	.729	1.31	
		Semantic-relation	-1.25	1.000	0.89		
	Unbalanced		German	Match vs. False-friends	8.36	<.001	2.46
				Mismatch vs. False-friends	9.31	<.001	2.73
		Italian	Match vs. False-friends	1.20	1.000	1.11	
			Mismatch vs. False-friends	1.28	1.000	1.12	
		Italian vs. German	Match	3.56	.045	1.36	
	Mismatch		2.41	1.000	1.23		
	False-friends		10.19	<.001	3.01		
	Semantic-relation	1.28	1.000	1.12			
Unbalanced vs. balanced		Italian	False-friends	1.12	1.000	1.11	
			Semantic-relation	0.46	1.000	1.04	
		German	False-friends	-6.41	<.001	0.48	
	Semantic-relation		-2.08	1.000	0.83		
Group × Grade × Language	Balanced	I	Italian vs. German		0.86	1.000	1.06
		III	Italian vs. German		0.72	1.000	1.05
		I vs. III	Italian		-0.91	1.000	-1.50
		I vs. III	German		-1.67	1.000	0.90
	Unbalanced	I	Italian vs. German		8.78	<.001	1.81
		III	Italian vs. German		4.30	<.001	1.31
		I vs. III	Italian		-0.95	1.000	0.94
		I vs. III	German		-5.45	<.001	0.68

Note. Post-hoc tests were conducted using the Bonferroni correction. For each comparison we report log rate-ratios for count dependent variables (z), p-values (with level of significance set at $p < .05$), and odds-ratios as measure of effect size.

Table 3

Response times: Pairwise comparisons for significant main effects and interactions. Only the pairwise comparisons discussed in text are reported

Main effect or interaction	Group	Grade	Language	Condition	<i>t</i>	df	<i>p</i>	Effect-size (Cohen's <i>d</i>)		
Grade		I vs. III			3.13	72.99	.002	0.73		
Condition				Match vs. Mismatch	-4.65	509.00	<.001	0.41		
				Match vs. False-friends	-13.04	509.09	<.001	1.16		
				Mismatch vs. False-friends	-8.40	509.09	<.001	0.74		
				Match vs. Semantic-relation	-8.90	509.09	<.001	0.79		
				Mismatch vs. Semantic-relation	-4.25	509.09	<.001	0.38		
Group × Condition	Unbalanced			Mismatch vs. Semantic-relation	-2.51	509.00	.342	0.22		
	Balanced			Mismatch vs. Semantic-relation	-3.47	509.00	.016	0.31		
Group × Grade × Language × Condition	Unbalanced		I	Italian	Match vs. False-friends	-5.40	509.00	<.001	0.48	
					Mismatch vs. False-friends	-2.90	509.00	1.000	0.26	
					Match vs. Semantic-relation	-4.19	509.00	.017	0.37	
				German	Mismatch vs. Semantic-relation	-1.68	509.00	1.000	0.15	
					Match vs. False-friends	-0.18	509.71	1.000	0.02	
					Mismatch vs. False-friends	0.10	509.71	1.000	0.01	
			Italian vs. German	Italian vs. German	Mismatch vs. Semantic-relation	-1.26	509.00	1.000	0.11	
					Match	-1.87	509.00	1.000	0.17	
					Mismatch	0.34	509.00	1.000	0.03	
				Italian vs. German	False-friends	3.25	509.71	.617	0.29	
					Semantic-relation	0.77	509.00	1.000	0.07	
					Match vs. False-friends	-4.12	509.00	.022	0.36	
			III	Italian	Italian	Mismatch vs. False-friends	-3.22	509.00	.675	0.29
						Mismatch vs. Semantic-relation	-2.41	509.00	1.000	0.21
						Match vs. False-friends	-5.24	509.00	<.001	0.46
				German	German	Mismatch vs. False-friends	-3.29	509.00	.532	0.29
						Mismatch vs. Semantic-relation	0.31	509.00	1.000	0.03
						Match	-1.79	509.00	1.000	0.16
I vs. III			Italian	Italian vs. German	-2.84	509.00	1.000	0.25		
				False-friends	-2.91	509.00	1.000	0.26		
				Semantic-relation	-0.12	509.00	1.000	0.01		
			German	German	Match	1.74	181.10	1.000	0.26	
					Mismatch	2.77	181.10	1.000	0.41	
					False-friends	2.50	181.10	1.000	0.37	
			Semantic-relation	2.25	181.10	1.000	0.33			
			German	Match	1.76	181.10	1.000	0.26		

		Mismatch	0.64	181.10	1.000	0.10	
		False-friends	-1.61	188.87	1.000	0.23	
		Semantic-relation	1.67	181.10	1.000	0.25	
		Match vs. False-friends	-4.69	509.00	.002	0.41	
	Italian	Mismatch vs. False-friends	-2.96	509.00	1.000	0.26	
		Mismatch vs. Semantic-relation	-2.61	509.00	1.000	0.23	
		Match vs. False-friends	-6.00	509.00	<.001	0.53	
	I	German	Mismatch vs. False-friends	-3.82	509.00	.075	0.34
			Mismatch vs. Semantic-relation	-1.37	509.00	1.000	0.12
			Match	0.32	509.00	1.000	0.03
	Italian vs. German	Mismatch	-0.15	509.00	1.000	0.01	
		False-friends	-1.00	509.00	1.000	0.09	
		Semantic-relation	1.09	509.00	1.000	0.10	
		Match vs. False-friends	-5.69	509.00	<.001	0.50	
	Balanced	Italian	Mismatch vs. False-friends	-4.22	509.00	.014	0.37
			Mismatch vs. Semantic-relation	-1.83	509.00	1.000	0.16
			Match vs. False-friends	-5.69	509.00	<.001	0.50
	III	German	Mismatch vs. False-friends	-3.35	509.00	.435	0.30
			Mismatch vs. Semantic-relation	-1.13	509.00	1.000	0.10
			Match	0.29	509.00	1.000	0.03
	Italian vs. German	Mismatch	-0.30	509.00	1.000	0.03	
		False-friends	0.57	509.00	1.000	0.05	
		Semantic-relation	0.39	181.10	1.000	0.06	
		Match	1.91	181.10	1.000	0.28	
	I vs. III	Italian	Mismatch	2.07	181.10	1.000	0.31
			False-friends	1.24	181.10	1.000	0.18
			Semantic-relation	2.58	181.10	1.000	0.38
			Match	1.89	181.10	1.000	0.28
	German	Mismatch	1.97	181.10	1.000	0.29	
		False-friends	2.28	181.10	1.000	0.34	
		Semantic-relation	2.12	181.10	1.000	0.31	
		Match	0.78	181.10	1.000	0.12	
	I	German	Mismatch	-0.53	181.10	1.000	0.08
			False-friends	-3.17	188.45	.873	0.46
			Semantic-relation	-0.67	181.10	1.000	0.10
			Match	1.02	181.10	1.000	0.15
	Unbalanced vs. balanced	III	Mismatch	0.88	181.10	1.000	0.13
			False-friends	0.72	181.10	1.000	0.11
			Semantic-relation	-0.09	181.10	1.000	0.01

