Bilingualism: Language and Cognition 22 (2), 2019, 331–348 © Cambridge University Press 2018 doi:10.1017/S1366728918000378

Bilingual memory, to the extreme: Lexical processing in simultaneous interpreters*

MICAELA SANTILLI
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
MARTINA G. VILAS
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
EZEQUIEL MIKULAN
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
MIGUEL MARTORELL CARO
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
EDINSON MUÑOZ
Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile
LUCAS SEDEÑO
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
AGUSTÍN IBÁÑEZ
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
Universidad Autónoma del Caribe, Barranquilla, Colombia
Center for Social and Cognitive Neuroscience (CSCN), School of Psychology, Universidad Adolfo Ibáñez, Santiago de Chile, Chile
Centre of Excellence in Cognition and its Disorders, Australian Research Council (ACR), Sydney, Australia
ADOLFO M. GARCÍA
Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
Faculty of Education, National University of Cuyo (UNCuyo), Mendoza, Argentina

(Received: April 06, 2017; final revision received: February 06, 2018; accepted: February 27, 2018; first published online 16 April 2018)

This study assessed whether bilingual memory is susceptible to the extreme processing demands of professional simultaneous interpreters (PSIs). Seventeen PSIs and 17 non-interpreter bilinguals completed word production, lexical retrieval, and verbal fluency tasks. PSIs exhibited enhanced fluency in their two languages, and they were faster to translate words in both directions. However, no significant differences emerged in picture naming or word reading. This suggests that lexical enhancements in PSIs are confined to their specifically trained abilities (vocabulary search, interlingual reformulation), with no concomitant changes in other word-processing mechanisms. Importantly, these differences seem to reflect specifically linguistic effects, as both samples were matched for relevant executive skills. Moreover, only word translation performance correlated with the PSIs' years of interpreting experience. Therefore, despite their tight cooperation, different subcomponents within bilingual memory seem characterized by independent, usage-driven flexibility.

Keywords: bilingual memory, lexical processing, simultaneous interpreters, expertise

---

* This work was partially supported by grants from CONICET; CONICYT/FONDECYT Regular (1170010); Programa Interdisciplinario de Investigación Experimental en Comunicación y Cognición (PIIECC), Facultad de Humanidades, USACH; Proyecto Basal USACH USA1398_ME112214; FONDAP 15150012; and the INECO Foundation.

Micaela Santilli and Martina G. Vilas are both lead authors of this work, with equal contribution.

Address for correspondence:
Adolfo M. García, Ph. D., Institute of Cognitive and Translational Neuroscience & CONICET, Pacheco de Melo 1860, C1126AAB, Buenos Aires, Argentina
adolfomartingarcia@gmail.com

Supplementary Material can be found online at https://doi.org/10.1017/S1366728918000378

Introduction

Cognitive domains are highly flexible and adaptive, especially if subjected to continuous demanding conditions. This is clearly shown by expertise studies. For example, visual selective attention is enhanced in habitual videogame players (Green & Bavelier, 2003), whereas problem-solving and motor error detection mechanisms are boosted in chess masters (Bilalić, McLeod & Gobet, 2008) and professional tango dancers (Amoruso, Ibáñez, Fonseca, Gadea, Sedeno, Sigman, García, Fraiman & Fraiman, 2016; Amoruso, Sedeno, Huepe, Tomio, Kamienkowski, Hurtado, Cardona, Alvarez Gonzalez, Rieznik, Sigman, Manes & Ibáñez, 2014), respectively. However, little is known about the impact of sustained extreme demands on another well-characterized cognitive domain: bilingual memory – namely, the lexico-semantic system of individuals who speak both a native and a non-native language (L1 and L2, respectively). A key model to examine this issue is afforded by experts in simultaneous interpretation, a most exacting form of bilingual activity characterized by cumulative, dynamic, and concurrent processing of oral input and output under strict time constraints (Chernov, 2004). Building on these premises, we assessed the susceptibility of bilingual memory to expertise-related changes by comparing the lexical skills of professional simultaneous interpreters (PSIs) and non-interpreter bilinguals (NIBs).

Bilingual memory is cross-theoretically conceived as a complex of interfacing components (Dijkstra & van Heuven, 2002; Dong, Gui & Macwhinney, 2005; French & Jacquet, 2004; Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz & Green, 2010; Paradis, 2004; Van Hell & De Groot, 1998; van Heuven, Dijkstra & Grainger, 1998). Macroanatomically, lexical processes for both L1 and L2 in high-proficiency bilinguals are subserved by overlapping regions (Perani, Paulesu, Galles, Dupoux, Dehaene, Bettinardi, Cappa, Fazio & Mehler, 1998), irrespective of the age of L2 acquisition (Abutalebi, Cappa & Perani, 2005; Perani et al., 1998). However, different components of this overall system possess specific, partly dissociable functions. For example, although the L1 and the L2 lexico-semantic systems continuously coactivate and exchange information (for a review, see Kroll, Dussias, Bice & Perrotti, 2015), they are subserved by relatively independent neural substrates (Chee, Soon & Lee, 2003; Klein, Zatorre, Chen, Milner, Crane, Belin & Bouffard, 2006; Lucas, McKhann & Ojemann, 2004; Ojemann & Whitaker, 1978; Rapport, Tan & Whitaker, 1983). Similarly, subdomains within these systems (e.g., semantics, phonology) interact during verbal processing, but they rely on partially autonomous networks (Marian, Spivey & Hirsch, 2003; Pillai, Araque, Allison, Sethuraman, Loring, Thiruvaiyaru, Ison, Balan & Lavin 2003; Sasanuma, Sakuma & Kitano, 1992; Teichmann, Turc, Nogues, Ferrieux & Dubois, 2012). Moreover, at least some of the neural circuits engaged during interlingual processes (e.g., word translation) are different from those involved in single-language processes (e.g., word reading) (Borius, Giussani, Draper & Roux, 2012; García, 2015b).

Beyond strictly linguistic mechanisms, speech production and comprehension in L1 and L2 also rely on cognitive control processes, which are activated depending on the interactional context and the bilingual status of the interlocutor (Green & Weib, 2014). As proposed by several models, control mechanisms regulate lexical competition processes in bilingual memory. In single- or dual-language contexts, bilinguals must select a specific language schema and suppress non-target language items (Green & Abutalebi, 2013), a requisite that renders lexical processing more costly but seems to enhance executive functioning at large (Bialystok, 2009; Bialystok, Craik, Grady, Chau, Ishii, Gunji & Pantev, 2005; Bialystok, Craik, Klein & Viswanathan, 2004; Calvo, Ibáñez & García, 2016). Indeed, executive functions play a key role in the inhibition of words in the non-target language, be it L1 or L2 (Rodriguez-Fornells, De Diego Balaguer & Munte, 2006), but control demands are not always equivalent for both languages. In low- or mid-proficiency NIBs, switching from the dominant language (typically, L1) to the less dominant one (typically, L2) proves more effortful than doing so in the opposite direction (Costa & Santesteban, 2004). However, this asymmetrical switching cost is sensitive to the level and type of bilingual proficiency, as it can be neutralized in highly proficient NIBs (Costa & Santesteban, 2004) and even reversed in PSIs (Proverbio, Leoni & Zani, 2004).

Despite its integrity as a functional system, then, bilingual memory is an amalgam of specialized mechanisms that can be subjected to differential use-related demands, adaptations, and enhancements. Many of the connections holding these subsystems together are extremely taxed in PSIs, who must continually reformulate diverse forms of oral discourse from one language into another, sometimes for several hours. Typically, in professional settings, simultaneous interpreting occurs at a speed which largely exceeds the ideal rate of 95–120 words per minute (Chernov, 2004; Gerver, 1975), with a delay between input and output of 2 to 4 seconds (Anderson, 1994; Gerver, 1976). Yet, PSIs typically achieve remarkable propositional correspondence, significantly outperforming NIBs (Barik, 1975; Dillinger, 1994; Gerver, 1975). This suggests that mechanisms implicated in finding and translating task-constrained words can become enhanced due to the recurrence and intensity of their activation. Indeed, word-level processing is more efficient in bilinguals who attained high levels of L2 proficiency (Ferré, Sánchez-Casas & Guasch, 2006; Guasch, Sánchez-Casas, Ferré & García-Albea, 2008; Sunderman & Kroll, 2006; Talamas, Kroll & Dufour, 1999) or who acquired their L2 at an early age (Montrul & Foote, 2014; Silverberg & Samuel, 2004), with both languages remaining susceptible to dynamic adaptations across the lifespan (Malt, Li, Pavelnko, Zh & Ameel, 2015). By the same token, interpreting expertise could represent another critical subject variable to investigate the susceptibility of bilingual memory to fine-tuned processing skills forged throughout time.

Previous studies focused on executive functions have shown that PSIs exhibit behavioral enhancements (Babcock & Vallesi, 2015; for a review, see García, 2014) and relevant neuroplastic adaptations (Elmer & Kuhnis, 2016; Hervais-Adelman, Moser-Mercer, Michel & Golestani, 2014). More particularly, other studies indicate that, relative to NIBs, PSIs are faster to perform lexical decisions on non-words (Bajo, Padilla & Padilla, 2000), categorize non-typical exemplars (Bajo et al., 2000), and translate words in both directions (Christoffels, de Groot & Kroll, 2006). Moreover, in semantic decision tasks, they exhibit distinct neurophysiological modulations only for word pairs presented in the professionally trained direction (Elmer, Meyer & Jancke, 2010). It seems, therefore, that interpreting expertise may entail changes only in those bilingual skills specifically taxed by expert performance (i.e., vocabulary search, interlingual reformulation).

This conjecture, however, requires direct testing, as its supporting data are scant, heterogeneous, and clouded by various confounds. Crucially, only Elmer et al. (2010) matched PSIs and NIBs for L2 proficiency, age of L2 acquisition, and amount of L2 exposure, and none of the available studies did so in terms of extralinguistic skills associated with L2 processing and translation performance, such as working memory (Kroll, Michael, Tokowicz & Dufour, 2002; Linck, Osthus, Koeth & Bunting, 2014). Moreover, no single study has systematically assessed how *different* mechanisms within bilingual memory are affected by the extreme processing demands characterizing the activity of PSIs.

