Title:

Testing bilingual educational methods: A plea to end the language-mixing taboo

Running head:

Language mixing at school

Authors:

Eneko Antón[1], Guillaume Thierry[2], Alexander Goborov[3], Jon Anasagasti[4]

and Jon Andoni Duñabeitia[1]

Affiliations:

[1] BCBL. Basque Center on Cognition, Brain and Language; Donostia (Spain).

[2] School of Psychology, Bangor University; Gwynedd Wales (United Kingdom).

[3] Moscow Innovative Language Centre; Moscow (Russia).

[4] Euskal Herriko Unibertsitatea – Universidad del País Vasco; Donostia (Spain).

Contact details:

Jon Andoni Duñabeitia

Basque Center on Cognition, Brain and Language (BCBL)

Paseo Mikeletegi 69, 2

20009 Donostia, SPAIN

j.dunabeitia@bcbl.eu

+34 943 309 300 (ext. 208)

Abstract

Language mixing in a given class is often avoided in bilingual education because of the generally held belief that *one subject* should be taught in only *one language* and *one person* should stick to *one language* in order to minimize confusion. Here, we compared the effects of mixing two languages and monolingual functioning on memory performance in immediate recall as a proxy for comprehension and attention during learning. In Experiment 1, non-balanced bilingual youngsters were provided with definition pairs introducing familiar objects in a single-language or in a mixed-language context. After each definition block, participants were asked to identify previously introduced objects presented amongst a stream of Old and New items. In Experiment 2, the same speaker produced the two definitions in the mixed-language context, thus violating the second principle introduced above. In both experiments we found no advantage for the single-language over the mixed-language context of exposure.

***Keywords**:* Bilingualism; Bilingual education; Code-switching; Language mixing.

Author Note: This research has been partially funded by grants PSI2012-32123 and SEV-2015-0490 from the Spanish Government, PI2015-1-27 from the Basque Government, ERC-AdG-295362 from the European Research Council, and by the AThEME project funded by the European Union (grant number 613465).

Introduction

Bilingual functioning - and more specifically bilingual schooling - have become a key phenomenon in regions of the world where two official languages co-exist in daily life and in places that have one official language and an increasing number of speakers of another –non-official– language (or several others). In these regions, bilinguals need to switch between languages in order to adjust to the demands of the communicative context in which they find themselves and of their interlocutor(s) (Auer, 2013; Gollan & Ferreira, 2009; Milroy & Muysken, 1995). Switching can be observed between sentences (i.e., inter-sentential code-switching), once within a given sentence (intra-sentential switching), or several times within the same sentence (e.g., by including tag phrases or words from one language in a sentence produced in the other language; Poplack, 1978/1981). Even though bilinguals often code-switch in everyday life, such practice is mostly avoided in circumstances where effective transmission and acquisition of information is required. Indeed, formal schooling mostly adopts the strategy of replacing natural bilingual interaction with segregated monolingual ones (i.e., they apply a *one subject-one language* rule). This rule, initially articulated by Lambert and Tucker (1972), advocates that academic subjects are taught consistently in one language throughout the academic year. Also, considering that a given academic subject is generally taught by one instructor, the *one subject-one language* rule generally becomes *de facto* a *one person-one language* rule: a teacher avoids mixing languages when teaching a particular subject, and s/he addresses students in the same language throughout the whole teaching experience, regardless of the subject (note that this rule resembles the

situation often encountered in relation to language learning in the context of bilingual parenting, the *one parent-one language* rule, first enunciated by Ronjat (1913) and applied with his own son).

This idea of keeping languages separate in bilingual education, also known as the *two solitudes* (Cummins, 2005; 2008), has been widely accepted as an appropriate methodology in bilingual schooling for several decades (see Lin, 2013; for a review). These rules originally applied to the case of children learning two languages simultaneously, to prevent cross-contamination, which could take the form of grammar mixing or an overuse of code switching (Jacobson & Faltis, 1990). However, in recent years, more permeability has been advocated during instruction, even of languages, considering that this may be emulating real conditions of language use more closely (Anderson, 2008; Lin & Martin, 2005; Cummins, 2005). Indeed, a systematic bilingual methodology has even been proposed in the case of language teaching (e.g., *translanguaging*[FOOTNOTE1]).

However, little is known about the potential side-effects of language-mixing on concept acquisition and manipulation in a context of formal schooling. Should learning be hindered when pupils have to learn mathematics in a mixed-language context? Would they get confused if a teacher was to describe historical events alternating between two known languages?

It should not be overlooked that language switching is cognitively demanding and that it requires additional resources and processes as compared to monolingual functioning. This makes language-switching a double-edged sword: One the one hand,

it is more ecological and makes greater use of the learner's linguistic resources, but on the other hand, it is likely to induce a cognitive cost in both perception and production (see Gullifer, Kroll, & Dussias, 2013, for a review). When bilinguals are asked to switch between languages while speaking, each language switch is accompanied by an increase in the time required to start speaking and by greater chances of making errors as compared to the non-switching conditions (see, among many others, Costa & Santesteban, 2004; see also Branzi, Martin, Abutalebi, & Costa; 2014; Verhoef, Roelofs, & Chwilla, 2009, for electrophysiological and neuroimaging data). Similarly, when bilinguals are presented with written input that involves language switches, they typically display longer recognition times and larger electrophysiological effects for switch as compared to non-switch trials (e.g., Macizo, Bajo, & Paolieri, 2012; Van der Meij, Cuetos, Carreiras, & Barber, 2011), even when switches are not perceived consciously, as is the case in masked priming experiment (e.g., Casaponsa, Carreiras, & Duñabeitia, 2015; Chauncey et al., 2008; Duñabeitia et al., 2010).