Note. Post-hoc tests were conducted using the Bonferroni correction. For each comparison we report mean differences for continuous dependent variables (t), fractional degrees of freedom (df) obtained using Satterthwaite's Degrees of Freedom Approximation (Kuznetsova, Brockhoff, & Christensen, 2016; Satterthwaite, 1946), p-values (with level of significance set at $p < .05$), and Cohen's d index (d) as standardized measure of effect size.

2) *Do the children show interference in responding to False-friends?*

Analysis showed a significant main effect of Condition in both accuracy ($\chi^2_{(3)}= 80.74, p < .001$) and RTs ($\chi^2_{(3)}= 185.62, p < .001$). Pairwise comparisons showed a significant difference between False-friends and control conditions in both accuracy and RTs and a significant difference between Match and Mismatch in RTs only (Tables 2 and 3).

2a) *Does only the L1 affect L2 or are the effects bidirectional?*

Tests showed a significant interaction between Language and Condition in accuracy only ($\chi^2_{(3)}= 51.97, p < .001$). Pairwise comparisons indicated a significant difference between False-friends and the control conditions (both Match and Mismatch) in L2 German only (Table 2).

2b) *Does L2 proficiency affect interference in both languages?*

Analysis showed significant interactions between Group, Language, and Condition in accuracy, and between all factors in RTs. Pairwise comparisons showed significantly lower accuracy in False-friends as compared to Match and Mismatch for both balanced and unbalanced bilingual children in German only (Table 2). Significantly slower responses to False-friends than to Match were also found, but these differed by grade for the two groups: in *first grade*, in Italian only, for the unbalanced bilingual children, and in both German and Italian for the balanced group; in *third grade*, in both languages for both groups (Table 3). Significantly slower responses to False-friends as compared to Mismatch were found for balanced bilinguals only, and in third grade only, in Italian.

Pairwise comparisons also showed that, in accuracy, the balanced and unbalanced groups performed similarly in False-friends in Italian, whereas in German the unbalanced bilingual children were significantly less accurate than the balanced children (Table 2). However, across groups within each grade similar RTs were found in both languages in all conditions (Table 3).

3a) *Do participants perform worse in Semantic-relation than in Mismatch?*

We found a main effect of Condition in both accuracy and RTs, a significant Language \times Condition interaction in accuracy, and a significant Group \times Condition interaction in RTs ($\chi^2_{(3)}= 8.47, p = .037$). Pairwise comparisons showed similar accuracy and RTs in Mismatch as compared to Semantic-relation (Tables 2 and 3). Interestingly, RTs were significantly different between these conditions in the balanced bilingual children only (Table 3), but not within grade or language.

3b) *Do the balanced bilingual children perform better than the unbalanced bilingual children in Semantic-relation in the L2?*

We found significant interactions between Group, Language, and Condition in accuracy and between all factors in RTs. Pairwise comparisons indicated no significant differences between groups for either accuracy or RTs, or within grade for RTs (Tables 2 and 3). Moreover, pairwise comparisons revealed no significant differences in accuracy in Semantic-relation between L1 and L2 for either group (Table 2).

4a) *Does grade level affect word recognition?*

We found a main effect of Grade in both accuracy ($\chi^2_{(1)}= 19.93, p < .001$) and RTs ($\chi^2_{(1)}= 9.75, p = .002$) and a significant Group \times Grade \times Language \times Condition interaction in RTs only. Pairwise comparisons showed significantly lower accuracy and slower RTs in Grade I, but no significant differences in RTs between grades in any one condition within language or group (Tables 2 and 3).

4b) *Is the difference between performance in the L2 in Grade I vs. Grade III significant only for the unbalanced bilingual children?*

In accuracy, tests showed a significant Group \times Language \times Grade interaction and pairwise comparisons confirmed significant differences in the L2 between grades for the unbalanced bilinguals only (Table 2). All other comparisons were non-significant. In RTs there was a significant Group \times Grade \times Language \times Condition interaction, but pairwise comparisons indicated no significant differences between grades for either group, within either language or condition (Table 3).

Discussion

Study 1 was designed to assess the accuracy and RTs of Italian-German bilingual children as they judged whether auditory stimuli matched visual images. Grouping was based on proficiency in German (balanced vs. unbalanced bilingual children; all were Italian native speakers) and grade (I vs. III).

Participants generally activated both their languages, although the experiments called for only one. The children experienced interference when presented with false friends, as evidenced by their less accurate and slower responses compared with control conditions. However, relative proficiency affected the children's performance.