Against this background, we conducted the first comprehensive examination of expertise-related differences in bilingual memory by comparing carefully matched PSIs and NIBs on phonological and semantic fluency, picture-naming, word-translation, and word-reading tasks, all in both L1 and L2. Guided by extant findings, we hypothesized that only those mechanisms implicated in time-constrained vocabulary search and word translation would become more efficient in subjects who underwent systematic interpreting practice. Thus, we predicted that, relative to NIBs, PSIs would produce more words in verbal fluency tasks and exhibit lower response latencies in word translation tasks, but not in picture naming or a shallow-processing task like word reading. Evidence for such selective enhancements could shed light on the organization and adaptability of the bilingual lexico-semantic system.

Materials and methods

Participants

The study comprised 36 participants, but two of them (both PSIs) were excluded because of technical problems during data acquisition. The final sample thus consisted of 34 native Spanish speakers from Argentina (32 female) who learned English through formal instruction. They all had normal or corrected-to-normal vision and no history of neurological or psychiatric disease. Half of them comprised NIBs ($n = 17$), that is, advanced students at an English teacher training program ($n = 14$) or English teachers ($n = 3$) with no experience in interpreting. The other half comprised PSIs ($n = 17$), mostly specialized in conference interpreting, with a mean of 14.65 years of experience ($SD = 12.09$).[1] Some of the participants (9 NIBs and 14 PSIs) reported possessing varying levels of competence in an additional language (mainly French, Italian or Portuguese, but also German, Argentine Sign Language, and Flemish in three individual cases). However, in professional settings, all PSIs interpreted only between Spanish and English.

All participants completed a self-report questionnaire providing demographic and language-related information, including interpreting expertise. They first reported their gender, age, age of L2 learning, and years of L2 study. They were also asked to indicate their level of competence in both languages as well as competence in L1-L2 and L2-L1 interpreting directions in a scale ranging from 1 (null) to 7 (optimal). It was indicated that those extremes of the scale denoted complete inability to perform even basic relevant tasks and high capacity to routinely deploy those skills at ease, respectively. Participants were also asked to estimate how many hours they spent each week watching, reading, or listening to media (e.g., TV, written texts, radio) in both languages and practicing SI in each direction. As shown in Table 1, PSIs were significantly more competent in L2-L1 and L1-L2 interpreting and they spent more hours interpreting in both directions each week. Crucially, however, the two samples were matched for gender, age, and critical variables known to modulate lexical performance, namely: L1 and L2 competence, hours of weekly exposure to media in L1 and L2, age

---

[1] A subset of this group provided information on the amount of time they spent practicing interpretation as students. Six subjects reported that, for each interpreting direction, they practiced between 3 and 5 hours per week. One subject mentioned having an extra hour of practice per week in the L2-L1 direction, and the remaining one devoted 10 hours of practice a week for each interpreting direction.

Table 1. *Participants' demographic, linguistic, executive, and interpreting profile.*

	NIBs	PSIs	NIBs vs. PSIs	
	n = 17	*n* = 17	*p*-value*	Cohen's *d*[#]
DEMOGRAPHIC DATA				
Gender (F:M)	(17:0)	(15:2)	.14	—
Age (years)	35.12 (14.88)	40.35 (11.87)	.27	.39
LANGUAGE BACKGROUND				
L1 competence[a]	6.71 (.47)	6.65 (1.46)	.87	0.05
L2 competence[a]	6.24 (.75)	6.65 (.61)	.08	0.6
Weekly exposure to L1 media[b]	1.94 (1.09)	2.12 (1.11)	.64	0.16
Weekly exposure to L2 media[b]	2.41 (1.28)	2.53 (1.59)	.81	0.08
Age of L2 learning (years)	7.31 (3.7)	7.03 (3.35)	.82	0.08
Years of study of L2	18.74 (12.69)	20.59 (9.81)	.64	0.03
EXECUTIVE FUNCTIONS				
Working memory span[c]	8.06 (1.6)	7.88 (1.54)	.75	0.11
Cognitive flexibility[d]	5.08 (.95)	5 (1.79)	.89	0.05
Overall executive skills[e]	26.5 (1.56)	25.71 (2.06)	.21	0.04
INTERPRETING EXPERTISE				
Competence in L2-L1 SI[a]	4.82 (.88)	6.53 (.8)	**< .001**	2.02
Competence in L1-L2 SI[a]	4.53 (1.01)	6.35 (.79)	**< .001**	2.02
Weekly dedication to L2-L1 SI[b]	.59 (.51)	3.38 (1.82)	**< .001**	2.08
Weekly dedication to L1-L2 SI[b]	.59 (.51)	3.38 (1.82)	**< .001**	2.08

Data presented as mean (*SD*) with the exception of gender. NIBs: non-interpreter bilinguals; PSIs: professional simultaneous interpreters; L1: native language; L2: foreign language; SI: simultaneous interpreting.

[a] Data from a self-rating scale ranging from 1 (null) to 7 (optimal).

[b] Data from a self-rating scale with the following ranks: 0 = null, 1 = little (from 1 to 5 hs), 2 = considerable (from 5 to 10 hs), 3 = intense (from 10 to 15 hs), 4 = very intense (from 15 to 20 hs), 5 = extremely intense (from 20 to 25 hs), 6 = excessive (more than 25 hs).

[c] Based on the working memory index of the INECO Frontal Screening battery (Torralva et al., 2009). This executive subdomain was subjected to individual analysis given its relevance for research on PSIs interpreters (I. K. Christoffels et al., 2006; Carolina Yudes et al., 2013).

[d] Based on the Wisconsin Card Sorting Test (Nelson, 1976).

[e] Based on the global score of the INECO Frontal Screening battery (Torralva et al., 2009).

******p*-values calculated with *t*-tests for independent samples (except for gender results, which were analyzed via a chi-square test.

of L2 learning, and years of study of L2. Moreover, they were similar in terms of relevant non-linguistic skills, such as working memory span, cognitive flexibility, and overall executive functioning (the measures used to tap into these domains are described in section "Materials and procedures"). Note, in this sense, that the NIBs were not specifically selected so that they would match the PSIs in the above measures. This was an empirical pattern emerging from our initial between-group comparisons – despite contradictory findings (García, 2014), similar results have been reported by Köpke and Nespoulous (2006) and Signorelli, Haarmann and Obler (2011).

Materials and procedures

All participants began the evaluation by answering the questionnaire described above. Then they performed two of the verbal fluency tests, followed by the executive function tasks. Then the picture-naming, reading, and translation tasks were assigned in randomized order across subjects. These were all performed on a personal computer in a dimly illuminated room, while an examiner monitored the participants' performance. The picture-naming task was designed and run on e-Prime software, whereas the reading and translations tasks were developed on Python programming language (www.python.org) with the Pygame development library (www.pygame.org). The protocol finished with the remaining verbal fluency tests. Note that the protocol's counterbalancing scheme, which was separately and identically applied for each group, ensured that no particular task, language or condition was at an inherent disadvantage due to previous-task exposure or fatigue effects – for a detailed account of the sequencing of instruments/tasks, see Tables S1 through S4 in the Supplementary Material (Supplementary Material). Details on the materials and procedures of each task are offered below.

Executive functions

First, overall executive skills were assessed through the INECO Frontal Screening (IFS) battery (Torralva, Roca, Gleichgerrcht, Lopez & Manes, 2009), a sensitive tool tapping on eight relevant domains, namely: (1) motor programming: subjects perform the Luria series (“fist, edge, palm”), first by copying the administrator and then on their own; (2) conflicting instructions: subjects are required to tap the table once when the administrator taps it twice, or twice when the administrator taps it once; (3) motor inhibitory control: subjects are told to tap the table only once when the administrator taps it once, but to do nothing when the examiner taps it twice; (4) numerical working memory: subjects are asked to repeat a progressively longer string of digits in the reverse order; (5) verbal working memory: subjects are asked to list the months of the year backwards, starting with December; (6) spatial working memory: the examiner presents four cubes and points at them in a given sequence; the subject is asked to repeat the sequence in reverse order; (7) abstraction capacity: subjects are read proverbs and asked to explain their meaning; (8) verbal inhibitory control: this task, based on the Hayling test, measures the ability to inhibit an expected response; in the first part, subjects are read three sentences and asked to complete them correctly, as quickly as possible; in the second part, they are asked to complete another three sentences with a syntactically correct but semantically incongruous word. The maximum global score on the IFS is 30 points.

Second, working memory skills were calculated by reference to the corresponding index from the IFS. Specifically, performance was measured as the sum of the results from two of the above-mentioned subtests: numerical working memory and spatial working memory. For the former, the total score over a maximum of 6 is calculated as the number of digits remembered in the exact order. For the latter, the participant can obtain up to 4 points depending on the number of correctly performed sequences. Thus, the maximum score in the WM index is 10 points.

Finally, cognitive flexibility was assessed with the Wisconsin Card Sorting Test (Nelson, 1976). In this stimulus categorization task, four guide cards are presented and, in each trial, the participant must place a new card below one of them. The cards vary according to three parameters: color (red, blue, green or yellow), shape (triangle, circle, star or cross) and number of figures (one, two, three or four). In each trial, the examiner sets a tacit categorization parameter (e.g., based on color) and subjects have to infer the underlying rule being used to find the correct position for the cards. The examiner only states whether the selected position is correct or not, so that the participant must infer the rule being used. At a certain point the examiner changes the rule (e.g., the categorization parameter switches to shape) without informing the participant, and the subject has to adapt his operative mental schema to infer the new rule and respond accordingly. In the present study, statistical analyses were based on the number of categories correctly inferred and completed.

Verbal fluency

Vocabulary search skills and word retrieval efficiency under time pressure were assessed via two phonological and two semantic fluency tasks, one in each language. Whereas the former focuses on form-level search skills (using phonemes as cues), the latter taps into conceptually mediated search mechanisms (within a predefined semantic field). Before the picture-naming, reading, and translation tasks, participants were randomly assigned the phonological condition in one language and the semantic one in the other. The remaining conditions were performed after the three above-mentioned tasks – for a detailed account of these tasks’ counterbalancing scheme, see Table S2 in the Supplementary Material (Supplementary Material).

