



# Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech

Irina Pivneva<sup>1,2</sup>, Caroline Palmer<sup>1,2</sup> and Debra Titone<sup>1,2\*</sup>

<sup>1</sup> Department of Psychology, McGill University, Montreal, QC, Canada

<sup>2</sup> Centre for Research on Brain, Language, and Music, McGill University, Montreal, QC, Canada

## Edited by:

Judith F. Kroll, Penn State University, USA

## Reviewed by:

Miriam Gade, University of Zurich, Switzerland

Noriko Hoshino, Kobe City University of Foreign Studies, Japan

Gerrit Jan Kootstra, Radboud University Nijmegen, Netherlands

## \*Correspondence:

Debra Titone, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC H3A 1B1, Canada.  
e-mail: dtitone@psych.mcgill.ca

Bilingual language production requires that speakers recruit inhibitory control (IC) to optimally balance the activation of more than one linguistic system when they produce speech. Moreover, the amount of IC necessary to maintain an optimal balance is likely to vary across individuals as a function of second language (L2) proficiency and inhibitory capacity, as well as the demands of a particular communicative situation. Here, we investigate how these factors relate to bilingual language production across monologue and dialogue spontaneous speech. In these tasks, 42 English–French and French–English bilinguals produced spontaneous speech in their first language (L1) and their L2, with and without a conversational partner. Participants also completed a separate battery that assessed L2 proficiency and inhibitory capacity. The results showed that L2 vs. L1 production was generally more effortful, as was dialogue vs. monologue speech production although the clarity of what was produced was higher for dialogues vs. monologues. As well, language production effort significantly varied as a function of individual differences in L2 proficiency and inhibitory capacity. Taken together, the overall pattern of findings suggests that both increased L2 proficiency and inhibitory capacity relate to efficient language production during spontaneous monologue and dialogue speech.

**Keywords:** bilingualism, dialogue, monologue, inhibition, proficiency

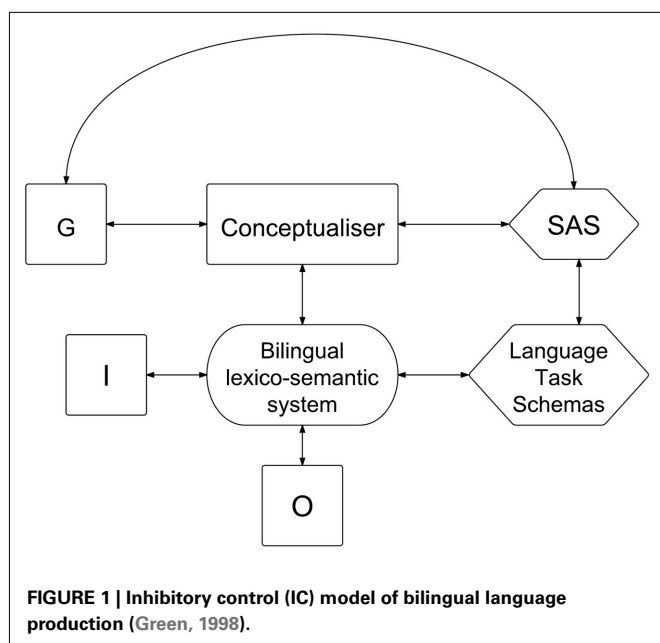
## INTRODUCTION

Speaking in one's first language (L1) is subjectively effortless, yet speech production involves a complex set of linguistic operations that require cognitive control (Kempen and Hoenkamp, 1987; Levelt, 1989). Speakers first conceptualize a message and then activate words in memory that are semantically and syntactically compatible with the message. Speakers then select from among this set the specific words that best convey the message, plan their articulation, and finally, implement the speech plan and produce their message at a rate of about 150–300 words per minute (Goldman-Eisler, 1968). These processes are incremental in that speakers transfer partially prepared fragments of the message from one stage to the next before completely preparing the message in its entirety. Thus, speakers begin articulating earlier parts of the message before fully activating and planning later parts of the message. The net effect of these cascaded and incremental speech processes is that native language production is quite cognitively demanding, in terms of word finding and word choice, grammatical and phonological realization, and overall fluency (Levelt, 1989; Dell et al., 1999; Griffin and Ferreira, 2006).

The production of fluent speech is likely to require even greater cognitive control for bilingual speakers, who face the challenges just described, as well as demands associated with knowing and using more than one language (Kroll et al., 2008; Colomé and Miozzo, 2010; De Groot, 2011). These added demands include a greater need to manage cross-language competition arising from

parallel activation of two languages (Kroll et al., 2006, 2008), less practice using inhibitory control (IC) during L2 speech production (Abutalebi and Green, 2007), and weaker links between conceptual and linguistic representations in the L2 and possibly L1 (Poulish and Bongaerts, 1994; Gollan et al., 2008). Indeed, recent work suggests that the added demands of bilingual language processing might lead to enhanced non-linguistic cognitive function for processes necessary to reduce cross-language competition, such as inhibitory capacity and selective attention (Bialystok et al., 2004; Bialystok, 2009).

In this study, we investigate how individual differences among bilinguals in L2 proficiency and inhibitory capacity modulate language production during spontaneous monologue and dialogue speech. Our theoretical framework derives from the IC model of bilingual language production, which is depicted in **Figure 1** (Green, 1998). A core assumption of this model is that language production is a communicative action that is analogous to non-linguistic physical actions (Green, 1998; Abutalebi and Green, 2007). Like physical actions, bilingual language production consists of mental task schemas, which are action sequences that are implemented by a conceptualizer (C). These task schemas achieve particular goals (G), which may be routine (L1 production) or non-routine (L2 production). For any given goal, parallel activation of multiple task schemas compete to control output (O). Consequently, the supervisory attentional system (SAS; Shallice and Burgess, 1996) suppresses routine goals via IC operations,



and monitors the successful implementation of non-routine goals, based on input from the bilingual lexico-semantic system. Accordingly, when a bilingual speaker engages in a dialogue with a monolingual speaker in their L2, the conceptualizer relays input (I) from the bilingual lexico-semantic system to the SAS, which, in turn, implements greater IC to globally suppress the irrelevant but more routine L1 dialogue language schema. As well, within the bilingual lexico-semantic system, IC fine-tunes the relative activation and inhibition of words within each language to select and output appropriate words for the dialogue.

Abutalebi and Green (2007) extended the IC model to incorporate neurocognitive evidence about bilingual language production. They identified a network of cortical regions (prefrontal, inferior parietal, and anterior cingulate cortices) and subcortical structures (basal ganglia, the head of the caudate nucleus in particular) that modulate competition between L1 and L2 knowledge activation during bilingual language production. Within this framework subcortical structures (basal ganglia) modulate the global activation of L1 or L2 task schemas, whereas frontal cortical structures modulate local activation of L1 and L2 lexical activation. Using this framework, the authors also make more specific claims about the role of L2 proficiency. When L2 proficiency is low, L2 language production is more controlled and less automatic (see also Favreau and Segalowitz, 1983; Segalowitz and Hulstijn, 2005; Segalowitz, 2010), thus requiring IC (prefrontal function, in particular; see also Petrides, 1998). In contrast, when L2 proficiency is high, L2 production is automatic and less dependent on IC, although L1 production effort might instead increase due to a weakening of the links between word forms and concepts in the L1 (Bialystok, 2001; Michael and Gollan, 2005; Gollan et al., 2008, 2011; Ivanova and Costa, 2008; Bialystok et al., 2010).

Thus, the IC model (Green, 1998) and its extension (Abutalebi and Green, 2007) make several logical predictions about the role of IC during bilingual language production: (1) L2 language

production should require greater IC than L1 production to the extent that L2 proficiency is low (and indeed, L1 language production may become more difficult as L2 proficiency increases); (2) these effects should interact with communicative task demands (i.e., a highly vs. less demanding communicative task should limit the resources available for IC to occur); and (3) bilinguals should successfully produce language insofar that they intrinsically possess IC capacity, after accounting for L2 proficiency.

Bilingual language production studies provide some support for these predictions, although many questions remain. Consistent with the first prediction, many studies show that L2 production (which is usually the less-dominant language) is indeed more effortful than L1 production (which is usually the more dominant language). This pattern of findings arises when bilinguals produce single words in response to pictures (Linck et al., 2008; Gollan and Ferreira, 2009; Hanulová et al., 2010; Sandoval et al., 2010), and also when they produce extended speech (Towell et al., 1996). Moreover, as L2 proficiency increases, language production in a less-dominant L2 improves (Poullisse and Bongaerts, 1994; Kormos, 2006; De Jong and Wempe, 2009). For example, at high L2 proficiency levels picture-naming speed and accuracy become more similar across L2 and L1 (Costa and Caramazza, 1999; Kroll et al., 2002; Costa and Santesteban, 2004). A similar pattern of effects is also seen during spontaneous speech production. For example, increased L2 proficiency is associated with increased articulation rate, longer utterance durations, shorter and less frequent silent pauses, and a greater number of words produced in the L2 when bilinguals narrated a story from a cartoon strip (Kormos and Dénes, 2004).

Increased L2 proficiency also relates to increased L1 processing effort when bilinguals produce single words in response to a picture (Gollan et al., 2005, 2008, 2011; Ivanova and Costa, 2008), overtly name visually presented words (Flege, 1999), or to general measures of functional language ability (i.e., subtractive bilinguals Lambert, 1974). Interestingly, our group recently found that these effects of increased L2 ability on L1 processing extend to eye movement measures of reading (Titone et al., 2011; Whitford and Titone, 2012). Presumably, such effects on L1 language processing arise because bilinguals who are highly proficient in their L2 use their L2 to a great extent, and as a consequence, use their L1 relatively less. Thus, over time and repeated L2 practice and use, L1 representations grow weaker while L2 representations grow stronger.

Returning to the second prediction of the IC model, there is also evidence that L1/L2 differences in language production are sensitive to increased task demands. For example, language production is more effortful during simultaneous interpretation, in which bilinguals must understand the utterance in one language and produce it in another (Christoffels and De Groot, 2004). As well, there is preliminary evidence of task demand effects for spontaneous speech when it is produced with or without a conversational partner. For example, bilinguals produce more disfluencies (e.g., uhs, ums) when answering speculation questions during a dyadic interview (e.g., What makes an ideal friend?) than when producing speech without a conversational partner (e.g., telling a story from a picture; Fehring and Fry, 2007). This suggests the possibility that a dialogue context may be relatively more

effortful than a monologue context, especially during L2 language production. This finding is interesting in light of recent work suggesting that dialogue speech can be less effortful than monologue speech because conversational partners provide additional sources of information that can facilitate speech planning, such as immediate feedback about communication success or lexical and syntactic priming across partners (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Costa et al., 2008; Hartsuiker and Pickering, 2008; Kootstra et al., 2010). While such facilitative interactive alignment effects are certainly possible, they are likely offset by other increased task demands of spontaneous dialogue, such as integrating language production and comprehension simultaneously, and making decisions about when to speak or listen, all within the time limits of normal conversational exchange (McFarland, 2001; Wilson and Wilson, 2005).