One could thus intuitively conclude that learning in a mixed-language context likely involves additional cognitive load as compared to learning in a monolingual context. According to the Cognitive Load Theory (see Sweller, 1988, 1989; see also Merriënboer & Sweller, 2005, for a recent review), three different kinds of cognitive loads can be distinguished: intrinsic, extraneous, and germane. The intrinsic cognitive load refers directly to the difficulty inherent to the task (e.g., calculating 548 + 975 is a task with a greater intrinsic cognitive load than 2 + 2). The extraneous cognitive load relates to the conditions in which information is transmitted (e.g., explaining the concept of a "triangle" by showing different pictures of triangles reduces the

extraneous load as compared to a series of definitions). The germane cognitive load results from the processing, construction, and automatization of new schemas (e.g., when students face complex geometry calculations for the first time, a step-by-step example showing how to use a previously explained formula in practice can facilitate the self-explanation and thus the acquisition of schemas to cope with that new procedure). While intrinsic cognitive load is brought about by the task itself and is thus difficult to manipulate by the vehicular language(s), both extraneous and germane cognitive loads can and should be adjusted to optimize learning. Recent findings suggest that keeping the extraneous cognitive load to a minimum is best to achieve efficient learning and avoid cognitive overload, by controlling the learning process, adjusting the task demands, and making them adequate, working with examples, using different sources of information, and reducing redundancy (see Merriënboer & Sweller, 2010). Similarly, the germane load should be optimized by increasing the variability of the tasks and encouraging self-explanation.

According to this view, mixing languages could hypothetically cause cognitive overload in bilinguals due to the increase in extraneous cognitive load it may imply. It could make learning more effortful, since receiving information in a mixed-language context could slow down the process of decoding and understanding, as compared to presentation of information in a monolingual context. However, what is the evidence supporting the assumption that language mixing is harmful to learning? Little is known to date about the impact that mixing languages may have on concept acquisition or learning, over and above the well-known switch cost described above. Inasmuch as this assumption may seem coherent, the intuitively adopted educational

'rules' to avoid language mixing during learning have not yet received any scientific validation. Indeed, to the best of our knowledge, no study has shown that learning achieved in a non-mixing context is better established, more stable in memory, or in any way superior or preferable as compared to knowledge acquired in a mixed-language context.

Quite on the contrary, some studies support the idea that gaining understanding and knowledge is potentiated through the use of two languages in a variety of ways, but beneficial effects of language mixing in academic contexts are mainly the finding of studies based on classroom observation and/or informal qualitative reports that describe the effects of the use of multiple language usage in different formal schooling settings (Baker, 2011; García, 2009; García & Wei, 2014). Even though the quantitative experimental data supporting this phenomenon is very scarce, the limited scientific evidence gathered so far on this matter suggests that mixing languages does not negatively affect concept acquisition nor learning performance. For instance, Baker and colleagues (Baker et al., 2012) investigated how Spanish-speaking English learners from grades 1 to 3 differed in their English reading achievement depending on the reading teaching methods. They investigated the reading achievement in these learners as a function of the number of languages used during the teaching, by comparing pupils set in a single-language (English-only) program and those set in a mixed-language (bilingual) program. The authors found that participants following the mixed-language bilingual approach showed highly similar reading achievement to those of participants in the single-language group. Interestingly, the differences between groups (if any) favored the mixed-language

program. Therefore, L1 assistance during L2 learning was not detrimental to learning and there were no significant differences with classic L2 learning methods. In fact, if there was a trend, it was in favor of language mixing methods.

While the study by Baker et al. (2012) is relevant, there are a number of limitations that preclude generalization from the use of language-mixing methods in bilingual school systems. The question of L1 use when learning to read in the L2 is very different from that of the use of two languages when teaching any (language-independent) subject. That particular study was aimed at assessing the impact of multilingual sources of information on literacy, rather than the acquisition and consolidation of concepts. The authors studied reading acquisition in monolingual and multilingual literacy, a very specific learning outcome that establishes the foundation for future concept acquisition. Hence, what still remains to be seen and understood is whether the use of two languages (vs. one) leads to reduced learning of concepts.

In a recent study, Antón, Thierry and Duñabeitia (2015) provided experimental data showing that mixing languages during learning did not measurably affect the outcome in a concept acquisition task. Antón et al. tested highly proficient balanced Basque-Spanish bilinguals and looked at the consequences of breaking the *one subject-one language* rule. Adult participants (Experiment 1) and children (Experiment 2) were set in an experimental context in which they were asked to learn new concepts represented by pictures of unfamiliar tools whose meaning was provided by two written definitions. Critically, these definitions could be provided either in a single language (the single-language context, SLC), or one in Basque and the

other one in Spanish (the mixed-language context, MLC). Half of the adults and children were randomly assigned to the SLC learning, and the other half were assigned to the MLC learning. Learning consolidation was assessed via direct and indirect measures of concept acquisition and integration, and results showed that MLC learners did not perform worse than SLC learners on any of the measurements. This led to the conclusion that there are no clear-cut benefits in preventing bilinguals from language mixing while learning.

The absence of any differences in the acquisition of new concepts and in their recall between the learners immersed in the two different learning contexts could be explained by the automatic and effortless language co-activation processes that balanced bilinguals boast. The automatic activation of translation representations in another language takes place in a few tens of milliseconds in the case of balanced bilinguals, presumably because both languages are always active even in a monolingual context (see Lagrou, Hartsuiker, & Duyck, 2013; Midgley, Holcomb, van Heuven, & Grainger, 2008; Spalek, Hoshino, Wu, Damian, & Thierry, 2014; Thierry & Wu, 2007; Wu, Cristino, Leek, & Thierry, 2013; Wu & Thierry, 2010, 2012). In fact, recent findings suggest that balanced simultaneous bilinguals access translation equivalents in their two languages with minimal cost (Duñabeitia, Dimitropoulou, Uribe-Etxebarria, Laka, & Carreiras, 2010; Duñabeitia, Perea, & Carreiras, 2010). This probably explains –at least in part– the lack of differences observed in the study by Antón et al. (2015), since bilinguals could have spontaneously and unconsciously accessed the translations of the multilingual input, thus overcoming the costs that are inherent to a switching context. The data from the study by Antón et al. are also in line

with the theoretical accounts indicating that, at sufficiently high levels of proficiency, bilingual individuals access language-independent semantic representations efficiently regardless of the input language (see Kroll, Michael, Tokowicz, & Dufour, 2002; Kroll, Van Hell, Tokowicz, & Green, 2010). If semantic access takes place in such automatic ways, bilinguals could find it easier to deal with the additional cognitive load imposed by the use of two languages.