Phonological fluency was assessed through the Controlled Oral Word Association Test COWAT (Benton & Hamsher, 1989). Participants were allotted 60 seconds to utter as many words as they could say starting with a specific phoneme (/f/, /a/, and /s/).[2] Importantly, the cumulative numbers of candidate items, considering all parts of speech (4,067 for Spanish and 4,663 for English) and only nouns (2,136 for Spanish and 2,284 for English), were similar between languages [all parts of speech: $t(4) = -0.378905), p = .72$; nouns: $t(4) = -0.163470), p = .88$], which indicates that task difficulty, in terms of potential responses, was similar between languages – these analyses were based on the 20,000 more frequent words in Spanish (Davies, 2008a) and English (Davies, 2008b). The order of cue phonemes was alternated between languages and participants. For analysis, the performance on each of the three cue phonemes was averaged for each subject. As regards the semantic fluency tests, participants were given 60 seconds to produce all the words they could think of belonging to the category ANIMALS. All responses were audio-recorded in .mp3 files, which were then transcribed by one examiner and checked by another one.

In line with standard analysis procedures for both tasks (Spreen & Strauss, 1991), repeated words, proper names,

---

[2] Participants were explicitly instructed to rely on the presented phonemes as cues, and not on particular letters that may correspond with them in print.

and words corresponding to the same lemma were not taken into consideration. Indeed, if repeated words were factored in the analyses, performance could be biased by constant reliance on one item (e.g., *sock*, *soup*, *sock*, *sock*, *steel*, *sock*), and the same could happen if valid responses included derivates of the same noun (e.g., *perro*, *perra*, *perros*, *perras*, *perrito*, *perrita*) or proper names (e.g., *Mariano*, *Mariana*, *Marianito*, *Marianita*). Under these parameters, individual scores for each condition were computed as the total number of valid responses given, with higher scores indicating better vocabulary retrieval and search skills.

Picture naming

The participants' lexicalization skills (i.e., going from concepts to words) were assessed with a picture-naming task. Unlike semantic fluency tasks, which tap into the ability to activate numerous concepts within a broad semantic field, this paradigm examines the capacity to produce specific words guided by image-evoked semantic constraints. The stimuli consisted of 64 black-and-white images from Cycowicz et al. (Cycowicz, Friedman, Rothstein & Snodgrass, 1997), all corresponding to concrete nouns. They were presented in two 32-item pseudorandomized blocks, each including 16 cognate and 16 non-cognate targets. The blocks were counterbalanced across participants, so that half responded to the first one in L1 and to the second one in L2, while the other half did so in the reverse order – for a detailed account of these tasks' counterbalancing scheme, see Table S3 in the Supplementary Material (Supplementary Material). The images in each set were similar in terms of name agreement ($t(62) = -1{,}17$, $p = .25$), familiarity ($t(62) = .42$, $p = .67$), image agreement ($t(62) = 1.87$, $p = .07$), and visual complexity ($t(62) = .69$, $p = .49$), based on normative data for the adult Argentine population (Manoiloff, Artstein, Canavoso, Fernandez & Segui, 2010). Similarly, the target words in each set were matched for length of the L1 target ($t(62) = .00$, $p = 1$), length of the L2 target ($t(62) = .00$, $p = 1$), frequency of the L1 target ($t(62) = .07$, $p = .94$), and frequency of the L2 target ($t(62) = .66$, $p = .51$). Moreover, there were no significant differences in length ($t(126) = .73$, $p = .47$) or frequency ranking ($t(126) = .44$, $p = .66$) between the full sets of L1 and L2 targets.

Each trial began with a fixation cross, shown for 300 ms. A picture was then displayed for 800 ms against a white background. The following trial was launched 1 second after a response was made, or after 2500 ms if no response was given. Participants were instructed to name each image aloud as fast and accurately as possible, in the language indicated by the instructions. An examiner kept track of response accuracy for each trial, while reaction times were recorded by the computer. The task lasted approximately 10 minutes.

Word translation

Word translation was assessed with a previously reported task (for full details, see García, Ibáñez, Huepe, Houck, Michon, Lezama, Chadha & Rivera-Rei, 2014). This paradigm offers an objective measure of the participants' ability to activate cross-linguistic equivalents of specific source-language items (i.e., words with considerable or maximal semantic overlap between languages). Two blocks of 64 English nouns (EN1, EN2) were used for backward translation (BT, from L2 to L1), and another two blocks of 64 equivalent Spanish nouns (SP1, SP2) were used for forward translation (FT, from L1 to L2) – for a detailed account of these tasks' counterbalancing scheme, see Table S4 in the Supplementary Material (Supplementary Material). To prevent translation priming effects, each participant was assigned only one English block and only one Spanish block (i.e., EN1 and SP2, or EN2 and SP1). All blocks had the same number ($n = 16$) of concrete cognates (e.g., *paper*, *papel*), abstract cognates (e.g., *comedy*, *comedia*), concrete non-cognates (e.g., *table*, *mesa*), and abstract non-cognates (e.g., *punishment*, *castigo*). Stimuli in each language were matched for frequency ranking ($p = 0.97$) and syllabic length ($p = 0.99$), and blocks for each language were additionally matched for frequency (Spanish: $p = 0.95$; English: $p = 0.98$) – data for these variables were extracted from (Davies, 2008a, b). The blocks were counterbalanced across participants, and stimuli within them were pseudorandomly distributed.

Each trial started with a fixation cross 300 ms prior to the stimulus, which was displayed in white letters against a black background for 200 ms. Participants were instructed to translate each word aloud, as fast and accurately as possible, into the language indicated at the beginning of each block. An examiner kept track of response accuracy for each trial, while reaction times were recorded by the computer. The examiner judged the responses with a control grid that included the specific word that was considered as valid in each case. The following five rejection criteria were considered: (i) no response (e.g., subject remains silent); (ii) hesitation or false start (e.g., *fury* → *fueg... furia*!); (iii) task confusion (e.g., subject reads when asked to translate, or vice versa); (iv) wrong translation (e.g., *fury* → *fuera*); and (v) non-predefined translation (e.g., *fury* → *ira*). In all cases, one examiner first went over each participant's responses to identify invalid trials, and then a second examiner checked the procedure to spot inaccuracies. In the few cases in which these emerged, they were settled by a third member of the team. The task lasted approximately 20 minutes. For full methodological details, see García et al. (2014).

Word reading

The word reading task was also taken from García et al. (2014). This paradigm taps into the participants' skills to articulate words prompted by their written representation, so that shallow (i.e., form-level) mechanisms suffice for task completion (Sasanuma et al., 1992; Teichmann et al., 2012). Stimuli consisted of two blocks of 64 nouns, one in each language. The two blocks were matched for number ($n = 16$) of concrete cognates, abstract cognates, concrete non-cognates, and abstract non-cognates, as well as frequency ranking ($p = .99$) and syllabic length ($p = .99$) – data for the latter two variables were extracted from Davies (2008a, b). The blocks were counterbalanced across participants, and stimuli within them were pseudorandomly distributed – for a detailed account of these tasks' counterbalancing scheme, see Table S4 in the Supplementary Material (Supplementary Material). Participants were instructed to read each word out loud, as fast and accurately as possible, in the same language of presentation. The stimulus presentation, data recording, and trial rejection procedures were exactly the same as the ones applied in the word translation tasks. The task lasted approximately 10 minutes. For full methodological details, including set-up features, statistical details, and trial rejection criteria, see García et al. (2014).

Statistical analysis

As in multiple previous studies in the field (Babcock & Vallesi, 2015; Bialystok, Craik & Luk, 2008; Christoffels et al., 2006; Friesen, Luo, Luk & Bialystok, 2015; Heikoop, Declerck, Los & Koch, 2016; Hernández, Costa, Fuentes, Vivas & Sebastián-Gallés, 2010; Hernández, Martin, Barceló & Costa, 2013; Ibáñez, Macizo & Bajo, 2010; Kousaie, Sheppard, Lemieux, Monetta & Taler, 2014; Padilla, Bajo & Macizo, 2005; Poarch & van Hell, 2012; Prior & Macwhinney, 2009; Stavrakaki, Megari, Kosmidis, Apostolidou & Takou, 2012; Yudes, Macizo & Bajo, 2011), experimental data for each task were separately analyzed using mixed-effects ANOVA and Tukey's HSD tests (to address necessary corrections in post-hoc contrasts). Effect sizes were calculated with partial eta squared (η^2). Also, to assess whether observed differences were directly related to interpreting skills, we reanalyzed results from all tasks yielding significant main effects of group or interactions through ANCOVAs using "competence in L2-L1 interpreting" and "competence in L1-L2 interpreting" (Table 1) as covariates. In addition, for all linguistic tasks yielding group differences, we performed correlation analyses between the outcome in each language/direction and the PSIs' years of interpreting experience. Finally, to assess the role of the participants' domain general skills on their lexical abilities, we ran multiple linear regressions between outcomes in the executive tasks (digit span, WCST, IFS) and their performance in each word-processing task. Standard benchmarks were used to discriminate among small ($\eta^2 = 0.01$), medium ($\eta^2 = 0.06$), and large ($\eta^2 = 0.14$) effects (Cohen, 1988). Alpha values were set at $p < .05$. All statistical analyses were performed on Statistica 10 (http://www.statsoft.com/).

Results

In each task, and in line with previous lexical processing studies (e.g., Ferré et al., 2006; García & Ibáñez, 2016; Guasch et al., 2008), participants whose mean scores were 2 standard deviations away from the group's mean in each task and condition were considered outliers and removed only from the corresponding analyses – no other scores were interpolated in their stead. Additional subject-level data points were removed due to incorrect logging by the software or inaccurate audio recordings. See Table S5 (Supplementary Material) for details about the percentage of participants excluded in each task.

Verbal fluency

To examine between-group differences in form-based lexical search, on the one hand, and semantically-driven lexical search, on the other, data from the phonological and the semantic fluency tasks were analyzed via two separate 2 x 2 mixed-effects ANOVAs, with group (NIBs and PSIs) as a between-subjects factor and language (L1 and L2) as a within-subject factor. The rates of rejected responses for each task were similar between groups and languages. Specifically, for NIBs, rejected responses in the phonological condition amounted to 14.8% in L1 and 13.8% in L2, and in the semantic condition they amounted to 3.1% in L1 and 1.7% in L2. For PSIs, rejected responses in the phonological condition amounted to 17.5% in L1 and 15.6% in L2, and in the semantic condition they amounted to 7.4% in L1 and 3.7% in L2. Neither condition yielded significant between-group differences (phonological fluency in L1: $\chi^2 = 0.037, p = 0.85$; phonological fluency in L2: $\chi^2 = 0.039, p = 0.84$; semantic fluency in L1: $\chi^2 = 0.947, p = 0.33$; semantic fluency in L2: $\chi^2 = 0.172, p = 0.68$).