Finally, there is preliminary evidence consistent with the third prediction of the IC model that individual differences in inhibitory capacity modulate bilingual language production, over and above the effects of L2 proficiency. Linck et al. (2008) found that bilinguals with greater inhibitory capacity vs. those without, as assessed by non-linguistic tasks, inhibited L1 activation during L2 production more efficiently, irrespective of L2 immersion environment, L2 proficiency, or L1/L2 script similarity. However, given that Linck and colleagues investigated single word production, an open question is whether individual differences in inhibitory capacity exert similar effects when producing extended spontaneous speech and in different communicative contexts.

Thus, the purpose of the present study is to investigate several questions about bilingual language production in the domain of spontaneous monologue and dialogue speech. Based on the IC model (Green, 1998) and its extension (Abutalebi and Green, 2007), we predicted that L2 vs. L1 language production would be more effortful overall; however, increased L2 proficiency would reduce this difference (Poulisse and Bongaerts, 1994; Green, 1998; Gollan et al., 2005; Fehring and Fry, 2007; Ivanova and Costa, 2008; Kroll et al., 2008; Linck et al., 2008). We also predicted that dialogue speech would be more effortful than monologue speech, particularly in the L2 vs. L1 context (Fehring and Fry, 2007). Finally, we predicted that individual differences in inhibitory capacity, while accounting for L2 proficiency, would interact with the language produced (L1 vs. L2) and task demands (monologue vs. dialogue). For example, it is possible that spontaneous speech produced in the most demanding condition (L2 dialogue) would require greater IC than speech produced in the least demanding condition (L1 monologue).

To test these predictions, we recorded participants as they spontaneously produced L1 and L2 monologue and dialogue speech (each participant performed in every condition). Participants also completed a battery that assessed their L2 proficiency and inhibitory capacity. To elicit spontaneous speech, we used a modified version of the Map task (Anderson et al., 1991), which is frequently used to study spontaneous speech in the context of natural dialogues (Brown and Miller, 1980; Macafee, 1983; Macaulay, 1985). In this task, each of two conversational partners receives a map that the other cannot see. One partner is assigned the role of instruction giver, and the other of instruction follower. Each map contains a starting point and black and white drawings of

landmarks, along with their word labels, that occasionally mismatch across the instruction giver and follower's map versions. Of note, the instruction giver's map has a route that must be verbally described so that the instruction follower can reconstruct the route on her own map. Because some of the landmarks mismatch across the maps, conversational partners spend time discussing these discrepancies (see Appendices A and B for examples of maps and speech output).

We modified the Map task procedure in the following ways. First, participants always served as instruction givers, and the same experimental confederate always served as the instruction follower. Second, we implemented a comparable monologue version in which participants instructed a "hypothetical" listener. Finally, all participants performed the task in their L1 and L2, with order counterbalanced across participants.

All speech output was digitally recorded and analyzed with respect to two kinds of measures: global language output measures, which provided information about the content of what was produced, and acoustic-temporal measures, which provided information about how the speech was produced in real time. Global language output measures consisted of the subjective impressions of trained raters regarding the clarity of speaker's instructions (clarity of semantic content), the fluency of the speaker (smoothness of speech, absence of interruptions, hesitations and self-repairs, and changes in speech rate), and the extent to which the speaker sounded native-like.

Acoustic-temporal measures were ascertained using software that we developed to extract from the speech recordings the number of vocalizations and their length, and the silent pause durations preceding each vocalization. We used these two indices to compute a ratio, which consisted of individual vocalization durations over their prior silent pause durations (VD/PPD) across all utterances (see Materials and Methods for further detail). We focused on the ratio between each vocalization duration and its prior silent pause duration, based on prior work suggesting that vocalization durations reflect speech output effort (Henderson et al., 1966; Goldman-Eisler, 1968; Kormos and Dénes, 2004; Kormos, 2006; Segalowitz, 2010), and that prior silent pause durations reflect speech planning effort (Lindsley, 1975; Chaffe, 1980; Levelt, 1983; Ferreira, 1991; Segalowitz, 2010). Given these findings, it stands to reason that a large ratio reflects a situation where a given vocalization is less effortful to plan than a vocalization having a small ratio. As well, examining this ratio, rather than vocalization duration or internal pause duration alone, has an advantage of standardizing any difference in vocalization durations that could arise due to within- or between-monologues or dialogues, participants, or languages.

## MATERIALS AND METHODS

### PARTICIPANTS

A total of 22 English-French and 20 French-English bilingual adults ( $N = 42$ ,  $M = 21.21$ ,  $SD = 2.52$ ; seven males, 35 females) from McGill University (Montréal, Canada) participated for course credit. Participants were healthy young adults, 18–35 years old, with normal or corrected-to-normal vision, and no self-reported speech or hearing disorders. Originally, we recruited 64 participants (32 English-French and 32 French-English) but

we excluded 22 participants (10 English–French and 12 French–English) for the following reasons. Four reported acquiring first language other than English or French (two from each group). Seven reported that L2 was currently their more-dominant language (all French–English). Seven reported on a L2 proficiency questionnaire that they would not choose to speak L2 at all (five English–French and two French–English). Three were excluded because of equipment failure during sound recording (all English–French). One participant did not complete a portion of the speech production task (French–English).

We used an adapted version of the language experience and proficiency questionnaire (LEAP-Q) to assess participants' L2 proficiency (Marian et al., 2007). At the time of testing, French–English bilingual participants reported learning French as their first language, rated it as their dominant language, and reported high proficiency in English. Similarly, English–French bilingual participants reported learning English as their first language, rated it as their dominant language, and reported high proficiency in French. For subsequent analyses, we used the rating sub-scales of the LEAP-Q to calculate a standardized L2 proficiency score, modeled after McMurray et al. (2010). **Table 1** summarizes self-assessed L2 proficiency measures.

## MATERIALS

We selected four pairs of maps from the Map task corpus (<http://groups.inf.ed.ac.uk/maptask/#maps>). Two maps were used

to elicit monologue speech for each participant, once in L1 and once in L2, and two additional maps were used to elicit L1 and L2 dialogue speech. Because the Map task corpus was created in English, we translated verbal labels into French and pasted them onto new maps.

## PROCEDURE

We randomly assigned participants to one of two counterbalancing streams (see **Figure 2**). As illustrated in **Figure 2**, we counterbalanced whether the Map task was performed first in the L1 or L2 separately for English–French and French–English participants. All participants completed the monologue version of the Map task in one language, followed by the dialogue version of the Map task in the same language. Then, they completed the monologue version of the Map task in the other language, followed by the dialogue version of the Map task in the same language. Half of the participants completed the Map task in L1 first (left panel of **Figure 2**) the other half of participants completed the Map task in L2 first (right panel of **Figure 2**). Following Map task administration, all participants completed a battery that assessed their inhibitory capacity, the vocabulary subtest of the Wechsler abbreviated scale of intelligence (WASI) and a language background questionnaire. Testing session lasted approximately 2 hours.

Across all monologue and dialogue versions, participants always served as the instruction giver. In the English version of

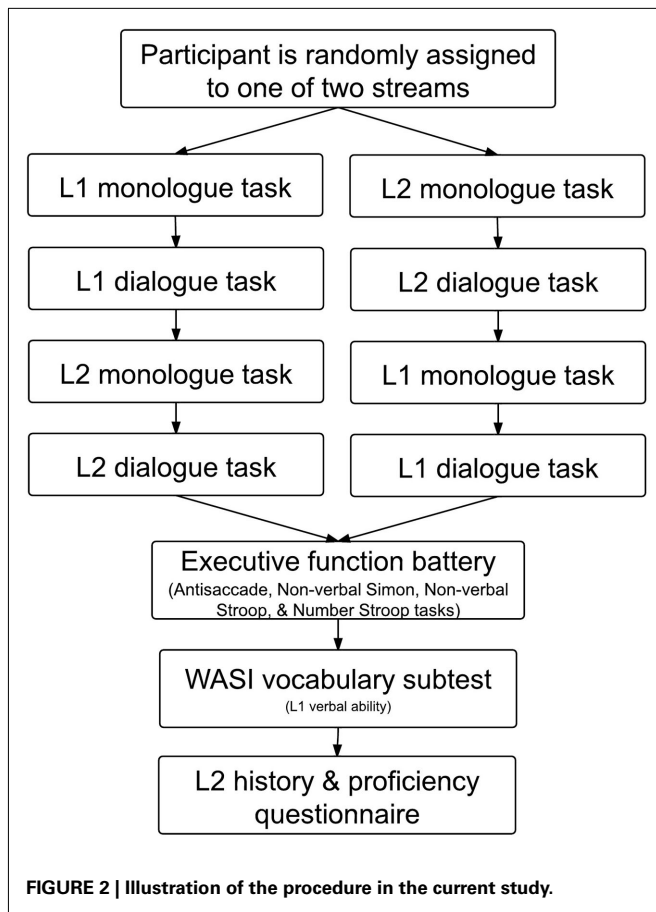
**Table 1 | Self-assessed L2 proficiency ratings, language history, and standardized L2 proficiency scores ( $n = 42$ ).**

|  | English–French ( $n = 22$ ) |           | French–English ( $n = 20$ ) |           |
|--|-----------------------------|-----------|-----------------------------|-----------|
|  | <i>M</i>                    | <i>SD</i> | <i>M</i>                    | <i>SD</i> |
| Rating scales (0–10)                                       |                             |           |                             |           |
| Speaking ability   | 7                           | 2         | 8                           | 2         |
| Reading ability  | 8                           | 2         | 8                           | 1         |
| Writing ability  | 7                           | 2         | 7                           | 1         |
| Translating ability  | 7                           | 2         | 8                           | 2         |
| Listening comprehension                                    | 8                           | 2         | 8                           | 1         |
| Pronunciation  | 7                           | 2         | 7                           | 2         |
| Fluency  | 7                           | 2         | 7                           | 2         |
| Vocabulary   | 7                           | 2         | 7                           | 2         |
| Grammatical ability  | 7                           | 2         | 7                           | 1         |
| Overall competence   | 7                           | 2         | 8                           | 1         |
| Sum of rating scales (0–100)                               | 71                          | 19        | 76                          | 13        |
| Standardized L2 proficiency score                          | −0.03                       | 0.93      | 0.27                        | 0.67      |
| Age of acquisition (years old)                             |                             |           |                             |           |
| Began acquiring L2*  | 5                           | 2         | 7                           | 4         |
| Became competent in L2                                     | 10                          | 4         | 12                          | 5         |
| Choose to speak L2 (%)**                                   | 17                          | 12        | 34                          | 21        |
| Degree of L1 interference when speaking in L2 (0–5)**      | 2                           | 1         | 3                           | 1         |
| Percent of present time spent functioning in each language |                             |           |                             |           |
| L1***  | 82                          | 9         | 48                          | 17        |
| L2***  | 14                          | 7         | 50                          | 17        |

\*Two-tailed independent samples *t*-test significant at  $p < 0.05$ .

\*\*Two-tailed independent samples *t*-test significant at  $p < 0.01$ .