As predicted, balanced bilingual children showed similar accuracy scores in Italian and German, even in False-friends, whereas unbalanced bilinguals produced significantly higher scores in their stronger language. This suggests that, for the balanced bilingual children, accessing words and managing the activation of the *unintended language* was equally difficult in each language. On the other hand, the relatively limited L2 experience of the unbalanced bilingual children was insufficient to successfully inhibit activation of the stronger language when hearing phonologically related words in the L2. More evidence comes from comparison of the groups in False-friends. Tests showed a significant difference in accuracy in L2 only. This suggests that, at a comparable processing speed, the groups differ only in how well they manage cross-linguistic interference from the L1 (cf. Poarch & van Hell, 2012). On the other hand, in the L1 the groups seem to be similarly able to cope with interference from the L2, regardless of L2 proficiency.

Moreover, when between-condition differences were analyzed for each group within language, both balanced and unbalanced bilingual children made significantly more errors in False-friends than in control conditions only in the L2. This suggests a subtle degree of Italian dominance even in the balanced bilingual group. Thus our task provided a more sensitive measure of language dominance as compared to questionnaire ratings or PPVT scores.

Analysis and interpretation of the RTs confirmed these findings: (i) We found significantly slower responses to False-friends as compared to Match for balanced bilinguals in both languages, in both grades, and for unbalanced bilinguals only in Italian in Grade I but in both languages in Grade III. Given that Italian was clearly dominant for the unbalanced bilingual children, these results demonstrate that differences in response speed between critical and control conditions emerge in children of this age only when they have gained sufficient proficiency. (ii) Although the difference between the control conditions was not significant, in Grade III the balanced bilingual children were significantly slower only in Italian in False-friends vs. Mismatch. This suggests that even our balanced bilinguals had some degree of Italian dominance.

Interestingly, these results also show that the ability of the unbalanced group to cope with interference improves with advances in proficiency between Grades I and III, becoming more similar across languages in Grade III. This suggests a gradual change in language

balance, reflecting the long-established finding that dominance changes with exposure and use (Mägiste, 1979).

We expected to find lower performance in Semantic-relation than in Mismatch, as we assumed that semantically related (but mismatched) words would be more difficult to reject. However, our results showed a significant difference between these conditions in the RTs of the balanced bilinguals only, and not within language or grade. In other words, it may be harder to reject mismatched pairs when the words or associated concepts are part of the same semantic network only once relative proficiency has been attained in both languages. Nevertheless, the advantage of *semantically unrelated* pairs is not greater in the dominant language, nor does it increase with more extensive experience (in Grade III).

Comparing the balanced and unbalanced bilingual children in Semantic-relation in the L2 provides more evidence that proficiency did not affect the semantic relatedness responses. We expected unbalanced bilingual children, with their sparser, less well-integrated L2 semantic network, to perform worse in German than the balanced group. But in fact the groups had similar error scores and RTs, reflecting comparable L2 semantic interference effects and, by implication, similarly strong L2 semantic organization. This seems to contradict the idea that ‘the more words a child knows, the more likely it is that the child will be able to organize [...] the words along relevant dimensions’ (DeAnda et al., 2016:664). Alternatively, our results might reflect the fact that our stimuli derived from basic vocabulary likely to be highly familiar to all the children in both languages.

Finally, we predicted that grade level would affect both speed and accuracy in each condition, and that advances between grades would also depend on proficiency. Our results showed that grade did not affect response speed within language or condition for either group. However, accuracy improved more between grades in the unbalanced than the balanced group. In fact, both groups showed stronger performance in Grade III in both languages, but only in the unbalanced bilingual children, in the L2, was accuracy significantly better in third than in first grade. These results suggest that two years of additional L2 experience provide more benefit to bilinguals with a greater need to consolidate their knowledge, or in other words, to strengthen within-L2 links, whether paradigmatic (connecting words belonging to the same word class and semantic category, like nouns referring to body parts or furniture or verbs referring to actions or experiences, say) or syntagmatic (connecting the elements of collocations or words that co-occur in sequences). This should enable the children to better resist interference from the L1. (For discussion of the memory processes involved in lexical consolidation, see James, Gaskell, Weighall, & Henderson, 2017).

On the whole, accuracy and RTs proved not to be equally sensitive to ease of lexical processing here. Accuracy captured more differences between conditions, groups, or grades, but only RTs showed significant interactions between all factors. Combining the two measures provided a comprehensive look at how language proficiency and balance affect cross-linguistic phonological and within-language semantic interference.

Study 2

In this study we tested Italian children learning French since preschool, to compare them with the similar Italian-German unbalanced bilingual children tested in Study 1. Our primary goal was to determine whether lexical similarity between L1 and L2 plays a role in the efficiency of word access, especially for False-friends, but the study allowed us to further test all our predictions with a separate group of bilingual children.

Methods

Participants. Fifty-one L2 French learners (24 in Grade I, 27 in Grade III) living in Aosta, in the Italian-French bilingual area in northwestern Italy, participated. These children did not differ significantly in age from Study 1 unbalanced bilinguals (Table 4). All participants had normal or corrected-to-normal vision and normal hearing and were free of learning difficulties or disabilities.

Table 4

Italian-French and Italian-German unbalanced bilingual children.

	Unbalanced Italian-French group ^a (N = 51)				Unbalanced Italian-French vs. Italian-German bilinguals	
	Grade I (N = 24)		Grade III (N = 27)		<i>t</i> (90)	<i>p</i>
Sex (male:female)	12:12		15:12			
Age (years)	6.2	(0.3)	8.6	(0.3)	-0.91	.365
Age at first L2 exposure (years)	3.3	(1.5)	3.7	(2.0)	1.33	.186
Parental education						
Mothers	12.75	(2.83)	12.96	(2.28)	1.47	.143
Fathers	12.88	(2.95)	13.30	(2.77)	0.17	.865
<i>Italian</i>						
Q ^b : Comprehension	4.30	(0.47)	4.48	(0.59)	0.93	.356
Q: Production	4.00	(0.46)	4.04	(0.54)	0.96	.345
PPVT	98.80	(13.98)	103.52	(8.32)	0.42	.675
<i>L2</i>						
Q: Comprehension	1.95	(1.15)	2.40	(1.95)	1.3	.196
Q: Production	1.65	(0.88)	2.24	(1.65)	0.03	.998
PPVT	52.67	(21.56)	77.96	(52.67)	1.31	.192
<i>Italian vs. L2</i>						
	<i>t</i>		<i>p</i>			
Q: Comprehension	<i>t</i> (50) = 16.35		<i>p</i> < .001			
Q: Production	<i>t</i> (50) = 16.00		<i>p</i> < .001			
PPVT	<i>t</i> (50) = 9.74		<i>p</i> < .001			