Results from the phonological fluency task revealed significantly better performance for PSIs over NIBs [$F(1, 28) = 16.75, p < .001$, partial $\eta^2 = .37$]. We also observed a main effect of language [$F(1, 28) = 15.27, p < .001$, partial $\eta^2 = .35$], with more words produced in L1 than in L2. The interaction between the two factors was not significant [$F(1, 28) = 2.12, p = .16$, partial $\eta^2 = .07$]. These results remained the same after the ANCOVA analysis for the main effects of group [$F(1, 26) = 9.65, p < .005$, partial $\eta^2 = .27$] and for the interaction effect [$F(1, 26) = 1.02, p = .32$, partial $\eta^2 = .04$]. However, the main effect of language did not survive the covariance analysis [$F(1, 26) = 0.12, p = .71$,

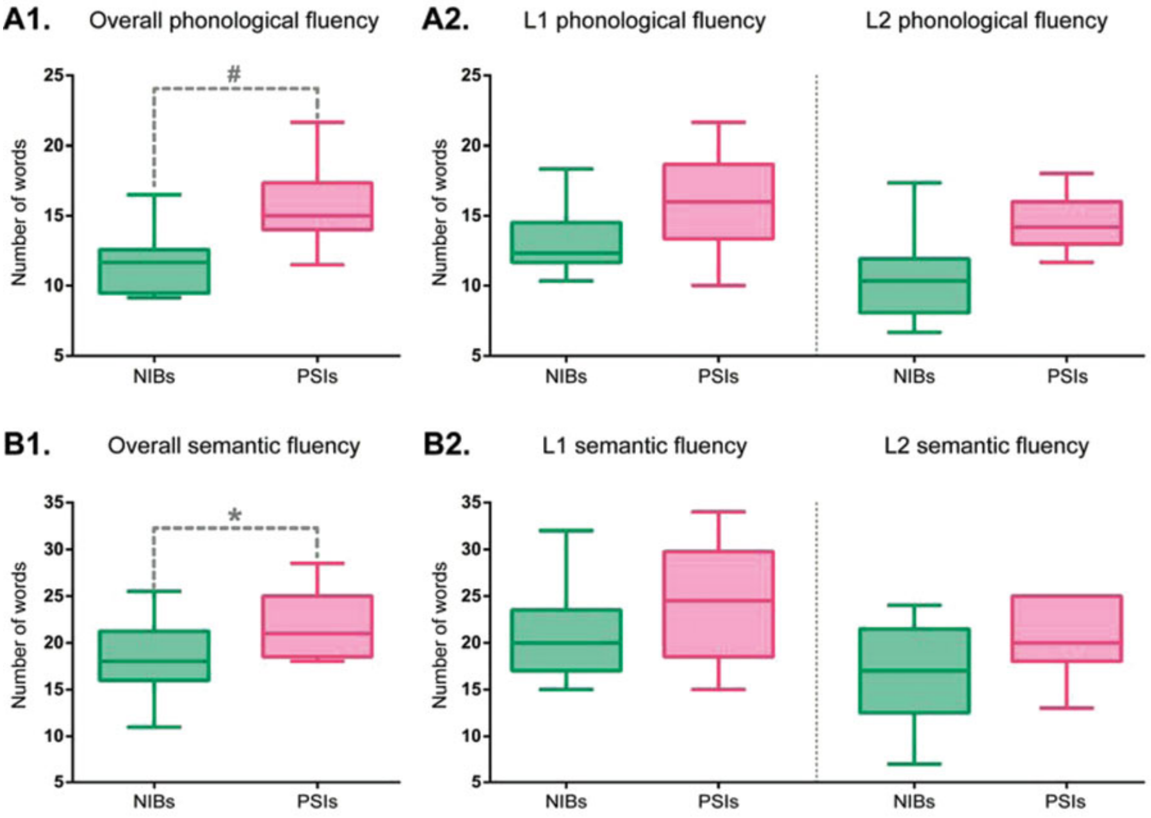

Figure 1. (Colour online) Phonological and semantic fluency results. **A1.** Overall phonological fluency: PSIs produced more words than NIBs when considering their performance in both languages. However, covariance results indicated that this pattern was not dependent on interpreting expertise. **A2.** Groups' performance in each language: no statistical differences were found in the language-by-group interaction. **B1.** Overall semantic fluency: PSIs outperformed NIBs. This difference disappeared after covariance analysis, suggesting that it depended on interpreting expertise. **B2.** Groups' performance in each language: as for phonological fluency, the language-by-group interaction was not significant. PSIs: professional simultaneous interpreters; NIBs: non-interpreter bilinguals. The asterisk (*) indicates significant differences at $p < .05$ for mixed ANOVA analysis only. The hash (#) identifies significant differences at $p < .05$ for both the mixed ANOVA and the ANCOVA analyses.

partial $\eta^2 = .005$]. These results are captured in Figure 1A. Also, a complementary analysis showed that, for PSIs, there was no association between the years of interpreting experience and phonological fluency performance in either language (see Table S6 in the Supplementary Material (Supplementary Material)). Neither was their performance predicted by any of the executive function measures (see Table S7a in the Supplementary Material (Supplementary Material)).

A similar pattern emerged in the semantic condition. PSIs produced significantly more words than NIBs [$F(1, 27) = 6.18, p = .019$, partial $\eta^2 = .19$], and production was larger in L1 than in L2 [$F(1, 27) = 11.43, p = .002$, partial $\eta^2 = .3$]. The group-by-language interaction did not reach significance [$F(1, 27) = 1.63, p = .75$, partial $\eta^2 = .004$]. However, the main effect of language [$F(1, 25) = .21, p = .65$, partial $\eta^2 = .008$], the effect of group [$F(1, 25) = 0.63$, $p = .435$, partial $\eta^2 = .025$], and the interaction effect [$F(1, 25) = 0.28$, $p = .6$, partial $\eta^2 = .01$] were not significant after covariation with interpreting competence. These results are graphically shown in Figure 1B. Moreover, no significant correlations emerged between the PSIs' years of experience and their semantic fluency performance in either language (see Table S6 in the Supplementary Material (Supplementary Material)). Finally, the PSIs' performance was not predicted by any of the executive function measures (see Table S7b in the Supplementary Material (Supplementary Material)).

In sum, these results indicate that vocabulary search skills were better for PSIs than NIBs (when guided by either phonological or semantic cues) and that both groups performed better in their L1 than in their L2. However, only differences in the semantic condition seemed to depend directly on interpreting competence. Furthermore, for PSIs, performance in none of the conditions was associated with the years of experience.

Table 2. *Picture naming accuracy.*

	NIBs $n = 17$	PSIs $n = 17$
L1 accuracy	.86 (.07)	.83 (.05)
L2 accuracy	.75 (.12)	.77 (.08)

Data presented as mean (*SD*). NIBs: non-interpreter bilinguals; PSIs: professional simultaneous interpreters; L1: native language; L2: foreign language.

Table 3. *Word translation accuracy.*

	NIBs $n = 17$	PSIs $n = 17$
BT accuracy	.85 (.07)	.87 (.05)
FT accuracy	.85 (.09)	.87 (.06)

Data presented as mean (*SD*). NIBs: non-interpreter bilinguals; PSIs: professional simultaneous interpreters; BT: backward translation; FT: forward translation.

Picture naming

Accuracy and reaction time (RT) data were analyzed via two separate mixed-effects two-way ANOVAs. Both analyses included group (NIBs and PSIs) as a between-subject factor and language (L1 and L2) as a within-subject factor.

Accuracy analyses revealed that both groups were similarly accurate [$F(1, 29) = .047, p = .83$, partial $\eta^2 = .002$]. They also showed an effect of language, with higher accuracy in L1 than in L2 [$F(1, 29) = 23.79$, $p < .001$, partial $\eta^2 = .45$]. The interaction between the two factors was not significant [$F(1, 29) = 2.65$, $p = .11$, partial $\eta^2 = .08$]. Mean scores and standard deviations are shown in Table 2.

Results showed that PSIs did not significantly differ from NIBs in their RTs [$F(1, 30) = 2.68, p = .11$, partial $\eta^2 = .08$]. Besides, naming was faster in L1 than in L2 [$F(1, 30) = 23.92$, $p < .001$, partial $\eta^2 = .44$]. The interaction between both factors was not significant [$F(1, 30) = 1.34$, $p = .26$, partial $\eta^2 = .04$]. After the covariance analysis, whereas the language effect remained significant [$F(1, 28) = 7.57$, $p = .010$, partial $\eta^2 = .21$], no significant differences were observed in the effect of group [$F(1, 28) = 0.02$, $p = .88$, partial $\eta^2 = .001$] and the interaction effect [$F(1, 28) = 0.39$, $p = .535$, partial $\eta^2 = 0.01$]. These results are captured in Figure 2A. Also, the PSIs' picture-naming performance was not predicted by any of the executive function measures (see Table S7c in the Supplementary Material (Supplementary Material)).

Overall, both groups evinced a similar performance in retrieving semantically constrained words. Also, both groups were characterized by greater efficacy and efficiency to perform the task in L1 than in L2, even after covariation analysis.

Word translation

Translation performance was analyzed through a 2 x 2 mixed-effects ANOVA, with group (NIBs and PSIs) as a between-subjects factor and directionality (BT and FT) as a within-subject factor.

Accuracy was similar between groups [$F(1, 30) = 0.68$, $p = .42$, partial $\eta^2 = .02$] and conditions [$F(1, 30) = .000$, $p = .99$, partial $\eta^2 < 001$]. The interaction between both factors was not significant [$F(1, 30) = .03$, $p = .87$, partial $\eta^2 < .001$]. Mean scores and standard deviations are shown in Table 3.