\*\*\*Two-tailed independent samples *t*-test significant at  $p < 0.001$ .



the Map task, we instructed participants in English to verbally guide the instruction follower through a printed route from start to finish, in English. In the French version of the Map task, we instructed participants in French to verbally guide the instruction follower, in French. In the monologue versions, we instructed participants to guide an imaginary person. In the dialogue versions, we instructed participants to guide their conversational partner, a confederate of the experiment. In the dialogue versions, the confederate reproduced the route on her version of the map, based strictly on the instructions of the participants. When participants and the confederate encountered discrepancies in labels across their versions of maps, unknown to the participant, the confederate was required to exclusively refer to the landmarks by the labels printed on her map.

Participants and the confederate performed the Map task in the same room. Participants and the confederate were instructed to speak at a normal rate, and faced away from each other to prevent gaze and posture coordination (Shockley et al., 2009). During monologues and dialogues, participants viewed maps on a 20" monitor located 71 cm away from where they were seated. Participants wore an AKG C420 PP MicroMic Series III headset microphone, while we used a Zoom H4 Handy Recorder to record their speech at 44 kHz in stereo, such that participants' voice was acoustically isolated to the left channel and the confederate's to the right channel.

## INDIVIDUAL DIFFERENCES MEASURES

To assess individual differences in inhibitory capacity, we administered an anti-saccade task (Hallett, 1978), a non-linguistic Simon (Simon and Ruddell, 1967), and Stroop (Stroop, 1935) tasks modeled after Blumenfeld and Marian (2011) and a Number Stroop task. To assess L1 verbal ability, we administered vocabulary subtest of the WASI (Wechsler, 1999).

### Anti-saccade task

This task assessed ability to inhibit the pre-potent tendency to look toward a peripherally presented target (Hallett, 1978). We used an Eye-Link 1000 tower mounted system (SR-Research, ON, Canada) with a sampling rate of 1 kHz to monitor and record fixation durations of the right eye. Participants were presented randomly intermixed pro-saccade and anti-saccade trials. At the onset of each trial, participants saw a small black fixation circle in the center of the computer screen, followed by a central fixation square that remained on the screen for 1000, 1250, or 1500 ms. The central fixation square was green to cue participants to engage in a pro-saccade trial, and red to cue participants to engage in an anti-saccade trial. Thus, contingent on the color of the central fixation square, participants looked toward (pro-saccade trials) or away (anti-saccade trials) from peripherally located black square targets. We computed an *Anti-saccade Cost* variable for each participant based on correct trials only (Bialystok et al., 2006), where we subtracted the average reaction time of all pro-saccade trials from the average reaction time of all anti-saccade trials.

### Non-linguistic simon and stroop tasks

We adapted these tasks from Blumenfeld and Marian (2011). Participants saw arrows on a screen. In the Simon task, the arrows pointed up or down. When the arrows pointed up, participants used their left hand to press a response button on the left, and when the arrows pointed down, participants used their right hand to press a response button on the right. Trials were congruent when the arrow appeared on the same side of the computer screen as the response and incongruent when the arrow appeared on the opposite side of the computer screen as the response. The Simon effect reflects the finding that participants execute a motor response more quickly and accurately when the left/right spatial location of the stimulus corresponds to the left/right spatial location of the response button (Simon and Ruddell, 1967). In the Stroop task, the arrows pointed left or right. When the arrows pointed left, participants used their left hand to press a response button on the left, and when the arrows pointed right, participants used their right hand to press a response button on the right. Trials were congruent when the arrow appeared on the same side as its pointed direction and incongruent when the arrow appeared on the opposite side as its pointed direction. The Stroop effect reflects the finding that participants execute a motor response more quickly and accurately when the semantic meaning of the stimulus corresponds to the required response (Stroop, 1935). We computed a cost score for the Simon and Stroop tasks separately, in which we subtracted the average reaction time on congruent trials from the average reaction time on incongruent trials. Only correct trials were included in these averages.

### Number stroop task

This task also assessed the ability to inhibit a strong automatic cognitive response. We presented a series of numbers ranging from one to four digits on a computer screen. Participants were instructed to use their dominant hand to press one of four response buttons that corresponded to the number of digits appearing on the screen. Trials were congruent when the quantity of digits corresponded to the depicted numbers (22 required response 2) and incongruent when the quantity of digits did not correspond to the depicted numbers (e.g., 222 required response 3). We computed a cost score for the correct reaction times for each participant by subtracting the average reaction time on congruent trials from the average reaction time on incongruent trials.

Descriptive statistics from each task appear in **Table 2**. Two-tailed independent samples *t*-tests revealed that performance did not significantly differ between English–French and French–English participants on all tasks ( $p > 0.05$ ). Using these measures of inhibitory capacity, we computed a standardized composite inhibition cost score (McMurray et al., 2010).

### WASI vocabulary subtest

Participants defined words in L1, which we scored and transformed into scaled score using age-appropriate norms.

## RESULTS

We constructed a series of linear mixed effect (LME) models, as implemented in the lme4 library (Bates, 2005) in R Project for Statistical Computing version 2.10.1 (Baayen, 2008; Baayen et al.,

2008; R Development Core Team, 2009). The models included as variables of interest the main effects and interactions of language (L1 vs. L2), speech type (monologue vs. dialogue), L2 proficiency (continuous), and inhibitory capacity score (continuous). All models had random intercepts for items (i.e., number of different maps) and participants (Baayen, 2008). All models had language group (English–French vs. French–English) as control variable to account for L2 vs. L1 linguistic differences between two groups. We excluded L1 verbal ability (WASI scaled scores) from the models reported below because there was only one instance where it accounted for a significant amount of variance. This was in the clarity of instructions measure (see below), where increased verbal ability was associated with higher ratings. Our dependent variables consisted of the global output measures and acoustic–temporal measures previously described. We first report the results for the global output measures, followed by results for the acoustic–temporal measures. Within each set of analyses, we first report the analyses that assess the contribution of L2 proficiency, followed by analyses that assess the added contribution of inhibitory capacity.

### GLOBAL OUTPUT MEASURES

Global output measures included the clarity of speaker's instructions and speaker fluency and nativeness. We selected and adapted these measures from the work of (Pinkham and Penn, 2006). To obtain these measures a team of independent raters (two native-English and two native-French) coded participants' speech files separately in monologues and dialogues and in L1 and L2. For each monologue or dialogue recording, the independent raters assigned a score from one to nine on the following dimensions, the clarity of speaker's instructions and speaker fluency and nativeness. Raters were trained on 20 English and French speech samples; however, they coded only speech samples that matched their native language. Interrater reliability on the training samples was high (Cronbach's alpha = 0.93). Descriptive statistics for each dimension are shown in **Table 3**.

### L2 proficiency and the clarity of instructions

**Table 4** presents the results of LME models for clarity of instructions. The clarity of instructions was lower in L2 ( $M = 6.67$ ) than L1 speech ( $M = 7.06$ ), resulting in a significant main effect of language ( $t = -2.02$ ,  $p < 0.05$ ). As well, the clarity of instructions was lower in monologues ( $M = 6.34$ ) than dialogues ( $M = 7.39$ ), resulting in a significant main effect of speech type ( $t = 2.96$ ,  $p < 0.01$ ). Finally, the clarity of instructions varied with the language of production and L2 proficiency, resulting in a significant two-way interaction between language and L2 proficiency ( $t = 2.09$ ,  $p < 0.05$ ). This interaction is depicted in **Figure 3**. The left panel of **Figure 3** shows that the clarity of instructions in monologues was significantly lower in L2 than in L1 speech for bilinguals with low L2 proficiency. Moreover, the L2 vs. L1 difference in the clarity of instructions decreased as L2 proficiency increased. Finally, the right panel of **Figure 3** shows that the clarity of instructions did not differ between L1 and L2 across all levels of L2 proficiency in dialogues.

**Table 2 | Minima, maxima, means, and SDs for individual difference measures.**

|                              | Min     | Max    | <i>M</i> | SD     |
|------------------------------|---------|--------|----------|--------|
| <b>L1 VERBAL ABILITY</b>     |         |        |          |        |
| WASI score                   | 8       | 18     | 14       | 3      |
| <b>SIMON TASK</b>            |         |        |          |        |
| Congruent                    | 376     | 660    | 489      | 65     |
| Incongruent                  | 424     | 728    | 527      | 73     |
| Cost                         | -48     | -68    | -38***   |        |
| <b>STROOP TASK</b>           |         |        |          |        |
| Congruent                    | 275     | 844    | 480      | 108    |
| Incongruent                  | 211     | 783    | 510      | 106    |
| Cost                         | 64      | 61     | -30***   |        |
| <b>NUMBER STROOP TASK</b>    |         |        |          |        |
| Congruent                    | 446     | 717    | 582      | 72     |
| Incongruent                  | 475     | 833    | 652      | 87     |
| Cost                         | -29     | -116   | -70***   |        |
| <b>ANTI-SACCADE TASK</b>     |         |        |          |        |
| Pro-saccade                  | 230     | 415    | 304      | 47     |
| Anti-saccade                 | 341     | 528    | 411      | 45     |
| Cost                         | -111    | -113   | -107***  |        |
| <b>INHIBITION COST SCORE</b> |         |        |          |        |
| English–French               | -0.0099 | 0.0087 | -0.0002  | 0.0048 |
| French–English               | -0.0070 | 0.0088 | 0.0002   | 0.0037 |

\*\*\*Two-tailed paired samples *t*-test significant at  $p < 0.001$ .

**Table 3 | Mean ratings and SEs of the mean/or clarity of instructions, speaker fluency, and nativeness for monologues and dialogues in L1 and in L2.**

|                         |           | Language | <i>M</i> | SE   |
|-------------------------|-----------|----------|----------|------|
| Clarity of instructions | Monologue | L1       | 6.62     | 0.26 |
|                         |           | L2       | 6.05     | 0.33 |
|                         | Dialogue  | L1       | 7.49     | 0.18 |
|                         |           | L2       | 7.28     | 0.18 |
| Speaker fluency         | Monologue | L1       | 7.52     | 0.17 |
|                         |           | L2       | 7.74     | 0.15 |
|                         | Dialogue  | L1       | 7.84     | 0.16 |
|                         |           | L2       | 6.58     | 0.22 |
| Speaker nativeness      | Monologue | L1       | 8.86     | 0.05 |
|                         |           | L2       | 6.98     | 0.25 |
|                         | Dialogue  | L1       | 8.88     | 0.05 |
|                         |           | L2       | 6.30     | 0.30 |