Note. Q = questionnaire; PPVT = Peabody Picture Vocabulary Test. (Standard deviations)

^a Like the Italian-German group, the Italian-French group preferred to use Italian, both inside and outside of school; they were learning to read and write in both Italian and French and L3 English.

^b Scale from 0 (no skills) to 5 (excellent skills).

Materials and procedure. Children were chosen to participate based on Study 1 criteria and were administered French adaptations of the questionnaire, tests, and experiments. Both the questionnaire ratings and the PPVT showed higher L1 than L2 scores;

the Italian-French and Italian-German unbalanced bilingual children did not differ significantly in age of L2 exposure, parental education, or in PPVT scores (Table 4).

Experiment. Study 2 replicated Study 1, as administered to the Italian-German unbalanced group, as closely as possible. Each group received 120 picture-word pairs, 60 in Italian, 60 in French. The words mismatched with the picture were chosen, controlled, recorded, and presented following the procedures of Study 1 (see Table A1). Accentual patterns were similar in 77% and 87% of the French and Italian word stimuli, respectively.

Results

Italian-French group only. The children's accuracy and RTs in both languages are shown in Figure 2, by grade and by condition (see Table A2 for more detail).

Percentage of correct responses and RTs to those responses were analyzed, after removing the items cited in Study 1, using two Mixed Model ANOVAs with Grade (I, III), Language (Italian, L2), and Condition (1, 2, 3, and 4) as fixed effects and Subject as random effect. As in Study 1, we present the results as responses to questions corresponding to our predictions.

1. *Do these unbalanced bilingual children respond better in Italian?*

Main effect or interaction	Group	Grade	Language	Condition	z	p	Effect-size (Odds-ratio)
<i>Italian-French only</i>							
Grade		I vs. III			-3.44	< .001	0.86
Language			Ita vs. L2		10.18	< .001	1.54
Condition				Match vs. Mismatch	0.34	1.000	1.02
				Match vs. False-friends	10.97	< .001	2.01
				Mismatch vs. False-friends	10.52	< .001	1.98
				Match vs. Semantic-relation	2.43	.091	1.14
				Mismatch vs. Semantic-relation	2.05	.241	1.12
Language × Condition	Italian			Match vs. False-friends	2.67	.209	1.23
				Mismatch vs. False-friends	2.56	.296	1.23
				Match vs. Semantic-relation	2.06	1.000	1.17
				Mismatch vs. Semantic-relation	1.94	1.000	1.16
	French			Match vs. False-friends	11.84	< .001	3.29
				Mismatch vs. False-friends	11.37	< .001	3.19
				Match vs. Semantic-relation	1.39	1.000	1.12
				Mismatch vs. Semantic-relation	0.98	1.000	1.08
	Italian vs. French			Match	2.56	.295	1.21
				Mismatch	2.77	.158	1.24
				False-friends	11.42	< .001	3.23
				Semantic-relation	1.84	1.000	1.16
<i>Italian-French vs. Italian-German</i>							
Language			Italian vs. L2		13.78	<.001	1.54
Grade × Language	I vs. III	I	Italian vs. L2		11.66	< .001	1.71
		III	L2		7.67	< .001	1.39
	I vs. III	Italian			-1.85	.384	0.93
		L2			-6.09	< .001	0.75

Analysis showed significantly higher performance in Italian than in the L2 in both accuracy and RTs (see Table 5 for accuracy, Table 6 for RTs).

2. Do the children show interference in responding to False-friends?

We found a main effect of Condition in RTs ($\chi^2_{(3)} = 73.04, p < .001$) and a significant Language × Condition interaction in accuracy ($\chi^2_{(1)} = 78.61, p < .001$). Pairwise comparisons showed slower RTs in False-friends as compared to Match and Mismatch, lower accuracy in False-friends as compared to the control conditions in French only and significantly higher accuracy in Italian than in the L2 in the False-friends condition only (Tables 5, 6).

Table 5

Accuracy: Pairwise comparisons for significant main effects and interactions. Only the pairwise comparisons discussed in text are reported

Note. Post-hoc tests were conducted using the Bonferroni correction. For each comparison we report log rate-ratios for count dependent variables (z), p-values (with level of significance set at $p < .05$), and odds-ratios as measure of effect size.

Table 6

Response times: Pairwise comparisons for significant main effects and interactions. Only the pairwise comparisons discussed in text are reported

Note. Post-hoc tests were conducted using the Bonferroni correction. For each comparison we report mean differences for continuous dependent variables (t), fractional degrees of freedom (df) obtained using Satterthwaite's Degrees of Freedom Approximation (Kuznetsova,