However, it should be acknowledged that the study by Antón et al. (2015) also suffers from some limitations that preclude any generalization of the results to the overall educational circumstances in bilingual schools. First, population-wise, they gathered evidence from a relatively rare type of bilinguals. All the participants tested were perfectly balanced and simultaneous bilinguals, and exposed to two (both official) languages every day in their society as well as at school. That situation is certainly specific to few purely bilingual communities (e.g., the Basque Country, Catalonia, or Wales), and it does not reflect the reality of the majority of bilingual communities in the world, where one language usually prevails over the other(s), despite their coexistence in bilingual schools (see, for example, Baker, 2011; García, 2009; García & Wei, 2014). It is plausible that no differences were found in Antón et al.'s study because the participants tested were balanced bilinguals that were used to manage their two languages on a daily basis. Unbalanced bilinguals might thus suffer a measurable impact from mixing languages. Therefore, the same hypothesis needs to be tested in different and more representative populations. Second, Antón et al.'s participants were presented with the definitions of the newly learnt objects in a written format, and this clearly deviates from the common practice in the educational

context where print is mainly used to reinforce and facilitate the transmission of information which is, in most cases, initially provided in the spoken form.

In the current study we tested non-balanced Russian-English bilingual minors immersed in a Russian-English bilingual schooling system from two different schools in Moscow. This allowed us to explore the consequences of language mixing in a bilingual setting where the two languages are distant from each other (Russian and English differ at phonological, orthographic, lexical, and syntactic levels), and where the second language (i.e., English) is not present in the society as an official tongue. The participants in the current study, despite being highly proficient in their second language, were not balanced in their mastery of the two languages, and are thus much more representative of main stream bilinguals worldwide. We thus asked whether the results obtained by Antón et al. (2015) would be replicated under such a change of conditions or, on the contrary, whether significant evidence in favor of a non-mixing context would arise.

Also, it is worth keeping in mind that the *one subject-one language* and the *one person-one language* rules commonly co-exist in the formal schooling systems, but they may play different roles in concept acquisition and memory recall. As explained above, the *one subject-one language* rule was tested in Antón et al.'s (2015) work, and despite the abovementioned limitations, the data showed no harmful effects of language mixing. However, the effects of language switching within the same speaker (i.e., the violation of the *one person-one language* rule) and its effect in the listener during concept acquisition have never been tested with scientific methods, despite

experiments showing switch cost effects in perception (e.g., Grainger & Beauvillain, 1987; Soares & Grosjean, 1984;Thomas & Allport, 2000).

For that reason, the current study aimed at exploring the consequences of breaking only the *one subject-one language* rule (Experiment 1), or both the *one person-one language* and the *one subject-one language* rules simultaneously (Experiment 2). This allowed us to explore the potentially detrimental effects of mixing languages and their possible conflation with speaker consistency. Participants in both experiments were exposed to spoken definitions of familiar objects that were known to them. These definitions provided information about the concepts either in a monolingual or in a mixed-language fashion, and participants had to comprehend the pieces of information in order to access the corresponding semantic representations (i.e., the concepts being defined), and to retain them in memory. Afterwards, participants' memory was tested by visually presenting them with pictures of different exemplars of the objects that had been previously defined in the exposure phase presented amongst other untrained everyday objects (namely, an Old/New judgment paradigm). In a nutshell, in the current study we tested bilingual youngsters' ability to understand, generalize and recall the information presented auditorily either in a single-language or in a mixed-language context. To do so, we explored whether or not language mixing during information exposure has an impact on immediate memory recall.

Experiment 1

Methods

Participants. 30 participants taken from two schools in Moscow were tested in this experiment. One participant was excluded from the final analysis due to technical issues arising during the data collection, and the final analysis was performed on the data from the remaining 29 participants (mean age: 14.38 years; 17 females).

Linguistic background data from all the participants were collected by asking their parents to complete a questionnaire, who also provided written informed consent. All of the participants had Russian as their native language, and they had been exposed to English as a second language since they were, on average, 4.93 years old (SD=1.69). The parents rated the participants' proficiency in Russian on a 1-to-10 Likert-like scale with a mean rating of 10, and their English proficiency (based on teachers' feedback, who were Russian-English bilinguals) with a mean rating of 8.34 (SD=0.48). All of the participants were exposed to both languages exactly in a 50%-50% basis in their school, receiving half of the tuition in Russian and half in English. Thus, even though participants were highly proficient and early exposed to English, they were not balanced bilinguals, showing better Russian than English skills.