**Table 4 | Linear mixed effects models for global output measures (clarity of instructions, speaker fluency, and nativeness) to illustrate interactions between speech type, language, and L2 proficiency.**

|   | Clarity of instructions |          |                 | Speaker fluency |          |                 | Speaker nativeness |          |                 |
|---|-------------------------|----------|-----------------|-----------------|----------|-----------------|--------------------|----------|-----------------|
|   | <i>b</i>                | SE       | <i>t</i> -Value | <i>b</i>        | SE       | <i>t</i> -Value | <i>b</i>           | SE       | <i>t</i> -Value |
| Fixed effects   |                         |          |                 |                 |          |                 |                    |          |                 |
| Intercept   | 6.34                    | 0.25     | 25.08***        | 7.32            | 0.21     | 35.77***        | 8.55               | 0.22     | 39.62***        |
| Speech type (monologue, dialogue) <sup>1</sup>                  | 0.95                    | 0.32     | 2.96**          | 0.30            | 0.22     | 1.32            | 0.02               | 0.26     | 0.08            |
| Language (L1, L2) <sup>2</sup>                                  | -0.65                   | 0.32     | -2.02*          | 0.19            | 0.22     | 0.86            | -2.02              | 0.22     | -9.04***        |
| L2 proficiency  | 0.59                    | 0.28     | 1.75            | 0.15            | 0.22     | 0.67            | -0.06              | 0.21     | -0.26           |
| Language group <sup>3</sup> (English–French vs. French–English) | 0.35                    | 0.23     | 1.53            | 0.42            | 0.22     | 1.91            | 0.65               | 0.20     | 3.22**          |
| Speech type × language  | 0.41                    | 0.45     | 0.90            | -1.45           | 0.32     | -4.59***        | -0.69              | 0.32     | -2.18*          |
| Speech type × L2 proficiency                                    | -0.30                   | 0.39     | -0.77           | -0.03           | 0.27     | -0.11           | 0.11               | 0.27     | 0.40            |
| Language × L2 proficiency                                       | 0.82                    | 0.39     | 2.09*           | -0.18           | 0.27     | -0.65           | 1.05               | 0.27     | 3.84**          |
| Speech type × language × L2 proficiency                         | -0.81                   | 0.55     | -1.46           | 0.37            | 0.39     | 0.96            | -0.08              | 0.39     | -0.20           |
| Random effects  |                         | Variance |                 |                 | Variance |                 |                    | Variance |                 |
| Subject   |                         | 0.00     |                 |                 | 0.23     |                 |                    | 0.16     |                 |
| Item  |                         | 0.00     |                 |                 | 0.00     |                 |                    | 0.02     |                 |
| Residual  |                         | 2.10     |                 |                 | 1.03     |                 |                    | 1.01     |                 |

\**p*MCMC < 0.05 level, \*\**p*MCMC < 0.01 level, \*\*\**p*MCMC < 0.001 level.

<sup>1</sup>Baseline = monologue.

<sup>2</sup>Baseline = L1.

<sup>3</sup>Baseline = English–French.

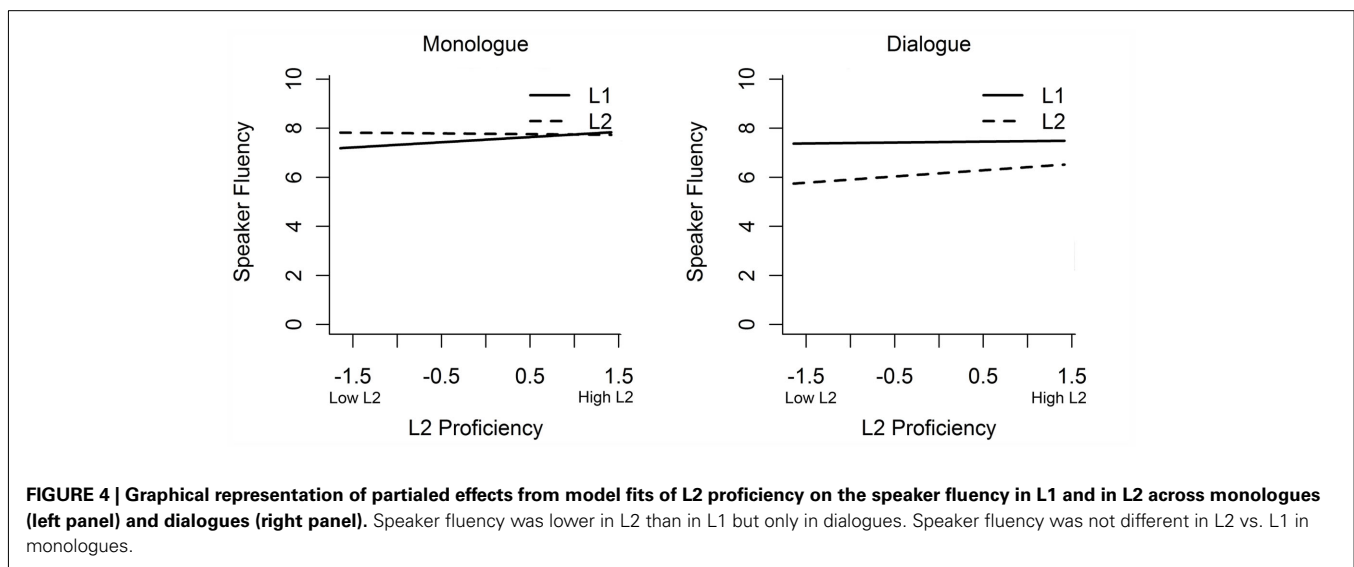
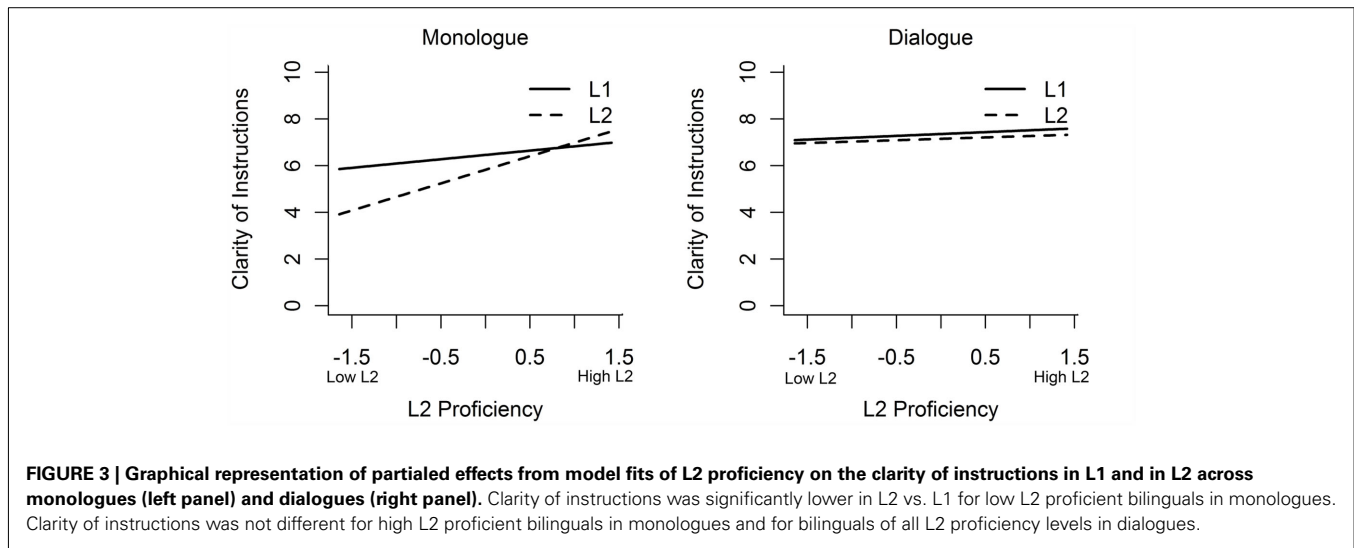
### L2 proficiency and speaker fluency

Speaker fluency varied as a function of language of production (L1 vs. L2) and speech type (monologue vs. dialogue), resulting in a significant two-way interaction between language and speech type ( $t = -4.59, p < 0.001$ ). As shown in **Table 3**, speaker fluency was lowest in L2 dialogues ( $M = 6.58$ ) as compared to L1 monologues ( $M = 7.52$ ), L2 monologues ( $M = 7.74$ ), and L1 dialogues ( $M = 7.84$ ), the latter of which did not differ. This interaction is depicted in **Figure 4**. The left panel of **Figure 4** shows that speaker fluency did not significantly differ between L1 and L2 monologues. The right panel of **Figure 4** shows that speaker fluency was significantly lower in L2 than L1 dialogues. L2 proficiency did not significantly predict speaker fluency for L1 vs. L2 and monologues vs. dialogues.

### L2 proficiency and speaker nativeness

Speaker nativeness was lower in L2 ( $M = 6.64$ ) than L1 speech ( $M = 8.87$ ), resulting in a significant main effect of language ( $t = -9.04, p < 0.001$ ). Speaker nativeness also varied as a function of the language of production (L1 vs. L2) and speech type (monologue vs. dialogue), resulting in a significant two-way interaction between language and speech type ( $t = -2.18, p < 0.05$ ). As shown in **Table 3**, speaker nativeness was lowest in L2 dialogues ( $M = 6.30$ ) followed by L2 monologues ( $M = 6.98$ ) and highest in L1 monologues ( $M = 8.86$ ) and L1 dialogues ( $M = 8.88$ ). Finally, speaker nativeness varied as a function of language of production and L2 proficiency, resulting in a significant two-way interaction between language and L2 proficiency ( $t = 3.84, p < 0.01$ ). **Figure 5** shows this interaction across left and right





panels. Speaker nativeness for L1 monologues and dialogues was high across all levels of L2 proficiency. Conversely, speaker nativeness for L2 monologues and dialogues varied as a function of L2 proficiency. Bilinguals with low L2 proficiency showed lower speaker nativeness than bilinguals with high L2 proficiency.

#### ***Inhibitory capacity and clarity of instructions, speaker fluency, and nativeness***

To assess whether individual differences in inhibitory capacity modulated global output measures, we included the composite inhibition cost score as a fixed effect to the models previously described. Thus, we constructed models with four-way interactions between language, speech type, L2 proficiency and inhibition cost score for clarity of instructions, speaker fluency, and speaker nativeness. Within these final models, inhibitory capacity did not significantly relate to any of the global output measures, neither as the main effect nor as part of the higher-order interactions (all  $t_s < 1.53$ ,  $p > 0.05$ ).