Main effect or interaction	Group	Grade	Language	Condition	t	df	p	Effect-size (Cohen's d)
<i>Italian-French only</i>								
Grade		I vs. III			3.23	48.9	.002	0.92
Language			Italian vs. L2		-5.42	341.2	<.001	0.59
Condition				Match vs. Mismatch	-3.15	339.05	.011	0.34
				Match vs. False-friends	-8.54	339.05	<.001	0.93
				Mismatch vs. False-friends	-5.39	339.05	<.001	0.58
				Match vs. Semantic-relation	-3.44	339.05	.004	0.37
				Mismatch vs. Semantic-relation	-0.30	339.05	1.000	0.03
<i>Italian-French vs. Italian-German</i>								
Group	Italian-French vs. Italian-German				3.64	87.91	<.001	0.78
Group × Language	Italian-French vs. Italian-German		Italian vs. L2		-6.57	613.25	<.001	0.53
	Italian-French vs. Italian-German		Italian		-1.08	610.30	1.000	0.09
	Italian-French vs. Italian-German		L2		1.76	129.53	.487	0.31
Group × Condition	Italian-French vs. Italian-German			Match	1.42	228.36	1.000	0.19
				Mismatch	2.78	228.36	.164	0.37
				False-friends	5.17	231.44	<.001	0.68
				Semantic-relation	1.89	228.36	1.000	0.25
Grade × Group × Condition	Italian-French vs. Italian-German	I vs. III		Match	1.33	229.43	1.000	0.18
				Mismatch	1.88	229.43	1.000	0.25
				False-friends	4.30	229.43	.003	0.57
				Semantic-relation	2.92	229.43	.461	0.39
	Italian-German vs. Italian-French	I vs. III		Match	1.61	227.50	1.000	0.21
				Mismatch	1.57	227.50	1.000	0.21
				False-friends	0.39	233.06	1.000	0.05
				Semantic-relation	1.81	227.50	1.000	0.24
	Italian-French vs. Italian-German	I		Match	0.78	227.50	1.000	0.10
				Mismatch	2.02	227.50	1.000	0.27
				False-friends	5.42	233.58	<.001	0.71
				Semantic-relation	1.75	227.50	1.000	0.23
Italian-French vs. Italian-German	III		Match	1.23	229.25	1.000	0.16	
			Mismatch	1.91	229.25	1.000	0.25	
			False-friends	1.85	229.25	1.000	0.24	
			Semantic-relation	0.91	229.25	1.000	0.12	

Brockhoff, & Christensen, 2016; Satterthwaite, 1946), p -values (with level of significance set at $p < .05$), and Cohen's d index (d) as standardized measure of effect size.

3. Do participants perform worse in Semantic-relation than in Mismatch?

Tests revealed a main effect of Condition in both accuracy ($\chi^2_{(3)} = 96.46$, $p < .001$) and RTs ($\chi^2_{(3)} = 73.04$, $p < .001$) and a significant Language × Condition interaction. We found similar accuracy but significantly slower RTs in Semantic-relation as compared to Match, similar accuracy and RTs in Mismatch and Semantic-relation, and similar accuracy but

significantly faster RTs in Match than in Mismatch (Tables 5 and 6). Moreover, accuracy in Semantic-relation was similar to that found in both control conditions, within each language, and was also similar in L1 and L2 (Table 5).

4. *Does grade level affect the children's word recognition in these tasks?*

We found higher accuracy ($\chi^2_{(1)} = 10.42, p = .001$) and faster RTs ($\chi^2_{(1)} = 10.46, p = .001$) in Grade III than in Grade I, but – unlike Study 1 – not specifically in the L2.

Italian-French compared with Italian-German unbalanced bilingual children.

Figure 2 presents the results for accuracy and RTs by language. Percentage of correct responses and RTs to those responses were analyzed in two Mixed Model ANOVAs with Group (Italian-German, Italian-French), Grade (I, III), Language (Italian, L2), and Condition (1, 2, 3, and 4) as fixed effects and Subject as random effect.

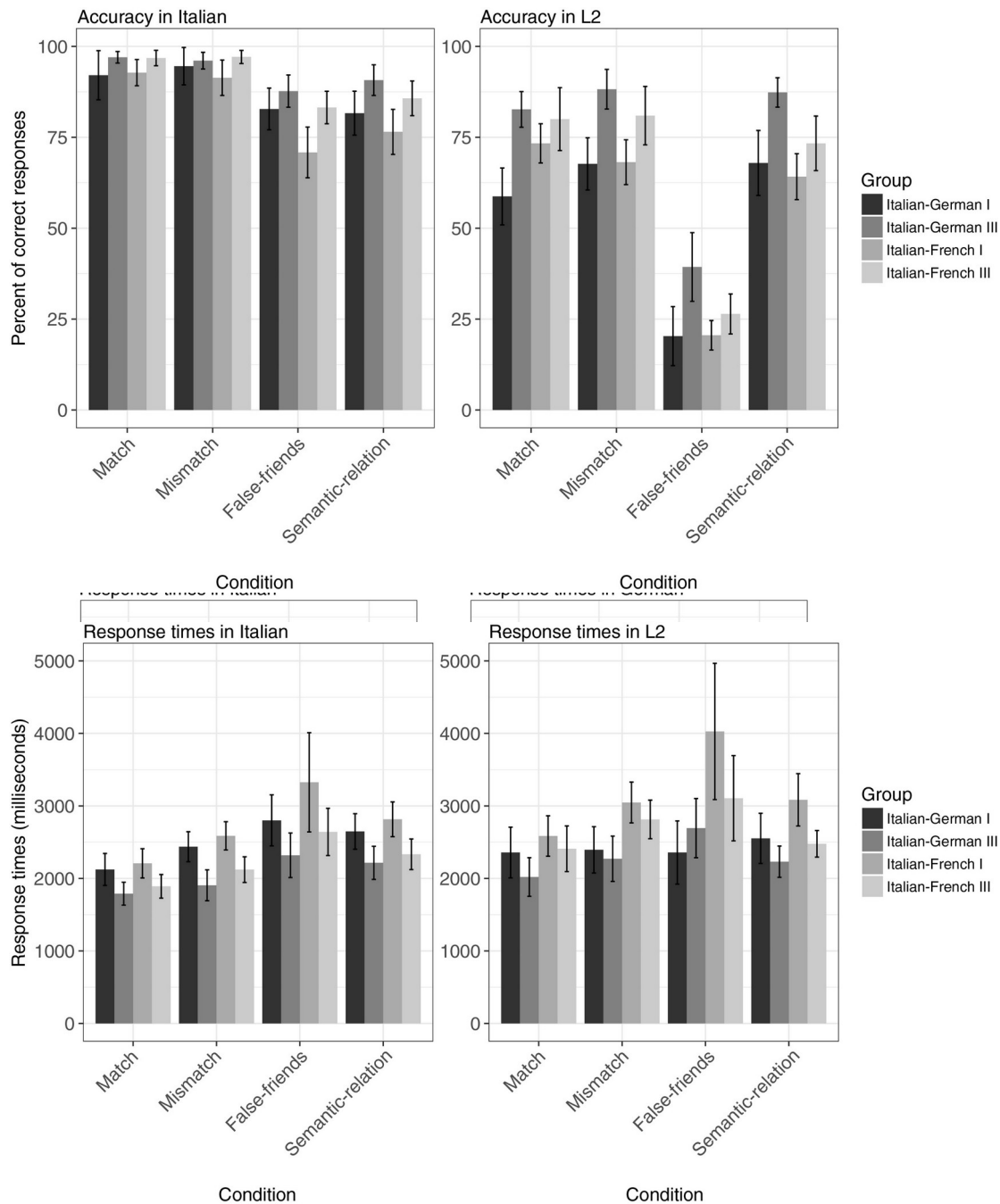
Pairwise comparisons confirmed significantly higher accuracy in the L1 than in the L2 both in general and within grades, and higher accuracy in Grade III than in Grade I in the L2 only (Table 5).