Instead, translation performance was significantly faster for PSIs than NIBs [$F(1, 32) = 6.006$, $p = .02$, partial $\eta^2 = .16$], and BT involved lower RTs than FT [$F(1, 32) = 12.49, p = .001$, partial $\eta^2 = .28$] . However, no significant RT differences emerged in the interaction between group and directionality [$F(1, 32) = 1.3, p = .26$, partial $\eta^2 = .04$]. Also, the effect of group did not survive covariation with interpreting competence [$F(1, 30) = 0.02, p = .89$, partial $\eta^2 = .001$.], which also yielded null effects of translation direction [$F(1, 30) = 0.25$, $p = .62$, partial $\eta^2 = .01$.] and interaction [$F(1, 30) = 2.45, p = .13$, partial $\eta^2 = .075$]. These results are graphically presented in Figure 2B. Furthermore, the amount of experience in PSIs was negatively correlated with translation speed in each direction (see Table S6 in the Supplementary Material (Supplementary Material)). Note, in addition, that the PSIs' performance was not predicted by any of the executive function measures (see Table S7d in the Supplementary Material (Supplementary Material)). Finally, regression analyses on data from both groups collapsed showed that only translation speed – as opposed to translation accuracy – could be predicted for both BT and FT in terms of the subjects' self-reported competence in the corresponding direction (Table S8a). However, separate regression analyses for each group (Table S8b and S8c) revealed that, only in PSIs, RTs for FT could be predicted by self-reported competence in that direction.

In short, both groups were similarly accurate in BT and FT. However, PSIs were characterized by faster performance in both directions, and they exhibited similar latencies for BT and FT. Once again, ANCOVA results indicated that these differences depended on the participants' level of interpreting competence, even though the PSIs' translation performance was not associated with their years of experience.

Word reading

The analysis of word-reading results followed the same statistical approach used for picture naming.

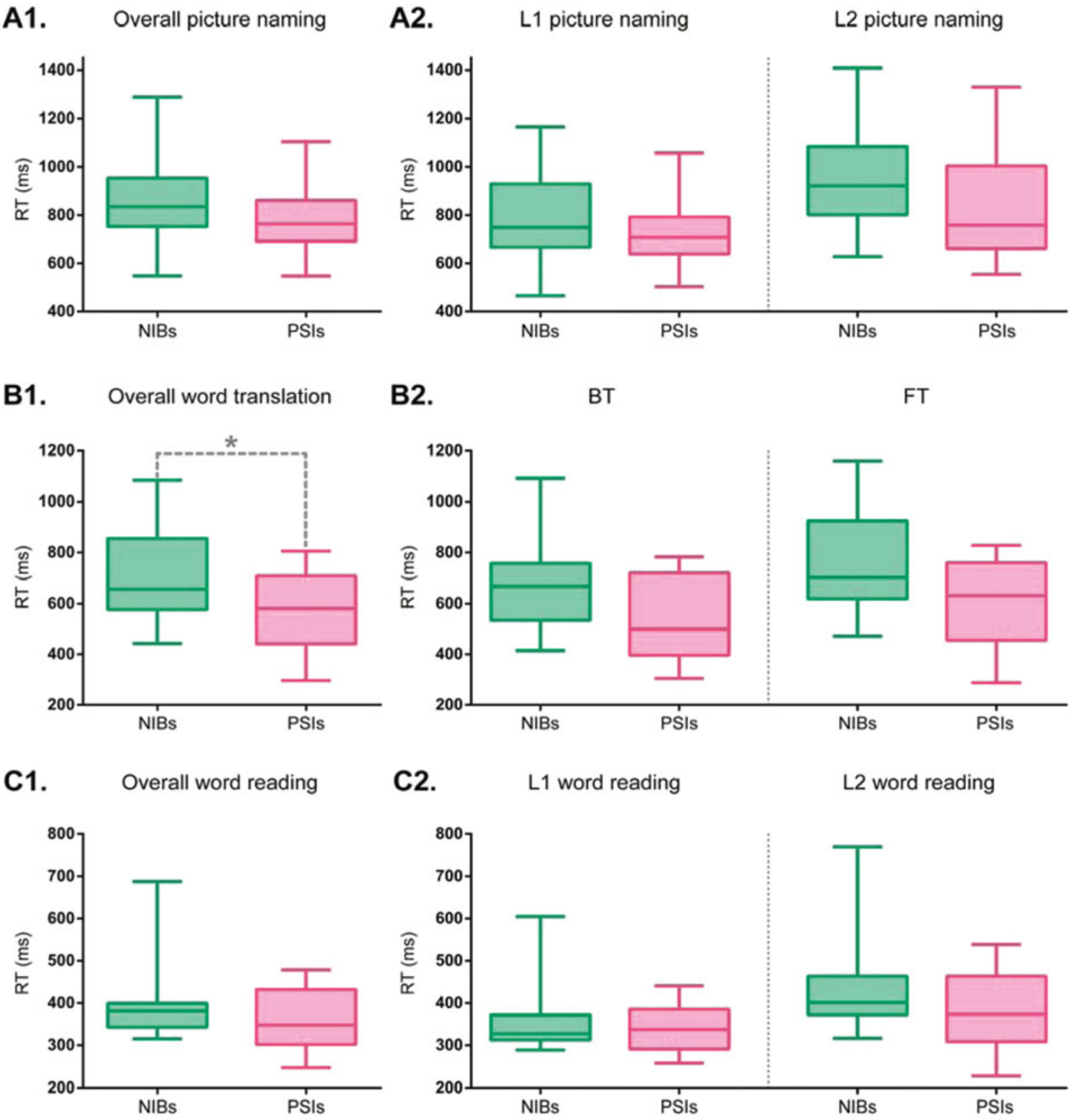

Figure 2. (Colour online) **Reaction time results for the picture naming, word translation, and word reading tasks. A1.** Overall picture naming: PSIs and NIBs presented similar reaction times. **A2.** Groups' performance in each language: the interaction effect between language and group was not significant. **B1.** Overall word translation: PSIs were significantly faster than NIBs. This difference disappeared after covariance analysis, suggesting an influence of interpreting expertise. **B2.** Groups' performance in each language: no statistical differences were found in the language-by-group interaction. **C1.** Overall word reading: The two groups showed similar reaction times. **C2:** Groups' performance in each language: the language-by-group interaction was not significant. PSIs: professional simultaneous interpreters; NIBs: non-interpreter bilinguals. The asterisk (*) indicates significant differences at $p < .05$ for mixed ANOVA analyses only.

Accuracy results showed no significant differences between groups [$F(1, 32) = 1.37$, $p = .25$, partial $\eta^2 = .04$] or languages [$F(1, 32) = 1.97$, $p = .17$, partial $\eta^2 = .06$]. The interaction between those variables was not significant either [$F(1, 32) = 2.27$, $p = .14$, partial $\eta^2 = .07$]. Mean scores and standard deviations are shown in Table 4.

RT results also showed similar performance for both groups [$F(1, 32) = 1.055$, $p = .31$, partial $\eta^2 = .03$], although processing was faster for L1 than L2 collapsing

Table 4. *Word reading accuracy.*

	NIBs $n = 17$	PSIs $n = 17$
L1 accuracy	.99 (.01)	.99 (.01)
L2 accuracy	.97 (.04)	.99 (.01)

Data presented as mean (*SD*). NIBs: non-interpreter bilinguals; PSIs: professional simultaneous interpreters; L1: native language; L2: foreign language.

both groups [$F(1, 32) = 29.8, p < .001$, partial $\eta^2 = .48$]. However, the interaction between group and language was not significant [$F(1, 32) = 2.19$, $p = .15$, partial $\eta^2 = .06$]. These results are shown in Figure 2C. Finally, word reading efficiency in PSIs was not predicted by any of the executive function measures (see Table S7e in the Supplementary Material (Supplementary Material)).

In sum, word-reading performance was similar for both groups, which also coincided in showing more efficient reading skills for L1 than L2 words.

Discussion

This study examined whether various operations within bilingual memory are sensitive to the sustained extreme demands characterizing the activity of PSIs. We found a two-fold pattern. On the one hand, PSIs exhibited better vocabulary search skills and they were faster to translate words in both directions. On the other hand, they showed no advantages for picture naming or word reading in either language. Such findings shed light on the cognitive particularities of PSIs and, more generally, on the adaptability of bilingual memory mechanisms to extreme processing conditions.

The selective impact of interpreting expertise on bilingual memory mechanisms

Verbal fluency tests revealed better lexical retrieval skills in L1 than in L2 for both groups. This replicates well-established patterns of asymmetrical performance in bilinguals with various levels of L2 and translation competence during single-language tasks (French & Jacquet, 2004; García et al., 2014; Kroll & Stewart, 1994). Yet, the most relevant finding was that, in both languages, PSIs showed greater efficiency than NIBs. This pattern emerged for both the phonological and the semantic fluency conditions, suggesting that sustained interpreting demands can enhance the ability to quickly find words satisfying task-specific constraints.

In line with these findings, PSIs have been observed to outperform NIBs in tasks which tax sublexical search mechanisms – e.g., lexical decisions on non-words (Bajo et al., 2000) – and meaning-based vocabulary search – e.g., categorization of non-typical exemplars (Bajo et al., 2000) and detection of semantic errors (Fabbro, Gran & Gran, 1991; Yudes, Macizo, Morales & Bajo, 2013). In this sense, note that lexical retrieval processes in professional interpreting can engage either form-level or conceptual routes at different times (De Groot & Christoffels, 2006; Paradis, 1994). The advantage we found on both fluency tasks may thus follow from the continual demands that PSIs place on both levels of processing during task-constrained lexical retrieval.

However, results from the picture-naming task indicate that no such advantage emerges when words must be accessed within the constraints of restrictive picture-evoked conceptual schemas. Indeed, the same result was also observed in the only previous assessment of picture-naming skills in PSIs (Christoffels et al., 2006). Thus, the semantic retrieval advantages exhibited by PSIs in the fluency task disappeared when words had to be activated based on schemas that satisfy highly specific semantic constraints.

The emerging pattern is that extreme demands placed by PSIs on bilingual memory affect mechanisms that are differentially taxed in professional settings. This is consistent with the results from the translation tasks. Although NIBs were as accurate as PSIs, the latter were significantly faster and, once again, this effect was not driven by their performance in either BT or FT. *Prima facie*, the latter result might seem at odds with Elmer et al.'s (2010) finding that PSIs exhibit specifically distinct electrophysiological markers of semantic processing in the L2-L1 direction – for PSIs, this was the only language combination failing to evince enhanced N400 modulations during semantic decision. However, note that the PSIs in that study had been trained specifically in L2-L1 interpreting, suggesting that each interpreting direction can manifest specific adjustments depending on the particular forms of expertise developed by the PSIs. This implication actually aligns with our new findings: given that professionals in the present investigation reported possessing higher skills to interpret into each of their languages and doing so much more frequently than NIBs (see Table 1), it seems that cross-language connections for each direction can also become more efficient when repeatedly subjected to extreme demands. Note, in this sense, that regression analyses (Table S8b and S8c) showed that only in PSIs were translation RTs predictable from self-reported competence in the corresponding direction – although results for BT were only marginally significant.