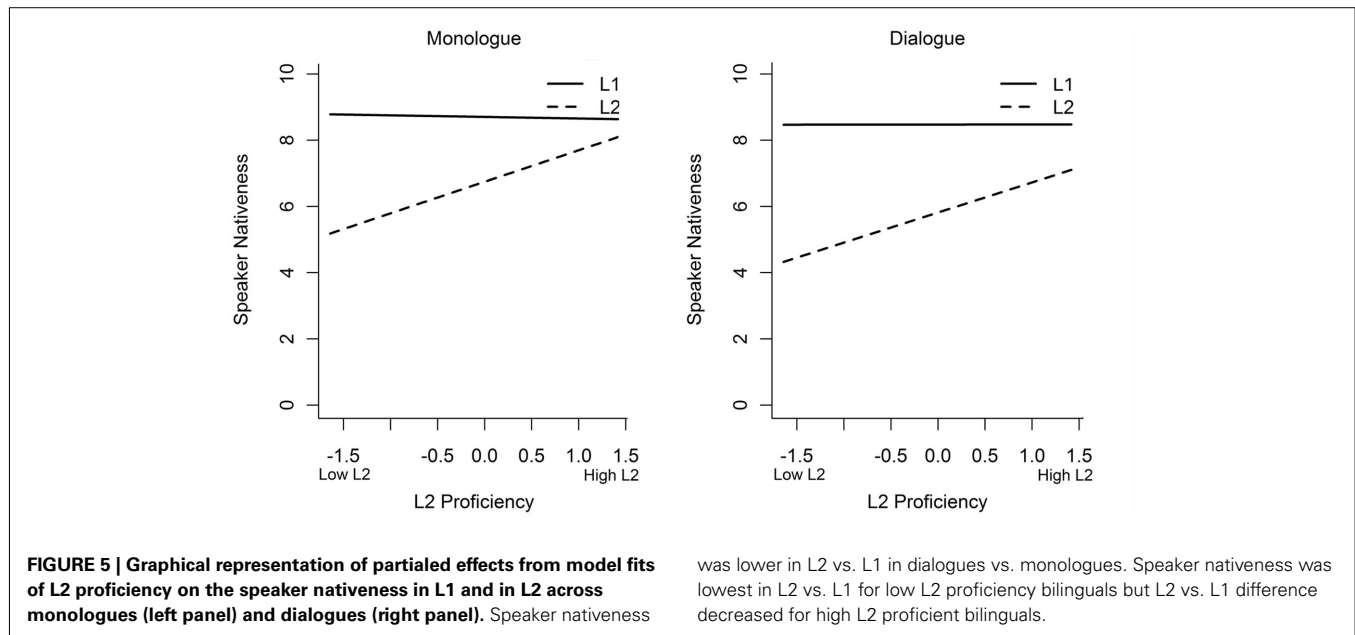
#### **ACOUSTIC-TEMPORAL MEASURES OF SPEECH PRODUCTION**

The acoustic-temporal measure of interest was the ratio of individual vocalization durations over their prior silent pause durations (VD/PPD). Again, we assumed greater ratios reflect increased efficiency of speech planning. First, we describe how we processed speech files to compute this measure.

#### ***Pre-processing of speech files***

To minimize cross-talk between conversational partners, we recorded speech at a relatively low volume. Thus, prior to analysis, we amplified the speech signal by 26 dB and removed inaudible speech below 40 dB. We used Soundforge (version 8.0, Sony Creative Software) to standardize the amplitude of the speech signal across monologues and dialogues, and to remove all instances of coughs and laughs. After this pre-processing stage, we used custom software to distinguish periods of vocalization from periods of silence for each speaker, based on prior work (Alpert et al., 1986; Welkowitz et al., 1990). For the purpose of this study, we





was lower in L2 vs. L1 in dialogues vs. monologues. Speaker nativeness was lowest in L2 vs. L1 for low L2 proficiency bilinguals but L2 vs. L1 difference decreased for high L2 proficient bilinguals.

only selected instances where silent pause preceded a vocalization duration uttered by the participant (see Appendix C). Independent periods of vocalization were registered when the speaker signal exceeded minimum amplitude for at least 250 ms. Periods of silence were registered when the speaker signal remained below minimum amplitude for at least 250 ms. These timing parameter estimates were based on prior work using similar automated speech processing methods and other studies of spontaneous speech (Goldman-Eisler, 1968; Alpert et al., 1986; Welkowitz et al., 1990; Wilson and Wilson, 2005; Kormos, 2006; Segalowitz, 2010). Initial silences (prior to the initial vocalization or following the final vocalization) and silences less than 250 ms were removed from estimates of the mean vocalization durations. Descriptive statistics for vocalization and silent pause durations are shown in Table 5.

Of note, our custom software also identifies in the speech signal switching pauses, turn-taking boundaries, and strong and weak interruptions. While these are important features of dialogue speech, we excluded them from the calculations of ratios to enable direct comparison of dialogue and monologue speech, the latter of which lacks these features.

#### **L2 proficiency and the ratio of vocalization durations to prior pause durations (VD/PPD)**

VD/PPD ratios were smaller for L2 ( $M = 3.33$ ) vs. L1 speech ( $M = 3.86$ ), resulting in a main effect of language ( $t = -3.93$ ,  $p < 0.001$ ). VD/PPD ratios were smaller in dialogues ( $M = 3.33$ ) than monologues ( $M = 3.86$ ), resulting in a main effect of speech type ( $t = -3.48$ ,  $p < 0.001$ ). Finally, VD/PPD ratios varied as a function of language of production, speech type, and L2 proficiency. This resulted in a significant three-way interaction between speech type, language, and L2 proficiency ( $t = -2.60$ ,  $p < 0.05$ ), shown in Figure 6. The left panel of Figure 6 shows that VD/PPD ratios for monologues were smaller in L2 than L1 for low L2 proficiency bilinguals. However, the L2 vs. L1 difference in VD/PPD ratios for monologues decreased as L2 proficiency increased. In

particular, as L2 proficiency increased, it appears that L2 VD/PPD ratios also increased while L1 VD/PPD ratios decreased. In contrast to monologues, there was no effect of L2 proficiency for dialogues. The right panel of Figure 6 shows that VD/PPD ratios were smaller for L2 vs. L1 speech, regardless of L2 proficiency.

#### **Inhibitory capacity and VD/PPD ratios**

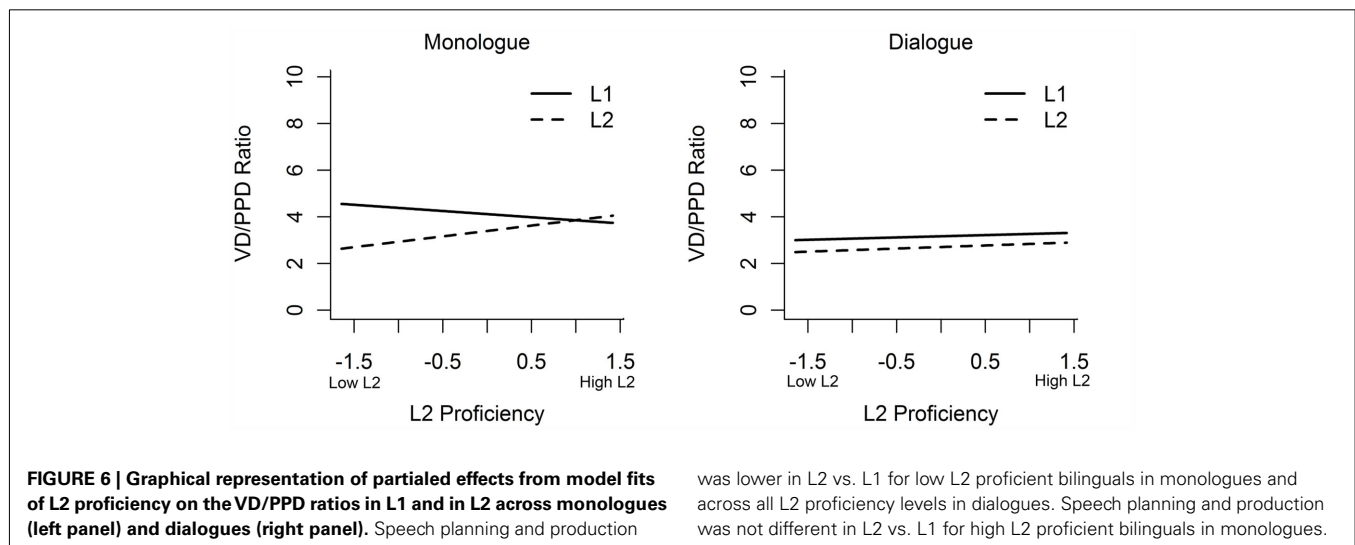
To investigate whether individual differences in inhibitory capacity relate to monologue and dialogue speech production, we added as a fixed effect the composite inhibition cost score to the three-way interaction (language  $\times$  speech type  $\times$  L2 proficiency) of the model just presented. Table 6 presents the results of this LME model. There was again a significant three-way interaction between language, speech type, and L2 proficiency, but no four-way interaction with inhibitory capacity ( $t = -0.85$ ,  $p > 0.05$ ). However, VD/PPD ratios decreased as inhibitory capacity decreased (inhibition cost increased), resulting in a main effect of inhibitory capacity ( $t = -2.20$ ,  $p < 0.05$ ). As well, inhibitory capacity interacted with speech type and L2 proficiency, resulting in a significant three-way interaction between speech type, L2 proficiency, and inhibitory capacity ( $t = 2.70$ ,  $p < 0.01$ ). This interaction is shown in Figure 7. As seen in the upper and lower left panels of Figure 7, VD/PPD ratio increased as both L2 proficiency and inhibitory capacity increased. In contrast, as seen in the upper right panel of Figure 7, VD/PPD ratios did not significantly vary for L1 dialogues as a function of L2 proficiency or inhibitory capacity. Finally, as seen in the lower right panel of Figure 7, VD/PPD ratios again increased as both L2 proficiency and inhibitory capacity increased.

## **DISCUSSION**

We investigated how individual differences in L2 proficiency and inhibitory capacity relate to bilinguals' spontaneous monologue and dialogue language production. There were several key findings pertaining to the role of L2 proficiency, task demands, and inhibitory capacity.

**Table 5 | Mean values (ms), SEs of the mean, and mean observation count for vocalization durations, prior silent pause durations, computed VD/PPD ratios, and total speech sample duration for monologues and dialogues in L1 and in L2.**

