Entropy characterizes bilingual diversity

Characterizing the social diversity of bilingualism using language entropy

Jason W. Gullifer & Debra Titone

Department of Psychology, McGill University

Centre for Research on Brain, Language and Music, McGill University

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (individual Discovery Grant, 264146 to Titone); the Social Sciences and Humanities Research Council of Canada (Insight Development Grant, 01037 to Gullifer & Titone); the National Institutes of Health (Postdoctoral training grant, F32-HD082983 to Gullifer, Titone, and Klein); the Centre for Research on Brain, Language & Music.

The authors thank Ashley M. DaSilva, Denise Klein, Pauline Palma, Mehrgol Tiv, Janet Van Hell, Robert D. Vincent, Naomi Vingron, and Junyan Wei for fruitful discussions regarding (language) entropy.

Address for correspondence

Jason W. Gullifer, McGill University, Department of Psychology, 2001 McGill College Avenue, Montréal, QC H3A 1G1, Canada. E-mail: jason.gullifer@mail.mcgill.ca.

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Abstract

Bilingual and multilingual individuals exhibit variation in everyday language experience. Studies on bilingualism account for individual differences with measures such as L2 age of acquisition, exposure, or language proficiency, but recent theoretical perspectives posit that the relative balance between the two or more languages throughout daily life (i.e., INTERACTIONAL CONTEXT) is a crucial determinant for language representation, access, and control. We propose an innovative measure to characterize this construct by using ENTROPY to estimate the social diversity of language use. Language entropy is computed from commonly-collected language history data and generalizes to multilingual communicative contexts. We show how language entropy relates to other indices of bilingual experience and that it predicts self-report L2 outcome measures over and above classic measures of language experience. Thus, we proffer language entropy as a means to characterize individual differences in bilingual (and multilingual) language experience related to the social diversity of language use.

Keywords: bilingualism, language entropy, social diversity, individual differences, interactional context

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Characterizing the social diversity of bilingualism using language entropy

Bilingual and multilingual individuals vary widely in their exposure to, and socially distributed use of language, particularly if they live in highly multilingual locations. Individual differences in static language experiences, such as age of language acquisition (AoA; e.g., Flege, Munro, & MacKay, 1995; Gullifer et al., 2018; Kousaie, Chai, Sander, & Klein, 2017; Luk, De Sa, & Bialystok, 2011; Piske, MacKay, & Flege, 2001; Subramaniapillai, Rajah, Pasvanis, & Titone, 2018; Titone, Libben, Mercier, Whitford, & Pivneva, 2011), and current language experiences, such as amount of second language (L2) exposure (e.g., Gullifer et al., 2018; Hartanto & Yang, 2016; Hofweber, Marinis, & Treffers-Daller, 2016; Jylkkä et al., 2017; Pivneva, Mercier, & Titone, 2014; Prior & Gollan, 2011; Subramaniapillai et al., 2018; Titone, Gullifer, Subramaniapillai, Rajah, & Baum, 2017), drive linguistic performance and executive control abilities. However, there is not yet consensus on the best practices for measuring current language experience (Baum & Titone, 2014; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012; Surrain & Luk, 2017; Takahesu Tabori, Mech, & Atagi, 2018; Tomoschuk, Ferreira, & Gollan, 2018) despite its theoretical importance (Abutalebi & Green, 2016; Green & Abutalebi, 2013), in large part because bilingualism and bilingual experience are not homogenous constructs.

There is now a long history of measuring the impact of bilingual experience on language acquisition and processing, often through the use of self-report assessment instruments that tap into various constructs (see e.g., H. P. Bahrick, Hall, Goggin, Bahrick, & Berger, 1994; Li, Sepanski, & Zhao, 2006; Li, Zhang, Tsai, & Puls, 2014; Marian, Blumenfeld, & Kaushanskaya, 2007; McLaughlin, 1977). However, standard operating procedures at present generally involve the computation of language proficiency along a single dimension (e.g., L2 proficiency). Few studies, particularly in cognitive domains, assess the impact of other background measures

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

commonly elicited by standardized language history questionnaires, such as daily exposure to known languages or daily language use in various communicative contexts (see e.g., critiques raised by Gollan et al., 2012; Surrain & Luk, 2017; Tomoschuk et al., 2018). Moreover, few researchers treat language experience measures in a continuous manner (see e.g., critiques raised by Baum & Titone, 2014), opting instead to dichotomize continuous variables into discrete groups. Although grouping may be warranted in some cases, such as when two distinct populations are being compared to address a particular question, it can be problematic in others, leading to a loss of information and poor statistical estimates, particularly when the phenomena under investigation exist on a continuum (MacCallum, Zhang, Preacher, & Rucker, 2002). One likely reason for these practices is that the sheer number of variables characterizing individual differences in L2 experience can be overwhelming and highly correlated, thus, creating challenges for researchers searching for a single dimension upon which to focus.