Materials. 28 real and well-known objects were selected, and defined by two features. For example, the object "backpack" was defined with the two features "you carry it on your back" and "you put books in it". All the definitions, written in Russian by a native Russian speaker, were rated by 20 native Russian speakers for their informativeness to see if they represented properly the real objects they were supposed to define. The definitions got a mean rating of 4.09 out of 5 (SD=0.44), meaning that they all were very informative. Then, to create the different exposure contexts (see Procedure for

further details), 14 definition pairs were kept in Russian (to be learned in a single-language context) while one of the definitions of the other 14 pairs were translated to English (to be learned in a mixed-language context). Informativeness ratings of the definitions in each language of the mixed-language context were also compared. English definitions had a mean informativeness rate of 4.11 (SD= 0.38), while Russian definitions had a mean informativeness rate of 4.04 (SD= 0.47), being the difference statistically not significant ($p>.76$). All the definitions were recorded by two balanced Russian-English bilingual females, and they were individually put in synchrony with speaking cartoon-like avatars. A single video clip was created for each definition, making sure that the same voice was always paired with the same avatar (i.e., avoiding person-language mixes). These video clips with speaking cartoon-like avatars were used to provide the participants with the definitions about the objects. The objects were divided in two groups according to the number of languages used in the definitions. In each object from the set of the single-language context (SLC), a different Russian-speaking avatar was assigned to each of the two definitions. In contrast, in each object from the set of the mixed-language-context (MLC), a Russian-speaking avatar provided one of the definitions and an English-speaking avatar provided the other definition. Importantly, the same avatar was never paired with different languages in the MCL, respecting the *one person-one language* rule not only within item, but also across the entire experiment (i.e., holding the person-language assignment constant).

For the Old-New tasks that were used to test memory recall and that took place after the corresponding exposure phases, images of two different exemplars of

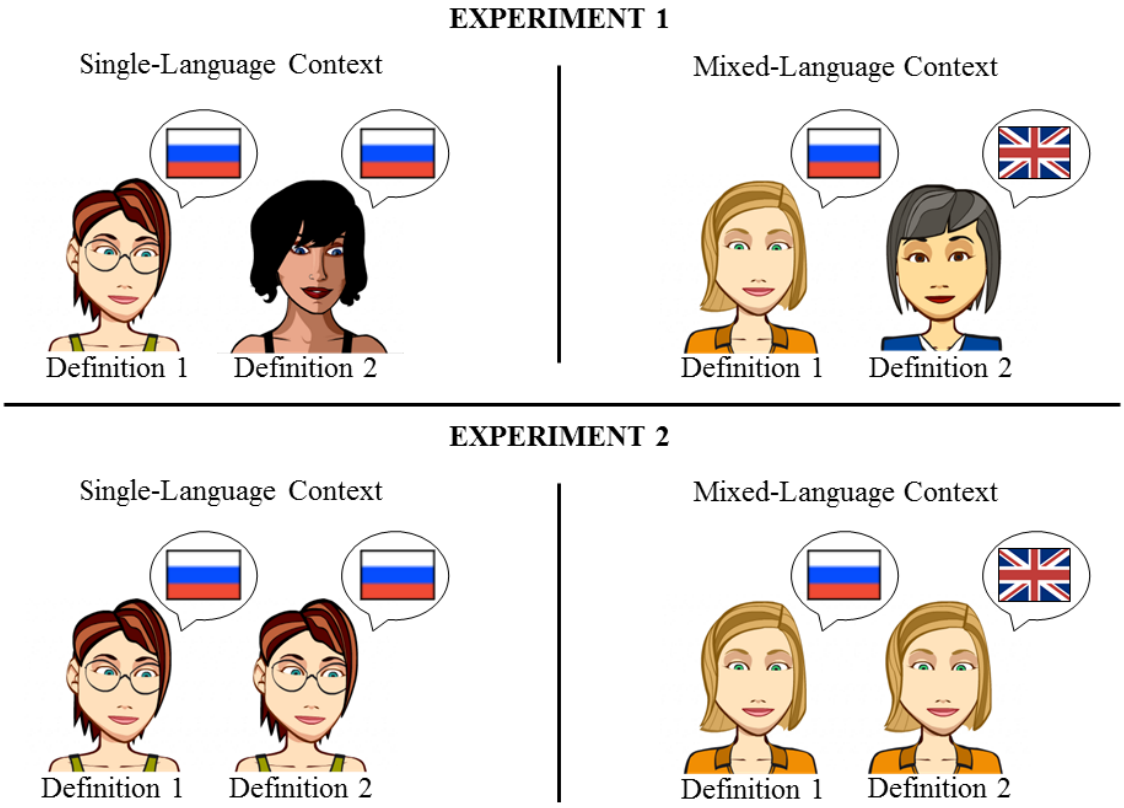
each of the objects described during exposure (e.g., two different images of a backpack) were selected from publicly available image databanks and included as "Old" items. Each of the "Old" items was paired with an unrelated real object, and their pictures were used as representative of the "New" items (see Figure 1 for an example of an "Old" and a "New" item). In order to avoid potential differences between the items used in the "New" and "Old" sets, the word frequencies of the English words corresponding to the objects were matched according to the existing corpora ($p>.68$; Davis, 2005).

Figure 1: Examples of the pictures used in the Old-New task. On the left, the two different pictures associated with the concept of backpack. This concept was described by two definitions in the exposure phase, and thus belonged to the "Old" condition. On the right, two pictures representing the concept of glove, which was not introduced in the exposure phase, making it a "New" item.

Procedure. The experiment was conducted over two days, and participants completed one session (single-language or mixed-language) each day, following a counterbalanced order (see Figure 2 for a schematic representation of the two sessions). The experiment was conducted within the dependencies of the students' schools during tuition hours. In both sessions, the procedure was identical, and

participants first completed in groups of 4-5 a computerized exposure phase in which the items were presented on a large screen associated with loudspeakers, and immediately after this, they individually completed the Old-New experiment individually on a PC. The exposure phase was organized as follows: participants were first informed about the structure of the session, and they were told that each trial of the exposure phase would start with a green square on the screen with a tone that would indicate that the two definitions of a given object would start. Automatically after this, the first avatar would produce one definition and when it finished, the second avatar would produce the second definition. This procedure was consecutively repeated twice for each item, and immediately after the last definition was played, a red square would appear on the screen to let participants know that the trial was over. After each trial was administered, the experimenter asked the participants to silently think about the object the avatars were referring to. In the SLC exposure phase, the same pair of two different avatars presented each item in Russian. In the MLC exposure phase, there were also two different avatars uttering the definitions in each trial, but each avatar provided the definitions in a different language (Russian or English) throughout the entire phase.