5. *Do bilingual children find it harder to keep their languages apart when the languages share more word forms?*

We found a significant Group \times Language interaction in RTs ($\chi^2_{(1)} = 12.67, p < .001$). Pairwise comparisons indicated that French and German learners differed in response times only in the L2 (more similar RTs were found in Italian), with the Italian-French group being significantly slower than the Italian-German group. Moreover, False-friends was the only condition in which the groups differed significantly: Italian-French bilinguals responded more slowly (Table 6). Italian-French children were significantly slower in French than the Italian-German children in German; only the Italian-French children responded significantly faster in Italian than in the L2 (Table 6).

Pairwise comparisons also showed a significant improvement between Grades I and III in speed of response in False-friends in the Italian-French group only; no significant differences were found in any other condition. Moreover, while the RTs to False-friends differed significantly for the two groups in Grade I, this was no longer the case in Grade III.

Figure 2. Accuracy and response times, by condition, in Italian and in the L2 for the Italian-German and Italian-French unbalanced bilingual children in Grades I and III.



Discussion

In Study 2 we assessed the performance of Italian-French unbalanced bilingual children (Grades I and III) in the experimental task used in Study 1 and compared the results with those of the Italian-German unbalanced group (Grades I and III) to assess the role of overall cross-language lexical similarity.

Our results show that, in general, the Italian-French group performed significantly better in the L1 than in the L2, in Grade III than in Grade I, and in the control conditions as compared to False-friends. These children activated both languages simultaneously, showing interference due to the presence of cross-linguistic phonologically similar words; however, extent of interference was affected by relative proficiency: Their accuracy in False-friends as compared to the control conditions was significantly more affected in their weaker language – just as was the case with the Italian-German unbalanced bilingual children in Study 1. Thus,

the effects of cross-linguistic interference do not seem to be bidirectional here, at least not before L2 proficiency reaches a certain level.

Semantic interference did not appear to be an issue: Comparison of the response times to Match and Mismatch revealed that these French learners found it more difficult to respond to mismatches than to matches, but no more so when mismatches were also semantically related. Moreover, the extent of this semantic interference did not appear to be modulated by relative proficiency, as these effects were not significant in either language.

Comparison of the Italian-French results with those of the Italian-German group with similar L2 proficiency showed that while the latter responded at a similar rate to the two languages, the Italian-French group was significantly slower in the L2 than in the L1 as well as in comparison to the Italian-German group in the L2. Moreover, a significant group difference was found in RTs in False-friends only: Knowledge of Italian and French, more closely related overall in terms of word forms, resulted in significantly slower responses than did knowledge of Italian and German. These results support our prediction: When languages share many closely related word forms, word recognition is hampered, especially in the less well-consolidated language.

There was also significantly better accuracy in Grade III as compared with Grade I, in the L2 only. However, given that we compared bilingual groups with similar proficiency, and that for Italian-French children accuracy was *not* significantly different between grades in the L2, these results must derive mainly from the Italian-German scores. Thus, these results suggest that two years of steady exposure to the L2 are sufficient for unbalanced bilingual children to make important advances in the consolidation of their lexical knowledge in that language, but the benefit to efficiency of lexical access is greater in the case of children learning two languages with relatively more distinct word forms.

General discussion

Balanced and unbalanced bilingual children were asked to judge, in each language, whether words presented auditorily matched accompanying visual images. In Study 1 we analyzed the performance of Italian-German children differing in the length and depth of their experience with L2 German, comparing responses to words cross-linguistically similar in form but not in meaning (false friends), words with related meanings (semantically related words), and controls, and testing first and third graders for potential effects of longer exposure to German. In Study 2 we tested Italian-French unbalanced bilingual children and compared the results with Study 1 findings to see how the relative cross-linguistic similarity of word forms affects performance.

We demonstrated false friends interference in perceptual processing in our school-age children. Von Holzen and Mani (2012) found subtle evidence of non-selective lexical access in younger pre-schoolers; Poarch and van Hell (2012) reported similar effects in a production study with children aged 4 to 8. One aim of this study was to investigate the development of an L2 in childhood and the shift from (Grade I) ‘L2 learner’ towards more balanced (Grade III) bilingual proficiency.

Language proficiency and experience

A key question for developmental bilingual studies is how differences in extent of experience with the less dominant language impacts on lexical access and processing. Studies of infants acquiring their first language(s) (e.g., Marchman, Fernald, & Hurtado, 2010) have shown that processing speed depends not on individual differences but on proficiency in each language. We asked to what extent proficiency mediates lexical access in school-age children.