Here, we hope to overcome some of these limitations by offering an innovative way to characterize current language experience by capitalizing on measures often assessed but not evaluated, which when combined with other data reduction techniques like principal component analysis, can yield a manageable set of individual differences variables. This approach involves a measure of LANGUAGE ENTROPY that indexes the relative balance or diversity in the daily usage of two or more languages. Higher entropy values relate to more balanced language use and greater language diversity.

Practically, language diversity can vary among global locations; among communicative contexts within a location; and, crucially, among individuals. For example, while some geographic areas house a bilingual population, the bilinguals may tend function in a COMPARTMENTALIZED fashion, restricting use of specific languages to specific communicative contexts and avoiding language mixing. In other areas, bilinguals may tend to function in an

INTEGRATED fashion, where all languages are used regardless of the communicative context and where language mixing (either across or within utterances) is common (see e.g., Beatty-Martínez & Dussias, 2017). Integrated bilingual language usage is the reputation of Montréal, the site of the present study (Heller, 1982; Higgins, 2004; Lamarre, 2002). The multilingual nature of Montréal is apparent to any resident or visitor and is generally supported by language demographic data from the Canadian census (Statistics Canada, 2017).

<Insert Figure 1 about here>

On the census, when Montréalers are asked to identify their "most commonly used language(s)", a vast majority report using French across communicative contexts (approximately 60-70% of respondents depending on the context), but a fair percentage also report the primary use of two or more languages (approximately 4-11%) in these contexts (proportionally more than Canada generally; 2-5%). These proportional data are illustrated in Figure 1A. In Figure 1B we illustrate the same proportional data as language entropy, which gives a sense of how language diversity varies by communicative context and geographic area. Across all contexts, Montréal exhibits higher entropy, and thus higher diversity than Québec, largely because of the dominance of the French language in Québec. Interestingly, Montréal has equal or lower language diversity relative to Canada as a whole for home-related contexts (i.e., mother tongue and primary home language). However, for the work-related context, Montréal exhibits substantially higher language diversity, indicative of more integrated usage in this context.

We note two caveats here that may result in the underestimation of language diversity in Montréal as illustrated. First, the census lacks questions about other communicative contexts that may reveal higher language diversity (e.g., social settings). Second, when Montréalers are asked about "other commonly used languages" (besides the most commonly used language(s); not illustrated here), English in Montréal experiences a significant boost in usage relative to other languages in the broader Canadian context. Together, these data support the notion that Montréal is highly bilingual and that there is substantial variability in language diversity among geographic locales and among different communicative contexts. Moreover, language entropy can provide a concrete estimate of diversity while simultaneously reducing the complexity of the data, making data visualization and modeling more tractable. Importantly, variability in language diversity holds theoretical implications for behavior, brain structure, and brain function (Abutalebi & Green, 2016; Beatty-Martínez & Dussias, 2017; Green & Abutalebi, 2013; Gullifer et al., 2018; Hartanto & Yang, 2016; H. Yang, Hartanto, & Yang, 2016; Poarch, Vanhove, Berthele, 2018), and our recent work shows that this variability is measurable at the level of the individual (Gullifer et al., 2018).