Figure 2: Example of the distribution of speakers and languages for a given item with its two definitions in each exposure context in Experiment 1 (upper panel) and in Experiment 2 (lower panel).



Immediately after the last trial in each exposure phase, participants were assigned a PC and performed an Old-New recognition task. Participants were asked to respond whether each of the images displayed on the screen corresponded, or not, to an object previously defined by the two avatars. (Note at this regard that participants had not been exposed to any visual image during the exposure phase). The images were individually presented centered on the screen, after a fixation cross that appeared on the screen for 500 milliseconds, and each target remained on the screen until a response was given. Two different keys from the keyboard were used to collect responses. Each version of each target was presented once during the experiment, and two different pictures represented each individual object from the exposure phase. This resulted in a set of 28 “Old” items and 28 “New” items. Trial order was randomly

varied across participants. The experiment was designed and administered using Experiment Builder© (SR Research, Ontario, Canada).

Results

We firstly analyzed the discrimination capacities of the youngsters in the two contexts according to the Signal Detection Theory, in line with preceding studies on memory recall (e.g., Nevin, 1969; Macmillan & Creelman, 2004). We calculated participants' discriminability indices (*d'*) as a function of a composite measure that takes into account the false alarm rates (i.e., the proportion of "Old" responses to "New" items) and the hit rates (i.e., the proportion of "Old" responses to "Old" items; see Table 1).

Participants in both contexts showed markedly high discriminability skills, demonstrating that they effectively recalled the "Old" items, while correctly rejecting the "New" items (i.e., successful recall; SLC: *d'*=3.49, SD=.49; MCL: *d'*=3.49, SD=.64). The *d'* of the participants in each of the two contexts were compared, and results indicated that the difference was not significant [$t(28)=.03$, $p>.98$, Cohen's d= -.01]. A Bayesian Null Hypothesis Testing was also carried out, and the Bayes Factor (BF_{01}) was computed (see Rouder et al., 2009; Wetzels et al., 2011). Results showed that the null hypothesis was 5 times more probable than the alternative, with a BF_{01} of 5.06.

In order to further explore whether potential differences could have emerged between the two contexts, response latencies were also analyzed. The latency of the responses to the "Old" and to the "New" items in the two contexts were analyzed following a 2*2 design including the factors Type (Old vs. New) and Context (single-

language vs. mixed-language context). Prior to the analysis, reaction times were trimmed and individual data points that exceeded the mean per condition and participant in more than 2.5 standard deviations were removed (3.75% of the data). Mean reaction times in each condition and in each context are presented in Table 1.

The ANOVA on the reaction times showed that the "Old" items were responded to faster than the "New" items [main Type effect: $F(1,28)=20.38$, $p<.01$], but there was no effect of Context [$F(1,28)=1.66$, $p>.21$, Cohen's $d=-.22$], nor an interaction between these two factors [$F(1,28)=2.73$, $p>.11$, Cohen's $d=-.39$]. Although these data suggest that there were no significant differences between the SLC and MLC blocks, we decided to further explore this hypothesis by computing the Bayes Factors (BF_{01}) by running a paired-sample Bayesian t-test analysis on the Old-New effects (i.e., the difference in RTs in the responses to the "New" and the "Old" conditions). Results showed that the null hypothesis was much more likely to explain these data than the alternative hypothesis, with a BF_{01} of 12.17 (namely, the null was 12 times more probable than the alternative).

Table 1: Mean reaction times (in milliseconds), percentages of hits (i.e., "Old" responses to "Old" items), percentages of false alarms (i.e., "Old" responses to "New" items) and *d'* values across contexts in Experiment 1. Standard deviations are provided in parentheses.

	Reaction times (ms.)		Accuracy		
Context	New	Old	Hits	False Alarms	d'
Single-Language	1112 (292)	989 (213)	91.62 (5.83)	0.86 (2.46)	3.49 (0.49)
Mixed-Language	1235 (391)	1007 (297)	90.76 (10.35)	1.23 (2.57)	3.49 (0.64)

Experiment 2

Methods

Participants. A different group of 30 youngsters taken from the same two schools was tested in Experiment 2 (mean age: 13.27 years; 14 females). As in Experiment 1, linguistic background data and written informed consent were collected from the students' parents. All of them reported Russian being their native language, and participants were exposed to English as a second language from a mean age of 5.37 years (SD=1.50). Based on teachers' feedback, the parents rated the minors' Russian proficiency with a 10 out of 10, and with a mean rating of 8.27 (SD=0.58) for their English skills. As in Experiment 1, all participants were equally exposed to Russian and English (50% of the time to each language) during school hours.

Materials. The materials were the same as those used in Experiment 1.

Procedure. Overall, the procedure was the same as in Experiment 1, but in order to simultaneously violate the *one subject-one language* and the *one person-one language* rules, a critical manipulation was added. In the two exposure phases (SLC and MLC) of Experiment 1 two different avatars provided the definitions for each item, each one producing one definition. In contrast, and even if different avatars were also used across trials, in the two exposure phases of Experiment 2 the same avatar produced the two definitions of each item (see Figure 1). This allowed us to introduce an important manipulation in the MLC: while in the exposure phase of the MLC from Experiment 1 the two avatars that provided the definitions never alternated between languages (i.e., the same avatar used the same language throughout the entire phase), in the exposure phase of the MLC from Experiment 2 the same avatar switched

between languages, producing one definition in English and the other one in Russian, thus breaking the *one person-one language* rule.