Our findings confirmed that balanced experience with two languages results in equally strong word-concept links, expressed in the ability to access words or concepts with equal ease in both languages. Our balanced bilingual children generally performed in a similar way across languages, whereas unbalanced bilinguals performed significantly better in their dominant language. Proficiency affected the ease with which the *unwanted* language was inhibited, as reflected in the effect of false friends. While, for the balanced group, managing the activation of the unwanted language was similarly difficult across languages, unbalanced bilingual children had insufficient L2 experience to successfully cope with L1 interference. Additional evidence for the role of proficiency comes from comparing L2 performance: Stronger L2 skills and more experience in managing simultaneous language activation allowed the balanced bilinguals to better resist the lure of L2 false friends.

Note that it is not possible, based on this study alone, to disentangle the effects of age (or maturation) from increased exposure and use (or gains in lexical knowledge). However, the differences between our balanced and unbalanced bilinguals are in accord with the now widely held view that the extent of a child's baseline lexical knowledge is a critical factor in further development (see James et al., 2017).

We also found that bilinguals must gain sufficiently robust knowledge of the L2 before false friends affect their dominant language: Our unbalanced bilingual children showed no interference effect in their L1. This is in line with adult research (e.g., Costa et al., 2000) as well as with Poarch and van Hell (2012), who found that only more balanced multilingual children show cognate facilitation effects in picture naming in both L2 and (though less robustly) L1. However, in our study even the balanced bilinguals suffered more cross-linguistic phonological interference in L2 than in L1. This may mean that, despite their similar vocabulary scores, these balanced bilinguals access words more efficiently in Italian than in German. This differential strength is revealed only in tasks with more sensitive measures, such as RTs; thus, eye-tracking or ERPs, for example, would be even better suited to test these ideas.

Contrary to our expectations, rejecting picture-word pairs proved no more difficult when the associated concepts belonged to a single semantic network (Semantic-relation) than when they did not (Mismatch). Moreover, efficiency in responding to semantically related as compared to mismatched words did not tend to increase with greater language proficiency and longer periods of L2 exposure for either the Italian-German or the Italian-French unbalanced bilingual children.

Thus in this study the effect of semantic relationship was not significantly affected by vocabulary size, proficiency, or the particular language pair. This is surprising given that, as children learn new words and associate them with meanings, they build and organize their mental representations into a system thought to become ever more efficient as the vocabulary grows – but also to engender more semantic interference as a consequence (Singh, 2014). However, the lack of the expected effect might be due to the over-familiarity of the stimuli or a lack of sensitivity in the measures used.

Finally, across-grade comparison showed a general improvement in each group, but not in any one particular condition. Importantly, it showed that improvement was not the same for all children: Only unbalanced bilingual children showed significantly better performance in Grade III, perhaps because in their L2 they have more room for improvement.

Lexical similarity across the two languages

We predicted that Italian and French would be more difficult to keep apart in the case of false friends than more lexically distinct Italian and German. As in BLINCS (Shook & Marian,

2013), we see the two languages as self-organizing into densely connected within-language networks based on the features that different linguistic units have in common *within* a given language; these units are also connected *across* the languages of a bilingual through form-form as well as form-meaning links (with the meanings largely shared cross-linguistically). According to the BLINCS model cross-linguistic lexical similarity should affect lexical access: The word-form similarity of Italian and French creates mappings in which Italian and French words cluster together along a putative language-border; Italian and German have fewer such words.

The greater similarity of Italian and French – as compared to Italian and German – is evident even from a comparison of just our small subset of German and French stimuli (excluding false friends) and their Italian translation equivalents: While only 3% of Italian and German words overlap to the extent of three (sequential but not necessarily contiguous) phonemes – the criterion used here for identifying false friends – there is a 15% overlap between Italian and French. The larger pool of similar words would account for the greater difficulty shown by the Italian-French-speaking unbalanced bilinguals as compared with the Italian-German unbalanced group in the L2, both in general and in False-friends in particular.

On the other hand, performance in False-friends ‘improved’ more in the case of Italian-French children between Grades I and III, to the point that in Grade III the Italian-French and Italian-German bilingual groups no longer differed in response time. This suggests that although lexical similarity initially leads to greater difficulty in dealing with false friends, over the course of development the ability to cope improves more for those bilingual children who encounter more cross-linguistically similar words, allowing children learning two more similar languages to catch up with their peers after just two years.

The lexical similarity of the two languages also affected how much each group ‘improved’ from Grade I to III overall. In fact, although we found no interaction between Language and Group, the analysis of the performance of the two groups alone and then as compared showed greater advances between Grades I and III for the Italian-German than for the Italian-French children. This suggests that knowledge of languages more closely related in terms of number of similar word forms is slower to lead to improvement in *efficiency of lexical access* as compared to knowledge of two less similar languages. However, longitudinal studies or comparison of two languages more sharply contrasted in terms of word forms would be needed to test this proposal.

Conclusion

This study is the first to explore developmental aspects of both balanced and unbalanced bilingual exposure in the early school years. It has again shown non-selective lexical access, as expected. It has also provided insight into how the ability to inhibit the unwanted language is affected by proficiency. Balanced and unbalanced bilingual children were affected similarly in the L1, but in the L2 the unbalanced bilinguals improved more between grades; this suggests that being first exposed to a second language at some point after the first year of life irreversibly compromises neither L1 nor L2 attainment. By testing unbalanced groups learning German and French as their second language we have demonstrated the greater challenge posed by lexically similar languages, especially in the case of words cross-linguistically similar in form but not in meaning, although Italian-French bilinguals catch up with their Italian-German peers by Grade III. Further study is needed to clarify in more detail how learning trajectories differ between balanced and unbalanced bilingual children and in those learning more or less distinct language pairs.