Recent psycholinguistic and neurocognitive perspectives of bilingualism, such as the adaptive control hypothesis (Abutalebi & Green, 2016; Green & Abutalebi, 2013), predict that individual differences in how bilinguals use their two or more languages across different social settings (i.e., INTERACTIONAL CONTEXT) are critically important in determining how bilinguals represent, access, and control those languages. For example, bilinguals who use their two languages in primarily single language contexts (i.e., compartmentalized bilinguals) are predicted to have different language and executive control demands relative to bilinguals who use their two languages in dual language contexts (i.e., integrated bilinguals). We note that while novel, this focus is a successor in the spirit of seminal work on language mode by Grosjean (1997; 2001). Accordingly, recent work has begun to examine the social diversity of language use in relation to executive control capacity and language processing by sampling various groups of participants from locations that are known to differ in the social diversity of language use (e.g., Beatty-Martínez & Dussias, 2017) or by computing difference scores between self-report usage measures for the first language (L1) and second language (e.g., Birdsong, Gertken, &

Amengual, 2012; Hartanto & Yang, 2016, Poarch et al., 2018). While innovative, such approaches are limited in the ability to assess nuanced differences in the social diversity of language usage that likely exists at the level of the individual within a population or that arise from situations where people use more than two languages. Thus, despite the practical and theoretical significance of the social diversity of language use among bilinguals, the field has not converged upon an optimal way of characterizing this source of individual variation (see also Fricke, Zirnstein, Navarro-Torres, & Kroll, 2019; Kroll, Dussias, & Bajo, 2018; Takahesu Tabori et al., 2018).

We propose language entropy as an innovative way to measure individual differences in the social diversity of language use, including the interactional context of language usage. Entropy is a concept with its roots in physics: a property of physical systems that is proportional to number of different configurations, or states, of those systems. The concept was adapted for information theory by Claude Shannon as a means to quantify information content or uncertainty (Shannon, 1948). Entropy has been used previously in psycholinguistics to quantify lexical and syntactic complexity (e.g., del Prado Martín, Kostić, & Baayen, 2004; Hale, 2003; Levy, 2008), and we recently applied language entropy to study the neurocognition of bilingualism using resting state functional connectivity (see Gullifer et al., 2018). Entropy is useful psychometrically as it yields a continuous measure of diversity, and it is computed as a function of the probability with which a set of events or states occur. Information about individuals' language use or exposure within various communicative contexts (e.g., language use at home, work, in social setting, etc.) is frequently elicited by the standard language history questionnaires that have become ubiquitous within the field (Anderson, Mak, Chahi, & Bialystok, 2018; Birdsong, et al., 2012; Dunn, & Fox Tree, 2009; Li et al., 2014; Marian et al., 2007), and these data are inherently proportional (or, in the case that data are collected via Likert scale, can be

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

converted to a proportion). Thus, language entropy can be straightforwardly assessed for each communicative context, or globally across all contexts. We provide a fully-documented R package that computes language entropy (Gullifer & Titone, 2018), available at: https://github.com/jasongullifer/languageEntropy.

Language entropy values range from 0 to *log n*, where *n* is the number of languages that entropy is computed over (e.g., with two languages, max entropy is 1, with three it is approximately 1.585). A communicative context that is completely compartmentalized, where an individual reports using only one language, will have an entropy of 0, signifying no language diversity and very high predictability of an upcoming language within this context (i.e., a single-language interactional context). In contrast, a communicative context that is completely integrated, where two or more languages are used in perfect balance, will have maximum entropy, signifying high language diversity and very low predictability of an upcoming language within this context. Language entropy ranges continuously between 0 and the maximum[1], allowing for the assessment of language use that falls in between compartmentalized and integrated dual language contexts.

Thus, the goal of this paper is to establish a proof of concept, where we demonstrate the utility of using entropy as an estimate for the social diversity of language use in two ways. First, we show that entropy exhibits substantial variation across speakers and communicative contexts. Second, we address a classic question in L2 acquisition and bilingualism, namely, what factors predict self-perceived L2 accentedness and L2 abilities (e.g., similar to work by Flege et al., 1995). To this end, we analyze data from a large sample of bilinguals/multilinguals (N = 507) drawn from the highly multilingual city of Montréal, QC, Canada, and we replicate a classic finding using self-report data: that L2 AoA and L2 current exposure predict self-perceptions of L2 accentedness and L2 abilities. Crucially, we show that the social diversity of language use exhibits additional predictive power over these classic predictors for both outcome measures. Moreover, for L2 accentedness ratings, the impact of social diversity interacts with L2 AoA. Together, these results suggest that the social diversity of language usage is an important variable for future studies to consider. In the discussion we highlight potential applications for future research, guided by the adaptive control hypothesis.

Method

Participants

We analyzed language history data collected in the McGill Language and Multilingualism (MLL) lab over several years. From 2008 to 2015, approximately 507 bilingual or multilingual participants were tested who reported detailed language history information (including the relative exposure to and use of two or more languages). We report a qualitative analysis of participant characteristics in the results section below.