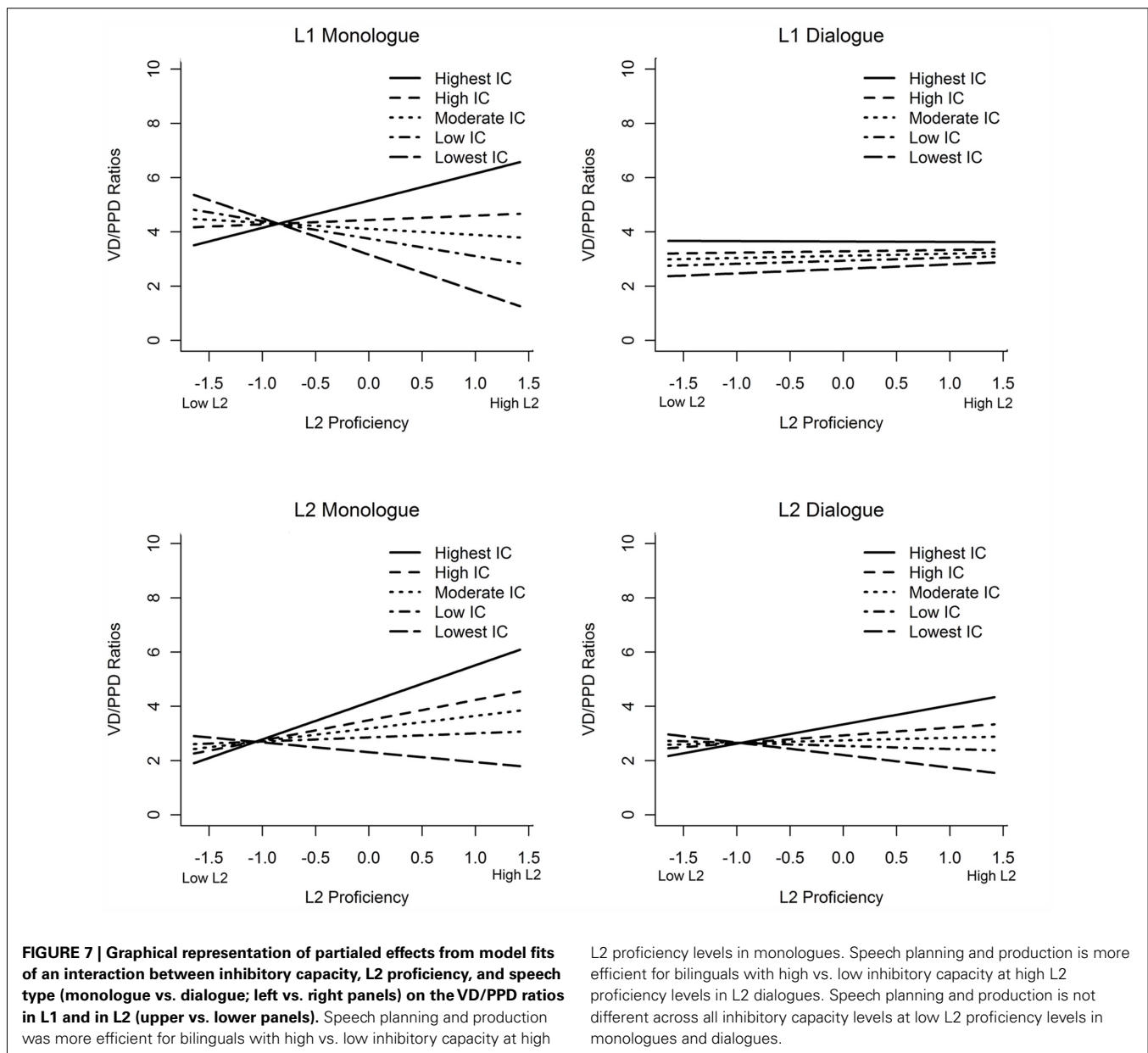
|                             |           | Language | Mean   | SE    | Mean observation count |
|-----------------------------|-----------|----------|--------|-------|------------------------|
| Vocalization duration       | Monologue | L1       | 2272   | 280   | 32.30                  |
|                             |           | L2       | 1959   | 274   | 40.72                  |
|                             | Dialogue  | L1       | 1624   | 237   | 52.14                  |
|                             |           | L2       | 1422   | 225   | 66.33                  |
| Prior silent pause duration | Monologue | L1       | 689    | 58    |                        |
|                             |           | L2       | 698    | 61    |                        |
|                             | Dialogue  | L1       | 584    | 63    |                        |
|                             |           | L2       | 596    | 59    |                        |
| VD/PPD ratios               | Monologue | L1       | 4.14   | 0.64  |                        |
|                             |           | L2       | 3.58   | 0.60  |                        |
|                             | Dialogue  | L1       | 3.58   | 0.61  |                        |
|                             |           | L2       | 3.08   | 0.56  |                        |
| Total sample duration       | Monologue | L1       | 151169 | 9292  |                        |
|                             |           | L2       | 154476 | 10775 |                        |
|                             | Dialogue  | L1       | 494494 | 27188 |                        |
|                             |           | L2       | 468151 | 21076 |                        |



Consider first the results for the global output measures. The clarity of instructions produced was higher when people spoke in their L1 than their L2, although increased L2 proficiency helped to close the gap between L1 and L2 clarity. Dialogue speech also was rated as clearer in content than monologue speech, which is consistent with recent work suggesting that dialogue speech is easier to produce than monologue speech and also that the goal of dialogue is to relay the message clearly to a conversational partner (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Hartsuiker and Pickering, 2008; Kootstra et al., 2010). Specifically, this finding suggests that the presence of a conversational partner was associated with enriched semantic content during language production, presumably because the conversational partner provided the speaker with ongoing feedback about when the content of their output was unclear.

The other two global output measures behaved somewhat differently from the clarity of instructions. Speech fluency (whether people spoke in a fluid or halting way) was generally high except for L2 dialogue speech, which is arguably the most cognitively demanding of the different language production conditions. This effect of speech fluency was unaffected by differences in L2 proficiency. Speaker nativeness, in contrast, was influenced by several factors: L2 proficiency, the language of speech, and whether a monologue or a dialogue was produced. L2 speech was rated as less native-like than L1 speech, and this difference was larger for dialogue than monologue speech. Finally, the difference between L1 and L2 speaker nativeness also decreased as L2 proficiency increased.

Taken together, the global output measures suggest that language knowledge (whether L1 or L2 production is adjusted



by individual differences in L2 proficiency) and task demands (whether people produce speech in a monologue or a dialogue) modulate the substance of what is produced during spontaneous monologue or dialogue speech. Absent here are any effects arising from individual differences in inhibitory capacity. This is potentially surprising given the IC model's focus on inhibition as a critical mechanism for bilingual language processing. However, it is possible that global measures of language production output are not the most appropriate level of analysis to observe an effect of inhibitory capacity. Rather, as clearly implied by the IC model, inhibition may have more local effects on the ongoing planning of individual vocalizations.

Indeed, we found clear evidence that the acoustic-temporal measures showed sensitivity to individual differences in inhibitory capacity. Recall, our primary acoustic-temporal measure was the

L2 proficiency levels in monologues. Speech planning and production is more efficient for bilinguals with high vs. low inhibitory capacity at high L2 proficiency levels in L2 dialogues. Speech planning and production is not different across all inhibitory capacity levels at low L2 proficiency levels in monologues and dialogues.

ratio between the duration of each vocalization and the duration of its prior pause (VD/PPD). Prior work suggests that there is a close linkage between the planning that takes place prior to a vocalization, and the nature of what is produced (Lindsay, 1975; Chaffe, 1980; Levelt, 1983; Ferreira, 1991; Segalowitz, 2010). Thus, a large value for this ratio should indicate that a speaker produced a given vocalization with relatively little planning effort. In contrast, a small value for this ratio should indicate that a speaker produced a given vocalization with relatively more planning effort. Consistent with our findings for the global output measures, monologues had higher ratios than dialogues. L1 speech also had higher ratios than L2 speech, although increased L2 proficiency reduced this difference overall. Unlike monologues, dialogues had more uniform ratios, as seen in Figure 6.

**Table 6 | Linear mixed effects models for the temporal measure.**

|   | VD/PPD ratios |                 |                 |
|---|---------------|-----------------|-----------------|
|   | <i>b</i>      | SE              | <i>t</i> -Value |
| <b>Fixed effects</b>  |               |                 |                 |
| Intercept   | 3.85          | 0.29            | 13.16**         |
| Speech type (monologue, dialogue) <sup>1</sup>                  | −0.60         | 0.17            | −3.48***        |
| Language (L1, L2) <sup>2</sup>                                  | −0.72         | 0.18            | −3.93***        |
| L2 proficiency  | −0.42         | 0.28            | −1.48           |
| Inhibitory capacity   | −119.51       | 54.29           | −2.20*          |
| Language group <sup>3</sup> (English–French vs. French–English) | 0.54          | 0.39            | 1.39            |
| Speech type × language  | 0.20          | 0.23            | 0.87            |
| Speech type × L2 proficiency                                    | 0.50          | 0.21            | 2.38*           |
| Language × L2 proficiency                                       | 0.71          | 0.22            | 3.25**          |
| Speech type × inhibitory capacity                               | 68.83         | 40.91           | 1.68            |
| Language × inhibitory capacity                                  | −13.80        | 42.82           | −0.32           |
| L2 proficiency × inhibitory capacity                            | −120.38       | 68.69           | −1.75           |
| Speech type × language × L2 proficiency                         | −0.71         | 0.28            | −2.60*          |
| Speech type × language × inhibitory capacity                    | 9.80          | 52.29           | 0.19            |
| Speech type × L2 proficiency × inhibitory capacity              | 136.17        | 50.54           | 2.7**           |
| Language × L2 proficiency × inhibitory capacity                 | −17.54        | 53.69           | −0.33           |
| Speech type × language × L2 proficiency × inhibitory capacity   | −55.43        | 64.88           | −0.85           |
| <b>Random effects</b>   |               |                 |                 |
|   |               | <b>Variance</b> |                 |
| Subject   |               | 1.30            |                 |
| Item  |               | 0.00            |                 |
| Residual  |               | 12.12           |                 |

VD/PPD ratios (ratio of vocalization durations over their prior silent pause durations) to illustrate interactions between speech type, language, L2 proficiency, and inhibition capacity.

\**p*MCMC < 0.05 level, \*\**p*MCMC < 0.01 level, \*\*\**p*MCMC < 0.001 level.

<sup>1</sup>Baseline = monologue.

<sup>2</sup>Baseline = L1.

<sup>3</sup>Baseline = English–French.

However, individual differences in inhibitory capacity also modulated VD/PPD ratios for monologues and dialogues. For monologues, increased inhibitory capacity appears to have blocked for L1 monologues the apparent decline associated with increased L2 proficiency. At the same time, increased inhibitory capacity appears to have enhanced the apparent growth associated with increased L2 proficiency (left panel of **Figure 7**). For dialogues, in contrast, increased inhibitory capacity seems to have facilitated overall VD/PPD ratios when people conversed in their L1. Increased inhibitory capacity also seems to have facilitated VD/PPD ratios when people who are high in L2 proficiency conversed in their L2 (right panel of **Figure 7**).

Thus, it appears that high L2 proficient bilinguals may expend more local effort at each vocalization in their L1 to maintain a high level of L1 global output clarity. In contrast, it appears that high L2 proficient bilinguals may expend more local effort at each

vocalization in their L2, and at the same time the global clarity is significantly reduced. Finally, bilinguals who have greater inhibitory capacity produce language more efficiently at the level of individual vocalizations, over and above the effects of L2 proficiency, as a function of communicative task demands.

These results are consistent with prior work showing that speech production is more effortful in a less-dominant language (Hernandez et al., 2000; Kormos and Dénes, 2004; Fehring and Fry, 2007; Gollan and Ferreira, 2009; Hanulová et al., 2010; Sandoval et al., 2010), and that L2 proficiency is an important determinant of L1 and L2 production performance (Poulisse and Bongaerts, 1994; Costa and Caramazza, 1999; Kormos and Dénes, 2004; Gollan et al., 2005, 2008, 2011; Ivanova and Costa, 2008). Such effects of increased L2 proficiency on both L2 and L1 production are consistent with the IC model, according to which L2 production should be more controlled and effortful than L1 production, especially when L2 proficiency is low (see also Segalowitz, 2010). Presumably, however, as L2 proficiency increases, L2 production becomes relatively more routine and less effortful, while L1 production may become relatively less so (Abutalebi and Green, 2007).

Another key finding was that dialogue speech appeared to be more effortful than monologue speech across several measures, especially during L2 production. Specifically, dialogue speech was less fluent and native-like, and required more effort to produce at the individual vocalization level, consistent with prior work (Fehring and Fry, 2007). Interestingly, the semantic clarity of what was produced in the L2 was greater for dialogues than monologues, presumably because speakers had the opportunity to better monitor their output through feedback from their conversational partner. In this way, our results are also consistent with prior work suggesting that dialogue speech production may be easier than monologue speech production due to interactive alignment processes (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Hartsuiker and Pickering, 2008; Kootstra et al., 2010).