A previous comparison between PSIs with NIBs yielded similar outcomes for both BT and FT (Christoffels et al., 2006). Those results, as well as our own, are consistent with the 'activation threshold hypothesis', which posits that the more a connection is engaged during bilingual processing, the lesser the amount of stimulation needed to activate it (Paradis, 1993, 1994). It seems plausible that crosslinguistic links in PSIs developed lower thresholds due to their systematic recruitment in working settings. In fact, reaction times for BT and FT were the only two variables showing significant (negative) correlations with the PSIs' years of interpreting experience (see Table S6), suggesting that translation speed increased in proportion to time spent in the profession. Interestingly, though, Christoffels et al. (2006) found no differences in translation performance when comparing PSIs with L2 teachers. However, this finding does not contradict the above interpretation.[3] In fact, L2 teachers can consistently rely on translation in their classes (Pekkanli, 2012) and develop greater (informal) translation competence than bilinguals who are not language professionals (García et al., 2014), Exp. 1). It is thus likely that translation-relevant routes become strengthened by any regular activity that continually taxes them.

On the other hand, PSIs showed no word-naming advantage in either language. Crucially, visual word recognition and phonological production subsystems (de Groot, Borgwaldt, Bos & van den Eijnden, 2002) are not taxed during simultaneous interpretation – reading skills have no critical bearing in this activity and the rate of production is not higher than that characterizing single-language communication (Chernov, 2004; Gerver, 1975). More particularly, word reading is a shallow task, which can be accomplished without semantic information or other mechanisms involved in vocabulary search (Sasanuma et al., 1992; Teichmann et al., 2012). In fact, while electrophysiological evidence indicates that PSIs, as opposed to NIBs, would be characterized by "a training-induced altered sensitivity to semantic processing within and across L1 and L2" (Elmer et al., 2010, p. 152), semantic manipulations fail to modulate word naming performance regardless of the subjects' experience in interlinguistic reformulation (García et al., 2014). Taken together, then, our findings indicate that the enhancements linked to interpreting expertise are confined to those skills that are specifically taxed during professional performance.

Finally, note that the enhancements of PSIs were not driven by any particular target language. Previous studies on executive skills have offered mixed results concerning this point, showing superior performance of PSIs in both languages for certain tasks (Christoffels et al., 2006), but not others (Chincotta & Underwood, 1998; Tzou, Eslami, Chen & Vaid, 2011). One of the possible factors underlying those discrepancies is whether both the L1 and the L2 are regularly used as source and target languages. Our present findings, stemming from lexical tasks, likely reflect the interpreters' greater expertise at working into their two languages. Indeed, relative to NIBs, PSIs perceived themselves as more competent in BT and FT and they spent significantly more hours a week interpreting in each direction. Thus, the interpreters' language-non-selective advantages may reflect the high semantic and cross-linguistic demands posed on both languages and directions.

Notably, between-group differences in interpreting competence emerged robustly, and with large effect sizes, even though NIBs judged their interpreting skills as roughly "intermediate" (with scores of 4.82 for L2-L1 and 4.53 for L1-L2, over a maximum of 7). This possible overestimation may have been driven by the caveats inherent to self-report measures. However, even if excessive, these values were significantly lower ($< .001$) than those of PSIs, who, in agreement with their sustained experience in the field, actually estimated their skills as near-optimal (6.53 for L2-L1 and 6.35 for L1-L2). Although this pattern aligns with the overall rationale of the study, it would be important to acknowledge the potential caveats of subjective competence assessments and complement them with objective measures (see "Limitations and avenues for further research" section).

In sum, PSIs were characterized by increased lexical efficiency, but only in tasks requiring quick vocabulary search or interlinguistic processing. These two operations, as opposed to word reading, are differentially taxed in the particular linguistic scenarios they recurrently face. Thus, extreme demands placed on bilingual memory do not lead to holistic reconfigurations of the system. On the contrary, processing enhancements seem confined to those subfunctions that become specifically recruited for expert performance.

Insights into the organization and flexibility of bilingual memory

Our results show that, despite their constant interplay and tight coactivation, specific mechanisms within bilingual memory can SELECTIVELY adapt to meet elevated processing demands. This overall finding underscores the functional independence of distinct mechanisms within the system, thus shedding light on its architecture.

First, the enhancement of translation performance in the absence of increased reading efficiency attests to the autonomy of interlinguistic (as opposed to intralinguistic) connections within bilingual memory. Our claim is supported by two separate strands of evidence. Crucially, behavioral evidence has repeatedly shown differential

---

[3] Note, in this sense, that only three subjects in our NIB group were L2 teachers, and that these subjects' translation performance was similar to that of the remaining NIBs ($N = 14$) – BT accuracy: [$t(15) = 1.402$, $p = .18$], BT response times: [$t(15) = -0.784$, $p = .44$], FT accuracy: [$t(15) = 1.701$, $p = .11$], FT response times [$t(15) = -1.316$, $p = .21$]. This indicates that NIBs' results in our study were not driven by the performance of the few L2 teachers included in the sample.

performance between both tasks in subjects with and without formal training in interlingual reformulation (García et al., 2014; Kroll & Stewart, 1994; Kroll et al., 2010). Also, neurological evidence indicates that reading and translation skills involve differential activity patterns (García, 2013; Price, Green & von Studnitz, 1999) and can become doubly dissociated following brain lesions (García, 2015b) or selectively inhibited with direct cortical stimulation of circumscribed cortical sites (Borius et al., 2012).

Second, the differential pattern between semantic- and form-level operations (e.g., semantic fluency vs. word reading) supports hierarchical models of bilingual memory – namely, models that recognize partially distinct processing mechanisms at the levels of meaning and form. This observation, indeed, is consistent with studies in various bilingual groups showing dissociations between conceptual and sublexical manipulations via translation, masked interlinguistic priming, equivalent recognition, and word association paradigms (for reviews, see Brysbaert & Duyck, 2010; García, 2015a; Kroll et al., 2010). Moreover, it aligns with clinical evidence of bilingual aphasics capable of establishing interlinguistic correspondences despite semantic access impairments (De Vreese, Motta & Toschi, 1988; Paradis, Goldblum & Abidi, 1982) – for a review, see García (2015b).

From a theoretical perspective, the evidence exposes the limitations of models which fail to introduce explicit distinctions between routes specialized for reading and translation, or between semantic and form-level systems (van Heuven et al., 1998). Conversely, it supports models which relate those oppositions with specific connections and strata, respectively (Dijkstra & van Heuven, 2002; Dong et al., 2005; Kroll & Stewart, 1994; Kroll et al., 2010; Van Hell & De Groot, 1998). Our findings contribute to the literature by showing that these functionally independent subsystems or operations within bilingual memory can selectively adapt to meet recurrent high demands placed on them.

Furthermore, present results suggest that the intensity and recurrence with which the above mechanisms are used can enhance relevant connections beyond the well-established effects of L2 proficiency, age of acquisition, or degree of exposure. In particular, PSIs in our study were matched with NIBs in terms of L2 competence, age of L2 learning, years of L2 study, and weekly hours of exposure to L2 media, which suggests that the observed effects were probably not driven by such factors. In the same vein, bilingualism *per se* has no beneficial effects on semantic fluency in adults (Bialystok et al., 2008; Friesen et al., 2015; Sandoval, Gollan, Ferreira & Salmon, 2010), and increased translation competence entails lexical processing enhancements even when samples are matched for L2 competence (García et al., 2014). Although objective assessments would be required to better assess the issue, self-report measures of L1 and L2 skills have been found to closely replicate reaction-time results (Langdon, Wiig & Nielsen, 2005), successfully predict language ability (Marian, Blumenfeld & Kaushanskaya, 2007), and reproduce statistical results of multilingual naming tests for classifying bilinguals into language-dominance groups (Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012). Therefore, the differences observed between groups seem to be specifically associated to interpreting expertise, as opposed to field-unspecific variables modulating bilingual performance – for compatible claims about executive skills, see Babcock and Vallesi (2015).

Indeed, ANCOVA results showed that between-group differences in semantic fluency and word translation disappeared after covariation with interpreting competence, suggesting a critical influence of this variable on such effects. The only exception to this pattern concerned phonological fluency differences, which did survive covariation with interpreting competence. Given that the groups were matched for bilingualism-related factors (e.g., L2 competence, age of L2 learning, years of L2 study), we surmise that phonological fluency may be more critically dependent on other factors non-linearly related to interpreting competence, such as phonological awareness. Also, the pattern may be more simply reflecting a ceiling effect in task performance. However, since our study did not include tests aimed to assess these factors, these claims remain conjectural and should be further addressed in future studies.

Finally, it has not escaped our attention that lexical processing also recruits executive functions (Christoffels et al., 2006; Christoffels, De Groot & Waldorp, 2003; Martin, Wiggs, Lalonde & Mack, 1994; McDowd, Hoffman, Rozek, Lyons, Pahwa, Burns & Kemper, 2011), which would in principle weaken the proposed association between interpreting expertise and adaptations of bilingual memory proper. However, reported semantic processing advantages in PSIs are independent of working memory (Yudes et al., 2013), and enhancements of the latter domain do not necessarily correlate with superior lexical retrieval (Christoffels et al., 2006). Similarly, although working memory capacity correlates with word translation performance in NIBs (Kroll et al., 2002), previous research indicates that PSIs' remarkable ability to engage in concurrent source-language comprehension and target-language production depends on lexical skills rather than on executive enhancements (Padilla et al., 2005). Also, results from graphical modeling show that translation and working memory skills are independent components which contribute separately to interpreting performance (Christoffels et al., 2003). More crucially, increased efficiency in the present study cannot be attributed to PSIs' executive skills, as both samples were matched for working memory, cognitive flexibility, and overall executive performance. In fact, multiple linear regressions in the present study showed that the PSIs' performance in each lexical task was not predicted by executive skills (see Tables S7a–S7e). Also, although some studies have found enhanced performance for PSIs than NIBs on certain executive domains (García, 2014), this pattern is not entirely systematic – for instance, non-significant differences have been reported by Köpke and Nespoulous (2006) and Signorelli et al. (2011). Consequently, we propose, sustained extreme demands during bilingual processing can hone specific lexical mechanisms in a *sui generis* fashion, irrespective of the contributions of domain-general functions.