Results

The same analysis routine as in Experiment 1 was followed. First of all, we computed the *d'* values for each participant in each context based on the proportion of hits and false alarms, and compared them across contexts (SLC and MLC; see Table 2). The discriminability skills of the participants in both SLC and MLC were high, suggesting good recall of learned items (SLC: *d'*=3.29, SD=.89; MCL: *d'*=3.06, SD=1.01). No differences were observed in the discrimination abilities in the different contexts [t(29)=1.48, p>.15, Cohen's d = .21]. As in the Experiment 1, a paired-sample Bayesian t-test analysis on the d' indices was run, and results showed that the null hypothesis was around 2 times more probable than the alternative, with a BF_{01} of 1.93.

As in Experiment 1, after trimming reaction times to discard outliers (4.01% of the data), response latencies were analyzed following a 2*2 design, and results were further checked using Bayesian Null Hypothesis Testing. Mean reaction times in each condition and each context are presented in Table 2. The ANOVA on the reaction time data showed a main effect of Type, demonstrating that the "Old" items were responded to faster than the "New" items [F(1,28)=20.30, p<.01]. No effect of Context (F<1, Cohen's d= .08) nor an interaction between the two main factors (F<1, Cohen's d= .04) were found. The analysis of the Old-New effect using a Bayesian t-test showed a Bayes Factor (BF_{01}) of 4.25, showing that, as in Experiment 1, the null hypothesis was four times more likely than the alternative.

Table 2: Mean reaction times (in milliseconds), percentages of hits (i.e., "Old" responses to "Old" items), percentages of false alarms (i.e., "Old" responses to "New" items) and *d'* values across contexts in Experiment 2. Standard deviations are provided in parentheses.

	Reaction times (ms.)		Accuracy		
Context	New	Old	Hit	False Alarm	d'
Single-Language	1715 (814)	1213(313)	89.52 (10.10)	3.69 (9.6)	3.29 (0.89)
Mixed-Language	1652(757)	1179(298)	86.06 (14.21)	5.41 (11.44)	3.06 (1.01)

Discussion

The deeply rooted belief that mixing languages during the transmission of information (e.g., during formal tuition) may be harmful for the listener is a scientifically ungrounded axiom that frequently governs the educational practice in bilingual schools across countries. However, and in contrast to this assumption, recent evidence from studies on language switching during concept acquisition suggest that the sequential alternation between languages does not have a dramatic impact in learning and immediate recall (see Antón et al., 2015). In the current study, the exposure and later recall of concepts in a mixed language context was compared to the same process in a single language context so as to extend the results by Antón et al. to a different language combination, a different bilingual setting that provides a harder test case for any potential harmful consequences of mixing two languages, and a different process, namely, memory consolidation. Russian native youngsters attending Russian-English bilingual schools were tested. First, English and Russian are markedly different on every linguistic level (sub-lexical, lexical and syntactic). Second,

and in contrast to the participants tested in the preceding study by Antón et al., the participants tested here are representative of non-balanced bilingual groups living in a society in which one of the languages is dominant (i.e., the official language), while the other is restricted to certain specific activities (e.g., at school). And third, in the current study, information was provided using spoken language, making it closer to formal tuition in school settings. Moreover, in the current study, the hypothetically different effects of breaking the *one subject-one language* and *one person-one language* were tested separately in two different experiments.

In Experiment 1, the *one subject-one language* rule was violated while keeping the speaker-to-language correspondence constant across trials (i.e., the same avatar always used the same language). Results showed that response latencies as well as accuracy in recall were similar in both language contexts, providing strong support against the idea that mixing languages would be more effortful and showing that alternation between languages during the introduction of concepts and their recall from memory did not have any measurable negative impact.

In Experiment 2 the same results were obtained and extended to the *one person-one language* rule. In this experiment, the same speaker (namely, the same avatar) alternated between languages in the mixed-language condition, thus violating the two "rules" at the same time. Results again demonstrated that there was no advantage in favor of the non-mixing context. Hence, the non-harmful effects of alternating languages can now be taken more reliably given that neither breaking the *one subject-one language* rule alone, nor the breaking both the *one subject-one*

language and the *one person-one language* rules lead to any impairment in object memorization as compared to the single-language-contexts.

It is worth noting that the potential differences between single-language and mixed-language exposure and recall of different items were tested here by investigating the impact of these methods within participant, thus minimizing the influence of any inter-personal extraneous factors. Except for the context, which was varied across sessions in a counterbalanced manner, everything remained constant between the experiments, and mixing languages did not result in a reduction in participants' capacity to recognize, memorize and recall the items. However, as we mentioned in the Introduction, code-switching has been repeatedly shown to induce a processing cost in language production, and critically, in language perception (see Grainger & Beauvillain, 1987; Soares & Grosjean, 1984; Thomas & Allport, 2000), so one may wonder how it is possible that the youngsters that took part in these experiments did not show detrimental effects of mixing languages, reflected in a poorer performance in the Old-New recognition tasks. In the study by Antón et al. (2015), the absence of differences between the acquisition and recall of concepts using mixed-language and single-language methodologies was explained by the relatively automatic and effortless mental translation processes that balanced bilinguals feature. In the current study, the participants were non-balanced bilinguals, and according to models of bilingual memory organization the access to semantic representations does not occur in parallel in this type of bilinguals and is typically L1-mediated (see Kroll & De Groot, 1997; Kroll et al., 2010). Hence, according to the view that the processing of L2 input requires L1-mediation through translation, it could be

tentatively proposed that the increased effort required by the processing of information in different languages may compensate for the inherent difficulty associated with processing linguistic material in a non-dominant language, ultimately yielding equivalent results across language contexts. The cognitively costlier or slower decoding process linked to mixed-language contexts might have caused the information to be better internalized and established in memory, possibly overcoming long-lasting detrimental effects of code switching.