Materials

All participants in this sample completed a language background questionnaire adapted from the LEAP-Q (Marian et al., 2007) or LHQ 2.0 (Li et al., 2014), allowing us to probe language usage within the Montréal context.

Language history information. For the purposes of the analysis and entropy computation, we extracted several types of background measures outlined below. We extracted basic demographic information, including L2 AoA and L2 exposure, which served as predictors in the analysis. We extracted data on “language exposure in different usage contexts” for the purposes of computing language entropy, which served as another class of predictors in the analysis. Finally, we extracted self-reported L2 accent perception and L2 abilities, which served as outcome measures in the analysis. For some of these extracted measures, we applied data

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

reduction techniques (i.e., aggregation or principal component analysis). A summary of the data, including aggregate measures or components, is available in Table 1.

<Insert Table 1 about here>

Basic demographic information. In the questionnaire, participants reported basic information about their demographics and language use. We extracted classic measures of L2 exposure, such as L2 AoA (based on the onset of learning) and global exposure to the L1, L2, and L3 (third language). Global L2 exposure is frequently used as a covariate in the MLL lab (e.g., Pivneva et al., 2014; Subramaniapillai et al., 2018), and these global exposure measures did not factor in to the computation of language diversity to allow for a comparison of the measures.

Language exposure in different usage contexts. Participants reported the extent to which they used the L1, L2, and L3 in a variety of communicative contexts in the home, at work, in social settings, for reading, and for speaking. The questionnaire elicited language use at home, work, and in social settings via Likert scales (e.g., "Please rate the amount of time you use each language at home"), with a score of 1 indicating "no usage at all" and a score of 7 indicating "usage all the time". We baselined Likert data at 0 by subtracting 1 from each response. Thus, a value of 0 reflects "no usage at all." We converted these data to proportions of usage by dividing a given language's score by the sum total of the scores within context. For example, a participant who reported (after baselining) the following data for language usage at home, L1: 6, L2: 5, L3: 0, would receive the following proportions for the home context, L1: 6/11, L2: 5/11, L3: 0/11.

Language use for reading and speaking were collected through percentage of use ("What percentage of time would you choose to speak each language?"), which totaled to 100% within a particular context. We converted percentages to proportions, and we used this proportional usage data to compute the diversity of language use in each context.

L2 accent perception. Participants reported the extent to which they believe they have an accent in the L2 and the extent to which they believe others identify them as nonnative speakers based on their accent. This information was elicited through two questions using seven-point Likert scales ("How much of a foreign accent do you have in L2?", 1 indicates *no accent*, 7 indicates *a strong accent*; and "Please rate how frequently others identify you as a non-native speaker based on your ACCENT in French", 1 indicates *never*, 7 indicates *all the time*). Across the sample, the self-accent perception and other accent perception were positively correlated (Spearman rho: 0.79, $p < 0.05$), and we computed a mean accent perception score (*M*: 3.62, *SD:* 1.79). We used this score as a dependent variable in the analyses.

L2 abilities. Participants answered a series of 20 questions that probed self-rated abilities in the L2 and L1, including speaking, reading, writing, translating, listening, pronunciation, fluency, vocabulary knowledge, grammatical knowledge, and overall competence. To reduce the complexity of these data, we conducted a principal components analysis (three components, determined via scree plot, with oblimin rotation) using the *psych* package {Revelle:2017um} for R (R Core Team, 2017). Variables related to the L2 loaded onto one component, and this component explained 39% of the variance in the data. Variables related to the L1 loaded onto two other components, and each component explained 15% of the variance in the data. We extracted the scores for the L2 component to serve as an index of self-rated L2 abilities, and we used this as a dependent variable in the analysis.

Computing language entropy. For each usage context (see "Language exposure in different usage contexts", above), we computed Shannon entropy (H) using the following equation $H = -\sum_{i=1}^{n} P_i log_2(P_i)$ and the methods available in the languageEntropy R package (Gullifer & Titone, 2018). Here, *n* represents the total possible languages within the context (i.e., 3) and P_i is the proportion that language_i is used within a context. To illustrate, if hypothetical bilingual reports using French 80% of the time and English 20% of the time within the work context, one computes language entropy by summing together $0.80 * log_2(0.80)$ and $0.20 * log_2(0.20)$, then multiplying by -1 to yield a positive entropy value. Thus, the hypothetical individual's language entropy in the work context would be approximately 0.72.