In sum, despite their joint and tight cooperation, different subcomponents within bilingual memory seem characterized by independent usage-driven flexibility. This complex system is susceptible to sustained extreme demands beyond the effects of L2 mastery, age of acquisition, and degree of exposure. Moreover, the selective adaptation of its inner mechanisms is not epiphenomenal to extralinguistic factors. These findings, derived from PSIs as a model of expert bilingual processing, can fruitfully extend the field's current research agenda.

Limitations and avenues for further research

Our work features a number of limitations that pave the way for further research. First, we were unable to include additional tasks tapping other bilingual memory mechanisms, such as lexical and semantic decision, equivalent recognition, or associated word production. These paradigms could be incorporated in replications or extensions of our study as an additional testing ground of its conclusions. Second, the present findings are blind to the impact of interpreting expertise on other verbal units (e.g., sentences or supra-sentential texts), which could further illuminate the issue. Third, data on language and interpreting competence in our study were obtained exclusively through self-report measures. While these are widely and profitably used in the field (Hulstijn, 2012), they may be biased by self-image (e.g., social desirability) factors. Hence, future assessments should also include relevant objective measures, as suggested in the specialized literature (Hulstijn, 2012), and they should also consider the impact of the amount of training received in each interpreting direction. Fourth, our study is blind to several phenomena constraining bilingual lexical processing, such as cross-language (e.g., interference) effects (Roelofs, Piai, Garrido Rodriguez & Chwilla, 2016). In principle, our strict counterbalancing of tasks, conditions, languages (and even phonemes in the phonological fluency tasks) suggests that the present results cannot be attributed to disproportionate interferences between languages for each group due to aspects of the study's design. However, it may be the case that asymmetrical patterns of interactions between the L1 and L2 (Heikoop et al., 2016; Slevca, Daveya & Linck, 2016) influenced the observed language effects (which consistently yielded better performance in L1). Further research would be needed to ascertain the specific role of potential cross-linguistic (including interference) effects during lexical processing in PSIs. Also, given that our research design precludes causal interpretations, it would be useful to replicate the study with a longitudinal approach and thus establish which of the observed differences are directly triggered by interpreting expertise. Finally, our behavioral approximation could be refined by adding neuroscientific measurements, as done in research tapping other aspects of interpreting expertise (Elmer, Hänggi & Jäncke, 2014; Elmer & Kuhnis, 2016; Elmer et al., 2010). More generally, the notion that extreme demands on lexical systems may selectively modulate a subset of their mechanisms could be extrapolated as a working hypothesis for other models of expert language use, including stenographers, choppers (fast-paced rappers), and verse improvisers (e.g., Basque *bertsolari* or Argentine *payadores*). Efforts in these directions could open new avenues of development for bilingual memory research and related fields.

Conclusion

This work extends available evidence on the organization and flexibility of bilingual memory by showing its susceptibility to sustained extreme demands. By comparing the lexical skills of PSIs and NIBs, we found that vocabulary search and translation (as opposed to reading) mechanisms within the system can selectively adapt to meet the exacting conditions characterizing the former group. Our results support hierarchical models that incorporate functionally autonomous routes for interlinguistic processes as well as distinct processing levels for semantic and sublexical information. Further research on lexical processing in PSIs could offer additional valuable insights into the architecture of bilingual memory and its capacity for usage-driven reconfiguration.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1366728918000378

References

Abutalebi, J., Cappa, S. F., & Perani, D. (2005). What can functional neuroimaging tell us about the bilingual brain. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of* *bilingualism: Psycholinguistic approaches,* pp. 497–515. New York.

Amoruso, L., Ibáñez, A., Fonseca, B., Gadea, S., Sedeno, L., Sigman, M., García, A. M., Fraiman, R., & Fraiman, D. (2016). Variability in functional brain networks predicts expertise during action observation. *Neuroimage*. doi: 10.1016/j.neuroimage.2016.09.041

Amoruso, L., Sedeno, L., Huepe, D., Tomio, A., Kamienkowski, J., Hurtado, E., Cardona, J. F., Alvarez Gonzalez, M. A., Rieznik, A., Sigman, M., Manes, F., & Ibáñez, A. (2014). Time to Tango: expertise and contextual anticipation during action observation. *Neuroimage, 98,* 366–385. doi: 10.1016/j.neuroimage.2014.05.005

Anderson, L. (1994). Simultaneous interpretation: Contextual and translational aspects. In S. Lambert & B. Moser-Mercer (eds.), *Bridging the Gap: Empirical Research in Simultaneous Interpretation,* pp. 101–120. Amsterdam: John Benjamins.

Babcock, L., & Vallesi, A. (2015). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Bilingualism: Language and Cognition,* 1–15. doi: 10.1017/S1366728915000735

Bajo, M. T., Padilla, F., & Padilla, P. (2000). Comprehension processes in simultaneous interpreting. In A. Chesterman, N. Gallardo San Salvador, & Y. Gambier (eds.), *Translation in Context,* pp. 127–142. Amsterdam: John Benjamins.

Barik, H. C. (1975). Simultaneous interpretation: Qualitative and linguistic data. *Language and Speech, 18,* 272–298.

Benton, A. L., & Hamsher, K. (1989). *Multilingual Aphasia Examination* (Second ed.). Iowa: Universiy of Iowa.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12,* 3–11.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology:Learning, Memory, and Cognition, 34,* 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *NeuroImage, 24,* 40–49.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging, 19,* 290.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts—Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56,* 73–102. doi: https://doi.org/10.1016/j.cogpsych.2007.02.001

Borius, P. Y., Giussani, C., Draper, L., & Roux, F. E. (2012). Sentence translation in proficient bilinguals: a direct electrostimulation brain mapping. *Cortex, 48,* 614–622. doi: 10.1016/j.cortex.2011.01.011

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition, 13,* 359–371. doi: 10.1017/S1366728909990344

Calvo, N., Ibáñez, A., & García, A. M. (2016). The Impact of Bilingualism on Working Memory: A Null Effect on the Whole May Not Be So on the Parts. *Frontiers in Psychology, 7,* 265. doi: 10.3389/fpsyg.2016.00265

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50,* 491–511.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65,* 171–237. doi: 10.1006/jecp.1996.2356

Chee, M. W., Soon, C. S., & Lee, H. L. (2003). Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. *Journal of Cognitive Neuroscience, 15,* 85–97. doi: 10.1162/089892903321107846

Chernov, G. V. (2004). *Inference and Anticipation in Simultaneous Interpreting: A Probability-prediction Model*. Amsterdam: J. Benjamins Publishing Company.

Chincotta, D., & Underwood, G. (1998). Simultaneous interpreters and the effect of concurrent articulation on immediate memory. *Interpreting: International Journal of Research and Practice in Interpreting, 3,* 1–20.

Christoffels, I., Degroot, A., & Kroll, J. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language, 54,* 324–345. doi: 10.1016/j.jml.2005.12.004

Christoffels, I. K., De Groot, A. M. B., & Waldorp, L. J. (2003). Basic skills in a complex task: A graphical model relating memory and lexical retrieval to simultaneous interpreting. *Bilingualism: Language and Cognition, 6,* 201–211.

Davies, M. (2008a). Corpus del español: 100 millones de palabras, 1200s-1900s. Retrieved October 13, 2013 http://www.corpusdelespanol.org/

Davies, M. (2008b). The corpus of contemporary American English: 425 million words, 1990-present. Retrieved October 4, 2013 http://corpus.byu.edu/coca/

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects. *Journal of Memory and Language, 47,* 91–124. doi: https://doi.org/10.1006/jmla.2001.2840

De Groot, A. M. B., & Christoffels, I. K. (2006). Language control in bilinguals: Monolingual tasks and simultaneous interpreting. *Bilingualism: Language and Cognition, 9,* 189–201.

De Vreese, L. P., Motta, M., & Toschi, A. (1988). Compulsive and paradoxical translation behaviour in a case of presenile dementia of the Alzheimer type. *Journal of Neurolinguistics, 3,* 233–259. doi: https://doi.org/10.1016/0911-6044(88)90015-2

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197. doi: 10.1017/S1366728902003012

Dillinger, M. (1994). Comprehension during interpreting: What do interpreters know that bilinguals don't? *The Interpreters' Newsletter, 3,* 41–58.

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8,* 221–238. doi: 10.1017/S1366728905002270

Elmer, S., Hänggi, J., & Jäncke, L. (2014). Processing demands upon cognitive, linguistic, and articulatory functions promote grey matter plasticity in the adult multilingual brain: Insights from simultaneous interpreters. *Cortex, 54,* 179–189. doi: https://doi.org/10.1016/j.cortex.2014.02.014

Elmer, S., & Kuhnis, J. (2016). Functional Connectivity in the Left Dorsal Stream Facilitates Simultaneous Language Translation: An EEG Study. *Frontiers in Human Neuroscience, 10,* 60. doi: 10.3389/fnhum.2016.00060

Elmer, S., Meyer, M., & Jancke, L. (2010). Simultaneous interpreters as a model for neuronal adaptation in the domain of language processing. *Brain Research, 1317,* 147–156. doi: 10.1016/j.brainres.2009.12.052

Fabbro, F., Gran, B., & Gran, L. (1991). Hemispheric specialization for semantic and syntactic components of language in simultaneous interpreters. *Brain and Language, 41,* 1–42.

Ferré, P., Sánchez-Casas, R., & Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Language Learning, 56,* 571–608. doi: 10.1111/j.1467-9922.2006.00389.x

French, R. M., & Jacquet, M. (2004). Understanding bilingual memory: models and data. *Trends in Cognitive Sciences, 8,* 87–93. doi: 10.1016/j.tics.2003.12.011

Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience, 30,* 238–250. doi: 10.1080/23273798.2014.918630

García, A. M. (2013). Brain activity during translation: A review of the neuroimaging evidence as a testing ground for clinically-based hypotheses. *Journal of Neurolinguistics, 26,* 370–383.

García, A. M. (2014). The interpreter advantage hypothesis: Preliminary data patterns and empirically motivated questions. *Translation and Interpreting Studies, 9,* 219–238. doi: 10.1075/tis.9.2.04gar

García, A. M. (2015a). Psycholinguistic explorations of lexical translation equivalents: Thirty years of research and their implications for cognitive translatology. *Translation Spaces, 4,* 9–28.