Furthermore, according to the Cognitive Load Theory, working memory can be effectively expanded using certain methods that involve multiple sources of information. For example, when the information is provided via different modalities (e.g., partly visual and partly auditory), its acquisition is more effective than in a unimodal setting (e.g., only visual or only auditory; see Frick, 1984; Mousavi, Low, & Sweller, 1995). We suggest that when different languages are used to introduce or transmit information, they behave as different information channels, boosting perceivers' working memory and potentially helping knowledge consolidation in the long run, over and above the immediate cost elicited by language switching. According to this view, language mixing should not be considered as an increased extraneous cognitive load that overburdens learners' capacities, but rather as germane cognitive load, which likely supports information acquisition. Furthermore, it is worth noting that in the current experiments the information provided by the two languages was complementary (i.e., both definitions related to the concept), which is a *sine qua non* condition for the memory-boosting modality effect to pop up (e.g., Leahy, Chandler, & Sweller, 2003).

This study represents a significant contribution to our understanding of how the concurrent use of multiple languages during exposure and comprehension (that are arguably early and necessary steps of learning) should be tested further as an ecological alternative to monolingual methods of teaching bilingual students. In this sense, the current study adds to a growing body of evidence suggesting that dynamic use of the two languages spoken by a bilingual in learning environments may result in an enrichment of the learning experience (i.e., the beneficial impact of translanguaging in the school; see Otheguy, García, & Reid, 2015, for a recent review). We focused the scope of our investigation to the recognition and immediate recall of concepts, and we tested whether memory recall would be affected by the context in which those concepts were defined and explained (single-language or mixed-language context). We acknowledge that in the current study we only tested a specific part of the learning process (memory recall) and that further studies should delve into other aspects of the construct. Nonetheless, the current study already provides important insights about the (im)permeability of concept assimilation, generalization, integration and recall processes to a bilingual exposure to information.

However, some limitations should also be acknowledged. In the current investigation, as in Antón et al. (2015), only immediate recall was tested, and consequently the processes explored mainly tap into short-term memory. Memory consolidation is a critical aspect of learning that needs further exploration, given that it better represents the reality of the educational context in which acquired knowledge is assessed at later stages of the academic year (e.g., at the interim and final examinations). Hence, future studies will have to elucidate whether language

mixing during tuition has any drawback (or boosting) effect in later retest phases. Besides, this study did not explore how a specific form of controlled translanguaging such as the one used here affects linguistic development and/or language learning, since we exclusively focused on how known concepts are consolidated in memory. We acknowledge that future studies should also explore whether mixing languages could represent any obstacle beyond the conceptual level, investigating its impact at the linguistic level in non-balanced bilinguals (e.g., the negative consequences of code-switching in the short or long term for cross-linguistic interference). Furthermore, it should also be considered that so far this research question has been exclusively approached using behavioral measures, which are interesting and important, but which do not allow for generalizations of the conclusions on their own, since a lack of behavioral differences does not necessarily imply a lack of differences at the neural level (see Duñabeitia & Carreiras, 2015, for discussion). And finally, it is worth noting that the degree of generalization of the current results and conclusions should be adequately circumscribed to the specific socio-educational contexts that have been tested so far. Nonetheless, considering the present results together with those reported by Antón et al., it seems reasonable to hypothesize that these findings are relatively stable across language combinations (e.g., Basque-Spanish, Russian-English), across age-ranges (e.g., children, adolescents and young adults), and across types of bilinguals (e.g., crib balanced bilinguals, sequential non-balanced bilinguals).

In the meantime, and until all other premises are tested, the available evidence overall points towards the abandon of the premises of the *one subject – one language* rule and of the *one person – one language* rule. The data suggest that

comparable comprehension and immediate recall can be achieved by bilingual learners whether or not languages are mixed used during tuition, with no detrimental effect of language mixing. In a situation in which no harmful effects of mixing languages are found, only the positive aspects of simultaneously using two languages remain, enabling bilingual learners to use their two languages in a naturalistic way, during learning as in everyday life. In the same way that active bilingualism involves language alternation, bilingual education should thus embrace language switching.

Footnotes

Footnote 1. According to some authors, translanguaging is a process of dynamic bilingualism based on the self-regulation of language selection and usage grounded on the "flexibility of bilingual learners to take control of their own learning" (García, 2014, p.80; see Otheguy, García, & Reid, 2015, for discussion). Other authors go further in their degree of concreteness and define translanguaging as "the planned and systematic use of two languages for teaching and learning inside the same lesson" (Lewis et al., 2013, p. 3), which would be in line with the manipulations implemented in the current study. However, considering the lack of homogeneity in the definitions of the term, we will avoid its use in the current article.

References

Anderson, J. (2008). Towards integrated second language teaching pedagogy for foreign and community/heritage languages in multilingual Britain. *Language Learning Journal, 36,* 79–89. doi:10.1080/09571730801988553.

Antón, E., Thierry, G., & Duñabeitia, J.A. (2015). Mixing languages during learning? Testing the one subject - one language rule. *PLoS ONE*, 10 (6), e0130069. doi: 10.1371/journal.pone.0130069.

Auer, P. (2013). *Code-Switching in Conversation: Language, Interaction and Identity*. Routledge.

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (Vol. 79). Multilingual matters.

Baker, D. L., Park, Y., Baker, S. K., Basaraba, D. L., Kame'enui, E. J., & Beck, C. T. (2012). Effects of a paired bilingual reading program and an English-only program on the reading performance of English learners in Grades 1–3. *Journal of School Psychology*, *50*(6), 737–758. doi: 10.1016/j.jsp.2012.09.002.

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. *Neuropsychologia*, 52, 102-116. doi: 10.1016/j.neuropsychologia.2013.09.022.

Casaponsa, A., Carreiras, M., & Duñabeitia, J.A. (2015). How do bilinguals identify the language of the word they read? *Brain Research,*1624, 153-166. doi: 10.1016/j.brainres.2015.07.035.