Our final key finding was that individual differences in inhibitory capacity modulated bilingual language production at the level of individual vocalizations and this interacted with communicative task demands. Specifically, bilinguals with higher inhibitory capacity were more efficient in planning and producing individual vocalizations than bilinguals with lower inhibitory capacity, particularly for monologue speech. In contrast, dialogue speech was generally more effortful overall. These findings are consistent with prior work showing that bilinguals with increased inhibitory capacity inhibit L1 during L2 production more efficiently than bilinguals with decreased inhibitory capacity, irrespective of L2 proficiency (Linck et al., 2008). Thus, consistent with the IC model, these findings suggest that increased L2 proficiency and inhibitory capacity are necessary for efficient bilingual speech planning and production.

While the results of this study improve our understanding of bilingual language production, there are several potential limitations that would be important to address in future work.

One potential limitation is that our particular use of the map task, where objects on the maps contained verbal labels, may have created a relatively low-demand communicative situation that

underestimated the normal challenges of spontaneous monologue and dialogue production. Thus, the effects of inhibitory capacity observed in this study might have been even more pronounced had we used a more demanding communicative task to elicit spontaneous speech. There are several features of our task that may have made it less demanding than expected: verbal labels on the maps; a single experienced confederate rather than a completely naïve conversational partner; the fact that dialogues always followed monologues may have preferentially advantaged dialogues over monologues. Regarding this latter point, however, there was little evidence of a dialogue advantage for any measure except the clarity of instructions.

In contrast, it is also possible that our dialogue speech condition may have been more demanding than normal because of the following. First, the confederate could interrupt the participant when encountering mismatches in map landmarks in dialogues. While no such mismatches were encountered during monologue speech, future work could assess whether presence vs. absence of mismatches in map landmarks in dialogues contributes to task difficulty. Second, participants and the confederate faced away from each other, thereby blocking any visual cues during conversational interaction. Given that conversational partners communicate more easily when the visual channel is available throughout dialogue speech (Doherty-Sneddon et al., 1997), it is possible that L1 and L2 dialogue speech may become less effortful when conversational partners can see each other as they speak. Thus, the results here for dialogue only generalize to auditory-only dialogue processes, such as when two people converse by telephone.

Another potential limitation is that it is possible that the dialogue speech condition had smaller VD/PPD ratios because of a higher likelihood of dialogues having shorter vocalizations than monologues. While it is possible that the ratio measure is compressed for dialogues vs. monologues because of the higher likelihood of shorter vocalizations for dialogues, we believe that the ratio measure has information to offer regarding the ease of language production in our study for several reasons. First, the conversation task used is one where longer turns are entirely appropriate to the extent that the content of what is produced is useful (i.e., having one person describe to another person where to go on a map). In this way, our communication task differs from normal conversation where there may not be as concrete a goal or topic, and interchanges may be more rapid and short. Second, the behavior of the ratio for dialogues alone shows that it responds in expected ways as a function of our independent variables, and in a similar way to monologues. Indeed, when we perform LME analyses

on the dialogues alone, we find a significant three-way interaction (language  $\times$  L2 proficiency  $\times$  inhibitory capacity interaction,  $t = -2.20$ ,  $p < 0.05$ ), suggesting that greater inhibitory capacity is associated with higher ratios for high L2 proficient bilinguals during L2 dialogues (see right panels of **Figure 7**). This effect is compatible with the monologue data where ratios were also higher as inhibitory capacity and L2 proficiency increased.

A final potential limitation concerns the independence of L2 proficiency and inhibitory capacity. Given prior work suggesting that bilinguals have better inhibitory capacity than monolinguals (reviewed in Bialystok, 2010), it is possible that bilinguals with high L2 proficiency might have greater inhibitory capacity than bilinguals with low L2 proficiency, by definition. This, in turn, would complicate our interpretation of the results for each variable individually. However, contrary to this hypothesis, the correlation between L2 proficiency and inhibitory capacity in our sample was not significant ( $r = -0.16$ ,  $p = 0.31$ ), perhaps due to the fact that all of the bilinguals tested here had some minimal high level of L2 proficiency to be able to produce spontaneous speech in an L2 monologue or dialogue context. As well, even presuming a statistically reliable relationship between L2 proficiency and IC, the LME approach would have allowed us to statistically disentangle the relative contributions of each to some extent, as these two variables are not likely to be perfectly correlated.

To conclude, the findings reported here suggest that individual differences among bilinguals in L2 proficiency and inhibitory capacity significantly modulate bilingual language production in monologues and dialogues, consistent with predictions of the IC model (Green, 1998; Abutalebi and Green, 2007) and prior work using other production tasks (Linck et al., 2008). Thus, our results establish a link between inhibitory capacity and bilingual language production among bilinguals, which is consistent with recent views suggesting that being bilingual enhances cognitive function (Bialystok et al., 2004; Bialystok, 2009). Finally, this study represents a first attempt at developing semi-automated methods to investigate the temporal dynamics of bilingual language production during more naturalistic conditions, such as during spontaneous monologue and dialogue speech.

## ACKNOWLEDGMENTS

This research was supported by an NSERC Discovery Award (PI: Titone). The authors gratefully acknowledge additional support from CIHR for the Frederick Banting and Charles Best CGS, Master's Award (awarded to Pivneva), and the Canada Research Chairs Program (awarded to Palmer and Titone).

## REFERENCES

- Abutalebi, J., and Green, D. W. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.
- Alpert, M., Homel, P., Merewether, F., Martz, J., and Lomask, M. (1986). Voxcom: a system for analyzing natural speech in real time. *Behav. Res. Methods* 18, 267–272.
- Anderson, A. H., Bader, M., Bard, E. G., Boyle, E. H., Doherty, G. M., Garrod, S. C., Isard, S. D., Kowtko, J. C., McAllister, J. M., Miller, J., Sotillo, C. F., Thompson, H. S., and Weinert, R. (1991). The HCRC map task corpus. *Lang. Speech* 34, 351–366.
- Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412.
- Bates, D. M. (2005). Fitting linear models in R: Using the lme4 package. *R News* 5, 27–30.
- Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. New York: Cambridge University Press.
- Bialystok, E. (2009). "Effects of bilingualism on cognitive and linguistic performance across the lifespan," in *Streitfall Zweisprachigkeit – The Bilingualism Controversy*, eds I. Gogolin and U. Neumann (Wiesbaden: VS Verlag für Sozialwissenschaften), 53–67.
- Bialystok, E. (2010). Bilingualism. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 559–572.
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

- Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354.
- Bialystok, E., Luk, G., Peets, K. F., and Yang, S. J. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogn.* 13, 525–531.
- Blumenfeld, H. K., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257.
- Brown, E. K., and Miller, J. E. (1980). *The Syntax of Scottish English*. Final Report to SSRC (UK), Project No. 5152, London.
- Chaffe, W. (1980). “Some reasons for hesitating,” in *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, eds H. W. Dechert and M. Raupach (The Hague: Mouton), 85–90.
- Christoffels, I. K., and De Groot, A. M. B. (2004). Components of simultaneous interpreting: a comparison with shadowing and paraphrasing. *Biling. Lang. Cogn.* 7, 1–14.
- Colomé, À., and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 96–109.
- Costa, A., and Caramazza, A. (1999). Is lexical selection language specific? Further evidence from Spanish-English bilinguals. *Biling. Lang. Cogn.* 2, 231–244.
- Costa, A., Pickering, M. J., and Sorace, A. (2008). Alignment in second language dialogue. *Lang. Cogn. Process.* 23, 528–556.
- Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.
- De Groot, A. M. B. (2011). *Language and Cognition in Bilinguals and Multilinguals: An Introduction*. New York: Psychology Press.
- De Jong, N. H., and Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behav. Res. Methods* 41, 385–390.
- Dell, G. S., Chang, F., and Griffin, Z. M. (1999). Connectionist models of language production: lexical access and grammatical encoding. *Cogn. Sci.* 23, 517–542.
- Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., and Bruce, V. (1997). Face-to-face and video-mediated communication: a comparison of dialogue structure and task performance. *J. Exp. Psychol. Appl.* 3, 105–125.
- Favreau, M., and Segalowitz, N. S. (1983). Automatic and controlled processes in the 1st-language and 2nd-language reading of fluent bilinguals. *Mem. Cognit.* 11, 565–574.
- Fehrer, C., and Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers. *Folia Linguistica* 41, 37–72.
- Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *J. Mem. Lang.* 30, 210–233.
- Flege, J. (1999). *Age of Learning and Second-Language Speech*. Hillsdale, NJ: Lawrence Erlbaum.
- Garrod, S., and Pickering, M. J. (2004). Why is conversation so easy? *Trends Cogn. Sci. (Regul. Ed.)* 8, 8–11.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous speech/Frieda Goldman-Eisler*. London, NY: Academic Press.
- Gollan, T. H., Bonanni, M. P., and Montoya, R. I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology* 19, 278–287.
- Gollan, T. H., and Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 640–665.
- Gollan, T. H., Montoya, R. I., Cera, C., and Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Mem. Lang.* 58, 787–814.
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., and Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis. *J. Exp. Psychol. Gen.* 140, 186–209.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.
- Griffin, Z. M., and Ferreira, V. S. (2006). “Properties of spoken language production,” in *Handbook of Psycholinguistics*, 2 Edn, ed. M. J. T. M. A. Gernsbacher (London: Elsevier), 21–59.
- Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Res.* 18, 1279–1296.
- Hanulová, J., Davidson, D. J., and Indefrey, P. (2010). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Lang. Cogn. Process* 26, 902–934.
- Hartsuiker, R. J., and Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychol. (Amst.)* 128, 479–489.
- Hartsuiker, R. J., Pickering, M. J., and Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychol. Sci.* 15, 409–414.
- Henderson, A., Goldman-Eisler, F., and Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Lang. Speech* 9, 207–216.
- Hernandez, A. E., Martinez, A., and Kohnert, K. (2000). In search of the language switch: an fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang.* 73, 421–431.
- Ivanova, I., and Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychol. (Amst.)* 127, 277–288.
- Kempen, G., and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cogn. Sci.* 11, 201–258.
- Kootstra, G. J., van Hell, J. G., and Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: evidence from bilingual monologue and dialogue. *J. Mem. Lang.* 63, 210–231.
- Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kormos, J., and Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32, 145–164.
- Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol. (Amst.)* 128, 416–430.
- Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Biling. Lang. Cogn.* 9, 119–135.
- Kroll, J. F., Michael, E., Tokowicz, N., and Dufour, R. (2002). The development of lexical fluency in a second language. *Second Lang. Res.* 18, 137–171.
- Lambert, W. (1974). *Culture and Language as Factors in Learning and Education*. Bellingham, WA: Western Washington State College.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition* 14, 41–104.
- Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: The MIT Press.
- Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374.
- Lindsay, J. R. (1975). Producing simple utterances – how far ahead do we plan. *Cogn. Psychol.* 7, 1–19.
- Macafee, C. (1983). *Glasgow*. Amsterdam: John Benjamin.
- Macaulay, R. (1985). *The Narrative Skills of a Scottish Coal Miner*. Amsterdam: John Benjamin.
- Marian, V., Blumenfeld, K. H., and Kaushanskaya, M. (2007). The language proficiency and experience questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967.
- McFarland, D. H. (2001). Respiratory markers of conversational interaction. *J. Speech Lang. Hear. Res.* 44, 128–143.
- McMurray, B., Samelson, V. M., Lee, S. H., and Tomblin, J. B. (2010). Individual differences in online spoken word recognition: implications for SLI. *Cogn. Psychol.* 60, 1–39.
- Michael, E., and Gollan, T. H. (2005). “Being and becoming bilingual: individual differences and consequences for language production,” in *The Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York: Oxford University Press), 389–407.
- Petrides, M. (1998). *Specialised Systems for the Processing of Mnemonic Information Within the Primate Frontal Cortex*. Oxford: Oxford University Press.
- Pickering, M. J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190.
- Pinkham, A. E., and Penn, D. L. (2006). Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry Res.* 143, 167–178.
- Poulisse, N., and Bongaerts, T. (1994). 1st language use in 2nd-language production. *Appl. Linguist.* 15, 36–57.
- R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing (Version 2.10.1): R Foundation for Statistical Computing*. Available at: <http://www.r-project.org>
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., and Salmon, D. P. (2010).