García, A. M. (2015b). Translating with an injured brain: Neurolinguistic aspects of translation as revealed by bilinguals with cerebral lesions. *Meta: Translators' Journal, 60,* 112–134.

García, A. M., Ibáñez, A., Huepe, D., Houck, A. L., Michon, M., Lezama, C. G., Chadha, S., & Rivera-Rei, A. (2014). Word reading and translation in bilinguals: the impact of formal and informal translation expertise. *Frontiers in Psychology, 5,* 1302. doi: 10.3389/fpsyg.2014.01302

García, A. M., & Ibáñez, A. (2016). Hands typing what hands do: Action-semantic integration dynamics throughout written verb production. *Cognition, 149,* 56–66.

Gerver, D. (1975). A psychological approach to simultaneous interpretation. *Meta: International Translator's Journal, 20,* 119–128.

Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Briskin (ed.), *Translation: Applications and Research,* pp. 165–207. New York: Gardner Press.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of Spoken Language Dominance: A Multi-Lingual Naming Test (MINT) and Preliminary Norms for Young and Aging Spanish-English Bilinguals. *Bilingualism, 15,* 594–615. doi: 10.1017/s1366728911000332

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534–537. doi: http://www.nature.com/nature/journal/v423/n6939/suppinfo/nature01647_S1.html

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25,* 515–530.

Green, D. W., & Weib, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience, 29,* 499–511.

Guasch, M., Sánchez-Casas, R., Ferré, P., & García-Albea, J. E. (2008). Translation performance of beginning, intermediate and proficient Spanish-Catalan bilinguals: Effects of form and semantic relations. *The Mental Lexicon, 3,* 289–308. doi: 10.1075/ml.3.3.03gua

Heikoop, K. W., DeClerck, M., Los, S. A., & Koch, I. (2016). Dissociating language-switch costs from cue-switch costs in bilingual language switching. *Bilingualism: Language and Cognition, 19,* 921–927.

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition, 13,* 315–325. doi: 10.1017/s1366728909990010

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language, 69,* 257–276. doi: 10.1016/j.jml.2013.06.004

Hervais-Adelman, A., Moser-Mercer, B., Michel, C. M., & Golestani, N. (2014). fMRI of Simultaneous Interpretation Reveals the Neural Basis of Extreme Language Control. *Cerebral Cortex, 25,* 4727–4739. doi: 10.1093/cercor/bhu158

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition, 15,* 422–433. doi: 10.1017/S1366728911000678

Ibáñez, A. J., Macizo, P., & Bajo, M. T. (2010). Language access and language selection in professional translators. *Acta Psychologica, 135,* 257–266. doi: 10.1016/j.actpsy.2010.07.009

Klein, D., Zatorre, R. J., Chen, J. K., Milner, B., Crane, J., Belin, P., & Bouffard, M. (2006). Bilingual brain organization: a functional magnetic resonance adaptation study. *Neuroimage, 31,* 366–375. doi: 10.1016/j.neuroimage.2005.12.012

Köpke, B., & Nespoulous, J.-L. (2006). Working memory performance in expert and novice interpreters. *Interpreting, 8,* 1–23. doi: doi:10.1075/intp.8.1.02kop

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioural Neuroscience, 8,* 250. doi: 10.3389/fnbeh.2014.00250

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. *Annual Review of Linguistics, 1,* 377–394. doi: doi:10.1146/annurev-linguist-030514-124937

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research, 18,* 137–171. doi: 10.1191/0267658302sr201oa

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition, 13,* 373–381.

Langdon, H. W., Wiig, E. H., & Nielsen, N. P. (2005). Dual-Dimension Naming Speed and Language-Dominance Ratings by Bilingual Hispanic Adults. *Bilingual Research Journal, 29,* 319–336. doi: 10.1080/15235882.2005.10162838

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: a meta-analysis. *Psychonomic Bulletin and Review, 21,* 861–883. doi: 10.3758/s13423-013-0565-2

Lucas, T. H., McKhann, G. M., & Ojemann, G. A. (2004). Functional separation of languages in the bilingual brain: a comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients. *Journal of Neurosurgery, 101,* 449–457. doi: 10.3171/jns.2004.101.3.0449

Malt, B. C., Li, P., Pavelnko, A., Zh, H., & Ameel, E. (2015). Bidirectional lexical interaction in late immersed Mandarin-English bilinguals. *Jounal of Memory and Language, 82,* 86–104. doi: 10.1016/j.jml.2015.03.001

Manoiloff, L., Artstein, M., Canavoso, M. B., Fernandez, L., & Segui, J. (2010). Expanded norms for 400 experimental pictures in an Argentinean Spanish-speaking population. *Behaviour Research Methods, 42,* 452–460. doi: 10.3758/brm.42.2.452

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50,* 940–967. doi: 10.1044/1092-4388(2007/067)

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: converging evidence from eyetracking and brain imaging. *Brain and Language, 86,* 70–82.

Martin, A., Wiggs, C. L., Lalonde, F., & Mack, C. (1994). Word retrieval to letter and semantic cues: a double dissociation in normal subjects using interference tasks. *Neuropsychologia, 32,* 1487–1494.

McDowd, J., Hoffman, L., Rozek, E., Lyons, K. E., Pahwa, R., Burns, J., & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology, 25,* 210–225. doi: 10.1037/a0021531

Montrul, S., & Foote, R. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism, 18,* 274–303. doi: 10.1177/1367006912443431

Nelson, H. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12,* 313–324.

Ojemann, G. A., & Whitaker, H. A. (1978). The bilingual brain. *Archives of Neurology, 35,* 409–412.

Padilla, F., Bajo, M. T., & Macizo, P. (2005). Articulatory suppression in language interpretation: Working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition, 8,* 207–219. doi: 10.1017/S1366728905002269

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics, 9,* 133–145.

Paradis, M. (1994). Towards a neurolinguistic theory of simultaneous translation: The framework. *International Journal of Psycholinguistics, 10,* 319–335.

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Paradis, M., Goldblum, M. C., & Abidi, R. (1982). Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain and Language, 15,* 55–69. doi: https://doi.org/10.1016/0093-934X(82)90046-3

Pekkanli, I. (2012). Translation and the Contemporary Language Teacher. *Procedia - Social and Behavioral Sciences, 46,* 955–959. doi: https://doi.org/10.1016/j.sbspro.2012.05.230

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (1998). The bilingual brain Proficiency and age of acquisition of the second language. *Brain, 121,* 1841–1852.

Pillai, J. J., Araque, J. M., Allison, J. D., Sethuraman, S., Loring, D. W., Thiruvaiyaru, D., Ison, C. B., Balan, A., & Lavin, T. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: preliminary findings. *Neuroimage, 19,* 565–576.

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 113,* 535–551. doi: 10.1016/j.jecp.2012.06.013

Price, C. J., Green, D. W., & von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain, 122,* 2221–2235.

Prior, A., & Macwhinney, B. (2009). A bilingual advantage in task switching. *Bilingualism: Language and Cognition, 13,* 253–262. doi: 10.1017/S1366728909990526

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: an ERP study. *Neuropsychologia, 42,* 1636–1656.

Rapport, R. L., Tan, C. T., & Whitaker, H. A. (1983). Language function and dysfunction among Chinese- and English-speaking polyglots: cortical stimulation, Wada testing, and clinical studies. *Brain and Language, 18,* 342–366.

Rodriguez-Fornells, A., De Diego Balaguer, R., & Munte, T. F. (2006). Executive control in bilingual language processing. *Language Learning, 56,* 133–190.

Roelofs, A., Piai, V., Garrido Rodriguez, G., & Chwilla, D. J. (2016). Electrophysiology of cross-language interference and facilitation in picture naming. *Cortex, 76,* 1–16. doi: 10.1016/j.cortex.2015.12.003

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*(*2*), 231–252. doi: 10.1017/S1366728909990514

Sasanuma, S., Sakuma, N., & Kitano, K. (1992). Reading kanji without semantics: Evidence from a longitudinal study of dementia. *Cognitive Neuropsychology, 9,* 465–486. doi: 10.1080/02643299208252068

Signorelli, T. M., Haarmann, H. J., & Obler, L. K. (2011). Working memory in simultaneous interpreters: Effects of task and age. *International Journal of Bilingualism, 16,* 198–212. doi: 10.1177/1367006911403200

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51,* 381–398. doi: https://doi.org/10.1016/j.jml.2004.05.003

Slevca, L. R., Daveya, N. S., & Linck, J. A. (2016). A new look at "the hard problem" of bilingual lexical access: evidence for language-switch costs with univalent stimuli. *Journal of Cognitive Psychology, 28,* 385–395.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests: Administration, norms and commentary*. New York: Oxford University Press.

Stavrakaki, S., Megari, K., Kosmidis, M. H., Apostolidou, M., & Takou, E. (2012). Working memory and verbal fluency in simultaneous interpreters. *Journal of Clinical and Experimental Neuropsychology, 34,* 624–633. doi: 10.1080/13803395.2012.667068

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An Investigation of Lexical Form, Meaning, and Grammatical Class. *Studies in Second Language Acquisition, 28,* 387–422. doi: 10.1017/S0272263106060177

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition, 2,* 45–58.

Teichmann, M., Turc, G., Nogues, M., Ferrieux, S., & Dubois, B. (2012). A mental lexicon without semantics. *Neurology, 79,* 606–607. doi: 10.1212/WNL.0b013e3182635749

Torralva, T., Roca, M., Gleichgerrcht, E., Lopez, P., & Manes, F. (2009). INECO Frontal Screening (IFS): a brief, sensitive, and specific tool to assess executive functions in dementia. *Journal of the International Neuropsychological Society, 15,* 777–786. doi: 10.1017/s1355617709990415

Tzou, Y.-Z., Eslami, Z. R., Chen, H.-C., & Vaid, J. (2011). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: evidence from Mandarin–English speakers. *International Journal of Bilingualism, 16,* 213–227. doi: 10.1177/1367006911403197

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language, 39,* 458–483. doi: https://doi.org/10.1006/jmla.1998.2584

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology, 2,* 309. doi: 10.3389/fpsyg.2011.00309

Yudes, C., Macizo, P., Morales, L., & Bajo, M. T. (2013). Comprehension and error monitoring in simultaneous interpreters. *Applied Psycholinguistics, 34,* 1039–1057. doi: 10.1017/S0142716412000112