Appendix. Supplementary material

References

- Abutalebi, J., Cappa, S., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, *4*, 179-190. <https://doi.org/10.1017/S136672890100027X>.
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Quantitative Psychology and Measurement*, *4*, 328. <https://doi.org/10.3389/fpsyg.2013.00328>.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing. *Journal of Memory and Language*, *68*, 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967, 1-27.
- Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48. <https://doi.org/10.18637/jss.v067.i01>.
- Boersma, P., & Weenink, D. (2017). Praat. [Computer program]. Version 6.0.25, retrieved 11 February 2017 from <https://www.praat.org>.
- Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary. *International Journal of Behavioral Development*, *38*, 317-322. <https://doi.org/10.1177/0165025414532559>.
- Brenders, P., van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners. *Journal of Experimental Child Psychology*, *109*, 383-396. <https://doi.org/10.1016/j.jecp.2011.03.012>.
- Conboy, B. T., & Mills, D. L. (2006). Two languages, one developing brain. *Developmental Science*, *9*, F1-F12. <https://doi.org/10.1111/j.1467-7687.2005.00453.x>.
- Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1283-1296. <https://doi.org/10.1037/0278-7393.26.5.1283>.
- Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, *9*, 137-151. <https://doi.org/10.1017/S1366728906002495>.
- DeAnda, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Lexical processing and organization in bilingual first language acquisition. *Psychological Bulletin*, *142*, 655-667. <https://doi.org/10.1037/bul0000042>.
- Dunn, L. M., & Dunn, L. M. (1981). *Examiner's manual for the Peabody Picture Vocabulary Test- Revised edition*. Circle Pines, MN: American Guidance Service.
- Floccia, C., Sambrook, T. D., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J., & Plunkett, K. (2018). III: Analyses and results for study 1: Estimating the effect of linguistic distance on vocabulary development. *Monographs of the Society For Research in Child Development*, *83*, 43-60. <https://doi.org/10.1111/mono.12350>.
- Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 305-315. <https://doi.org/10.1037/0278-7393.15.2.305>.
- Grosjean, F. (1998). Studying bilinguals. *Bilingualism: Language and Cognition*, *1*, 131-149. <https://doi.org/10.1017/S136672899800025X>.

- James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews*, 77, 1-13. <https://doi.org/10.1016/j.neubiorev.2017.01.054>.
- Kroll, J.F. (2008). Juggling two languages in one mind. *Psychological Science Agenda SCIENCE BRIEFS*. Retrieved from <https://www.apa.org/science/about/psa/2008/01/kroll.aspx> (accessed June 2018).
- Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule. *Bilingualism: Language and Cognition*, 9, 119-135. <https://doi.org/10.1017/S1366728906002483>.
- Kroll, J. F., Gullifer, J. W., & Rossi, E. (2013). The multilingual lexicon. *Annual Review of Applied Linguistics*, 33, 102-127. <https://doi.org/10.1017/S0267190513000111>.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest. R package version 2.0-32, URL <https://CRAN.R-project.org/package=lmerTest>.
- Kupisch, T. (2012). Specific and generic subjects in the Italian of German-Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition*, 15, 736-756. <https://doi.org/10.1017/S1366728911000691>.
- Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 952-965. <https://doi.org/10.1037/a0023217>.
- Mägiste, E. (1979). The competing language systems of the multilingual. *Journal of Verbal Learning and Verbal Behavior*, 18, 79-89. [https://doi.org/10.1016/S0022-5371\(79\)90584-X](https://doi.org/10.1016/S0022-5371(79)90584-X).
- Marchman, V., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, 37, 817-840. <https://doi.org/10.1017/S0305000909990055>.
- Marelli, M. (2017). Word-embeddings Italian semantic spaces. *Psihologija*, 50, 503-520. <https://doi.org/10.2298/PSI161208011M>.
- Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, 24, 173-193. <https://doi.org/10.1017/S0142716403000092>.
- McLaughlin, B. (1984). *Second-language Acquisition in Childhood*, 2nd ed.. V. 1: *Preschool children*. Hillsdale, NJ: Lawrence Erlbaum.
- Poarch, G. J., & van Hell, J. G. (2012). Cross-language activation in children's speech production. *Journal of Experimental Child Psychology*, 111, 419-438. <https://doi.org/10.1016/j.jecp.2011.09.008>.
- R Development Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. URL <https://www.R-project.org>.
- Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Development*, 48, 643-647. <https://doi.org/10.2307/1128667>.
- Satterthwaite, F. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2, 110-114. <https://doi.org/10.2307/3002019>.
- Sauval, K., Perre, L., Duncan, L. G., Marinus, E., & Casalis, S. (2017). Automatic phonological activation during visual word recognition in bilingual children. *Journal of Experimental Child Psychology*, 154, 64-77. <https://doi.org/10.1016/j.jecp.2016.10.003>.

- Schröter, P., & Schroeder, S. (2016). Orthographic processing in balanced bilingual children. *Journal of Experimental Child Psychology, 141*, 239-246. <https://doi.org/10.1016/j.jecp.2015.09.005>.
- Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for comprehension of speech. *Bilingualism: Language and Cognition, 16*, 304-324. <https://doi.org/10.1017/S1366728912000466>.
- Singh, L. (2014). One world, two languages. *Child Development, 85*, 755-766. <https://doi.org/10.1111/cdev.12133>.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages. *Psychological Science, 10*, 281-284. <https://doi.org/10.1111/1467-9280.00151>.
- Stella, G., Pizzoli, C., & Tressoldi, P. (2000). *Peabody. Test psicolinguistico*. Torino: Omega Edizioni.
- Thierry, G., & Wu, Y-J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*, 12530-12535. <https://doi.org/10.1073/pnas.0609927104>.
- Unsworth, S. (2013). Utrecht Bilingual Language Exposure Calculator (UBiLEC): Excel Spreadsheet. Unpublished material.
- Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*, 88-107. <https://doi.org/10.1016/j.jml.2010.08.006>.
- von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology, 113*, 569-586. <https://doi.org/10.1016/j.jecp.2012.08.001>.