- What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Biling. Lang. Cogn.* 13, 231–252.
- Segalowitz, N. (2010). *The Cognitive Bases of Second Language Fluency*. New York: Routledge.
- Segalowitz, N., and Hulstijn, J. (2005). “Automaticity in bilingualism and second language learning,” in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. B. B. de Groot (New York: Oxford University Press), 371–388.
- Shallice, T., and Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1405–1412.
- Shockey, K., Richardson, D. C., and Dale, R. (2009). Conversation and coordinative structures. *Top. Cogn. Sci.* 1, 305–319.
- Simon, J. R., and Ruddell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.
- Titone, D., Libben, M., Mercier, J., Whitford, V., and Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: the effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1412–1431.
- Towell, R., Hawkins, R., and Bazergui, N. (1996). The development of fluency in advanced learners of French. *Appl. Ling.* 17, 84–119.
- Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.
- Welkowitz, J., Bond, R. N., and Zelano, J. (1990). An automated-system for the analyses of temporal speech patterns – description of the hardware and software. *J. Commun. Disord.* 23, 347–364.
- Whitford, V., and Titone, D. (2012). Second-(language) experience modulates first- and second-language word frequency effects: evidence from eye movement measures of natural paragraph reading. *Psychon. Bull. Rev.* 19, 73–80.
- Wilson, M., and Wilson, T. (2005). An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* 12, 957–968.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 August 2011; accepted: 13 February 2012; published online: 16 March 2012.

Citation: Pivneva I, Palmer C and Titone D (2012) Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. *Front. Psychology* 3:57. doi: 10.3389/fpsyg.2012.00057

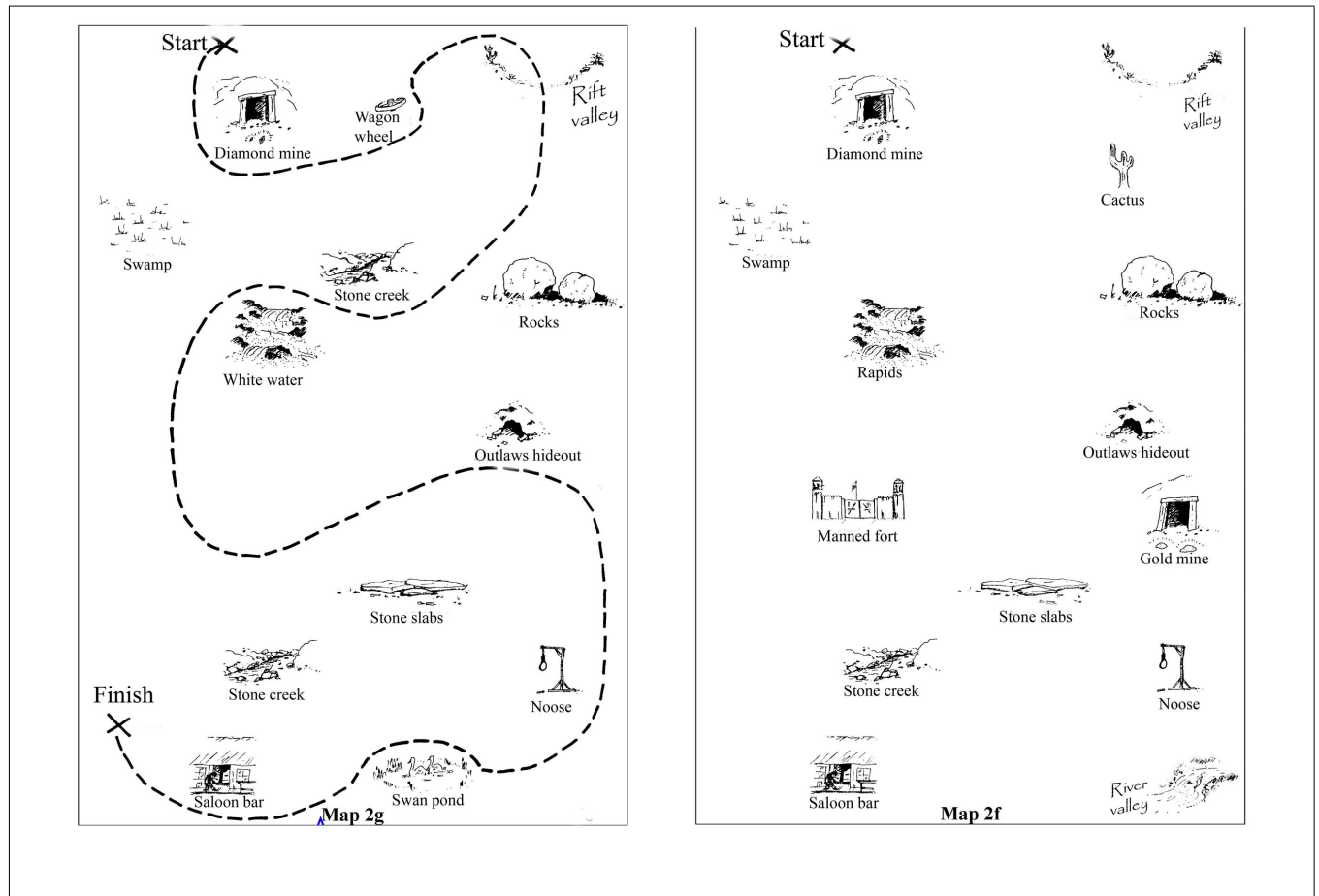
This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Pivneva, Palmer and Titone. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.



**APPENDICES**

**APPENDIX A**



**APPENDIX B****FE participant – L1 dialogue**

Participant: On va vers, euh... Vers la gauche, un peu. Donc en ligne droite, euh... A l'horizontale. Après, moi il y a l'eau vive. C'est une chute...

Confederate: Moi je touche les rapides. Est-ce que c'est correct? Moi j'ai les rapides ici.

Participant: Euh... Ben ça ressemble à des rapides sur ma photo mais ça s'appelle «eau vive», donc peut-être que c'est la même chose. Euh... On les contourne. Donc on passe en haut pour descendre le plus vers la gauche de la carte.

Confederate: Vers la gauche de la carte... OK.

Participant: On le contourne, oui. Vers le haut des... de l'eau vive pour descendre après ça. À gauche. Donc on descend quand même un peu, pour se rendre dans le dernier tiers de la carte, disons. Donc on descend à la verticale.

Confederate: A la verticale... OK.

(Translation)

Participant: We go towards, um... towards the left a little. So in a straight line... um... horizontally. After, I have white water. It's a fall.

Confederate: I touch the rapids. Is that OK? I have the rapids here.

Participant: Um... Well it looks like rapids on my picture but it's called "white water" So maybe it's the same thing. Um... so we go around them. So we pass above to go down to the left of the map.

Confederate: To the left of the map... OK.

Participant: We go around them, yes. Towards the top of the... the white water... To go down after that. On the left. So we still go down a little. To get to the last third of the map, let's say. So we go down vertically.

Confederate: Vertically... OK.

**EF participant – L1 dialogue**

Participant: And then do you have stone creek?

Confederate: Um... Yes, at the bottom of the page.

Participant: Um... No, OK. There is another one.

Confederate: OK.

Participant: Um... It's not far from the rocks but it's... It's basically, like, right at the center of the page. Where the rocks are but, like, towards the center. So you... You go under the stone creek...

Confederate: OK.

Participant: ... After the rocks. And then there is white water.

Confederate: OK. I have rapids...

Participant: OK. So... it's probably the same thing. And so you go over it and then...

**EF participant – L2 dialogue**

Participant: Après ça, on va se diriger comme dans une ligne diagonale allant vers le ruisseau des roches. Comme, à ce point-là, ça va être à ta droite.

Confederate: OK... Attends... Mon ruisseau des roches est comme vraiment en bas de la page.

Participant: Oui. Oh, OK! Non, non, non! Euh... le mien... le, le ruisseau de roche sur ma page, c'est comme... c'est à la même hauteur des roches, sauf c'est comme... Ils sont séparés de 3cm ou quoi.

Confederate: Ah OK OK! Moi j'ai des rapides qui sont vraiment un peu en bas. En bas des roches. C'est comme... de ruisseau de roche ou de ce que tu m'as dit... c'est comme entre les deux. C'est ça? Es que tu as des rapides?

Participant: J'ai des eaux vives. Ça a l'air des rapides.

(Translation)

Participant: After that, we are going to go in like, a diagonal line, going towards the stone creek. Like, at that point, it's going to be on your right.

Confederate: Ok... wait... My stone creek is like, really at the bottom of the page.

Participant: Yes. Oh, OK! No nono! Um... mine... The stone creek on my page. It's like... It's at the same height as the rocks. Except that it's like... They are separated by 3cm or something.

Confederate: Oh OK OK! I have rapids that are really a bit down. Below the rocks. It's like... From stone creek or from what you told me... It's like between the two. Right? Do you have rapids?

Participant: I have white water. It looks like rapids...

**FE participant – L2 dialogue**

Participant: You go towards the left of the sheet.

Confederate: Aha OK. So, I just go in a straight line?

Participant: In a straight line between the stone creek and the rocks.

Confederate: Stone creek? I only have a stone creek, like, at the bottom of the page. But not...

Participant: OK, well you go in a diagonal line, at the left of the rocks.

Confederate: At the left of the rocks... Like, how many centimeters am I away from the rocks?

Participant: um... 1.

Confederate: 1? OK, So I go diagonal. Like 45 degrees?

Participant: Um... Yeah.

Confederate: And where do I stop? At the rapids?

Participant: Um... not yet!

Confederate: Not yet, OK.

Participant: When you go down in a diagonal line for maybe about 5 cm.

Confederate: 1, 2, 3, 4, 5. OK.

Participant: And then you have to go at the top of the picture of the rapids.

APPENDIX C