Chauncey, K., Grainger, J., & Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: A masked priming study with event-related potentials. *Brain and language, 105*(3), 161-174. doi: 10.1016/j.bandl.2007.11.006.

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*(4), 491–511. doi: 10.1016/j.jml.2004.02.002

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal, 89,* 585–592.

Cummins, J. (2008). Teaching for transfer: Challenging the two solitudes assumption in bilingual education. *The Encyclopedia of language and education. Vol 5 Bilingual Education.* New York: Springer, 65-75. doi: 10.1007/978-0-387-30424-3_116.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior research methods, 37*(1), 65-70. doi: 10.3758/BF03206399.

Duñabeitia, J. A., & Carreiras, M. (2015). The bilingual advantage: Acta est fabula? *Cortex, 73*, 371-372. doi: 10.1016/j.cortex.2015.06.009.

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Research*, *1359*, 142–154. doi: 10.1016/j.brainres.2010.08.066.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, *57*(2), 98–107. doi: 10.1027/1618-3169/a000013.

Frick, R. W. (1984). Using both an auditory and a visual short-term store to increase digit span. Memory & Cognition, 12(5), 507-514. doi: 10.3758/BF03198313.

García, O. (2009). Bilingual education in the 21st century: A global perspective. Malden, MA and Oxford: Basil/Blackwell.

García, O. & Li Wei. (2014). *Translanguaging: Language, Bilingualism and Education*. New York: Palgrave Macmillan. doi: 10.1057/9781137385765.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 640–665. doi: 10.1037/a0014981.

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology, 39*(2), 295-319. doi:10.1080/14640748708401788.

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: lexical access in sentence context. *Frontiers in psychology, 4*. doi: 10.3389/fpsyg.2013.00278.

Jacobson, R., & Faltis, C. (Eds.). (1990). *Language distribution issues in bilingual schooling* (Vol. 56). Multilingual matters.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. *Tutorials in bilingualism: Psycholinguistic perspectives,* 169-199.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*(2), 137–171. doi: 10.1191/0267658302sr201oa.

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*(03), 373–381. doi: 10.1017/S136672891000009X

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). Interlingual lexical competition in a spoken sentence context: Evidence from the visual world paradigm. *Psychonomic Bulletin & Review*, *20*(5), 963–972. doi: 10.3758/s13423-013-0405-4.

Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment.* Rowley: Newbury House.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology, 17*(4), 401-418. doi: 10.1002/acp.877.

Lewis, G., Jones, B., & Baker, C. (2012). Translanguaging: Origins and Development from School to Street and beyond. *Educational Research and Evaluation, 18*(7), 641-654. doi: 10.1080/13803611.2012.718488.

Lin, A. (2013). Classroom code-switching: Three decades of research. *Applied Linguistics Review, 4*(1), 195-218. doi: 10.1515/applirev-2013-0009.

Lin, A. M. Y., & Martin, P. (2005). (Eds.). *Decolonisation, globalisation: Language-in-education policy and practice*. Clevedon, UK: Multilingual Matters.

Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research, 28*(2), 131-149. doi: 10.1177/0267658311434893.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Merriënboer, J. J. G. van, & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. doi: 10.1007/s10648-005-3951-0.

Merriënboer, J. J. G. van, & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, *44*(1), 85–93. doi: 10.1111/j.1365-2923.2009.03498.x.

Midgley, K. J., Holcomb, P. J., vanHeuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research*, *1246*, 123–135. doi: 10.1016/j.brainres.2008.09.078.

Milroy, L., & Muysken, P. (1995). *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-Switching*. Cambridge University Press. doi: 10.1017/cbo9780511620867.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of educational psychology, 87*(2), 319. doi: 10.1037/0022-0663.87.2.319.

Nevin, J. A. (1969). Signal detection theory and operant behavior: A review of David M. Green and John A. Swets' Signal detection theory and psychophysics. *Journal of the Experimental Analysis of Behavior*, 12(3), 475. doi: 10.1901/jeab.1969.12-475.

Otheguy, R., García, O. & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review, 6*(3), 281-307. doi: 10.1515/applirev-2015-0014.

Poplack, S. (1978/1981). Syntactic structure and social function of code-switching. In R.P. Duran (ed.), *Latino Discourse and Communicative Behaviour.* New Jersey: Ablex Publishing Corporation, 169-184.

Ronjat, J. (1913). *Le developpement du langage observé chez un enfant bilingue.* Paris: Champion.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review, 16*(2), 225-237. doi: 10.3758/PBR.16.2.225.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition, 12*(4), 380-386. doi: 10.3758/BF03198298.

Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., & Thierry, G. (2014). Speaking two languages at once: Unconscious native word form access in second language production. *Cognition*, *133*(1), 226–231. doi: 10.1016/j.cognition.2014.06.016.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. doi:10.1016/0364-0213(88)90023-7.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, *81*(4), 457–466. doi: 10.1037/0022-0663.81.4.457.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535. doi: 10.1073/pnas.0609927104

Thomas, M. S., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43*(1), 44-66. doi: 10.1006/jmla.1999.2700

Van Der Meij, M., Cuetos, F., Carreiras, M., & Barber, H. A. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology, 48*(1), 44-54. doi: 10.1111/j.1469-8986.2010.01039.x.

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition, 110*(1), 84-99. doi: 10.1016/j.cognition.2008.10.013.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology an empirical comparison using 855 t tests. *Perspectives on Psychological Science, 6*(3), 291-298. doi: 10.1177/1745691611406923.

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(22), 7646–7651. doi: 10.1523/JNEUROSCI.1602-10.2010

Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *NeuroImage, 59*(4), 3468–3473. doi: 10.1016/j.neuroimage.2011.11.049.

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition*, *129*(2), 418–425. doi: 10.1016/j.cognition.2013.08.005.