Theoretically, the entropy distribution has a minimum value of 0 that occurs when the proportion of usage for a given language is 1.0, representing a completely compartmentalized context. The distribution has a maximum value equal to log *n* (approximately 1.585 for three languages) when the proportion of use for each language is equivalent, representing a completely integrated context.

Our procedure resulted in five entropy scores for each participant, that pertained to language entropy for home, work, social, reading, and speaking. To reduce complexity of these data, we conducted a principal components analysis on the entropy data (two components, determined via scree plot, with oblimin rotation). The first component comprised reading, speaking, home, and social entropy, and this component explained 44% of the variance in the data. The second component comprised work with some cross-loading from social entropy, and this component explained 21% of the variance in the data. We extracted the component scores for each participant to serve as indices of language entropy at work and language entropy everywhere else.

Results

The data were prepared, plotted, and analyzed in R (R Core Team, 2017) using *tidyverse* (Wickham, 2017).

Qualitative analysis of participant characteristics. First, we offer a qualitative analysis of participant characteristics related to static historical language experience (i.e., L2 AoA) and current language experience (e.g., L2 exposure and language entropy), reported in Table 1.

In terms of static historical language experience, of the sampled 507 participants, we identified 51 individuals as simultaneous bilinguals, who acquired the L2 at or before 1 year of age, and 456 individuals as sequential bilinguals. Thus, the majority of the sample reported acquiring their L2 before the age of 15 (N = 494). Based on L2 AoA, we identified 240 participants as native English speakers, and 267 participants as native French speakers.

In terms of current language experience across the sample, participants reported, on average, being exposed to the L2 for roughly one third of the day. However, there was substantial variation in this measure indicated by the high standard deviation (M = 33.82, SD = 20.67). Participants reported, on average, minimal daily exposure to an L3 (M = 2.38, SD = 5.58). A qualitative analysis of language entropy measures suggests that, overall, participants were relatively integrated in terms of their bilingualism (see Table 1 for means and standard deviations). For example, 80-20 bilingualism would be reflected in an entropy value of 0.72 (approximately the average language entropy across contexts). Again, there was substantial variation across individuals within contexts, indicated by the large standard deviation. While not reported in the table, participants spanned the whole range of the entropy spectrum from completely compartmentalized (entropy: 0) to fully integrated (entropy: 1.585). There was also variation in language entropy across usage contexts, in that entropy was higher and less variable in social and work contexts relative to other contexts, with for example 65-35 bilingualism reflected in an entropy value of 0.94. We note that L2 AoA was generally not highly correlated with the entropy measures (range of Spearman rho: -0.17 – 0.03), though L2 exposure was (range: 0.13 – 0.44). See Supplemental Figure S1 for an illustration of the bivariate correlations between variables related to language experience, variables related to language entropy, and language entropy components. The results of this qualitative analysis pattern well with the components identified in the principal components analysis above, namely, that the work context loads onto a separate component with some cross-loading from the usage context. See Table 1 (Language entropy) for descriptive statistics.

Quantitative analyses assessing the utility of language entropy as a predictor. We used multiple linear regression to predict mean L2 accent ratings and scores on the L2 abilities component as dependent variables. For each dependent measure, we fit three nested models and selected the best model via a model comparison procedure. The first model (base model) included the L2 AoA and L2 exposure as fixed effects. The second model (additive model) additionally included non-work entropy and work entropy components. The third model (interaction model) additionally included all two-way interactions with L2 AoA (i.e., L2 exposure * L2 AoA, non-work entropy * L2 AoA, and work entropy * L2 AoA). Model comparisons using chi-squared tests assessed whether the addition of entropy measures and interactions with L2 AoA significantly improved model fit. We then examined the best-fitting models to assess the direction, magnitude, and significance of each fixed-effect slope estimate. All predictors in the models were centered and standardized.

Given the relatively few individuals in the sample who reported an L2 AoA greater than 15 years, it was possible that our results could be driven by the presence of outliers or high leverage points. Consequently, we conducted additional analyses using robust linear regression with Huber weights to attenuate the influence of observations with high residuals, allowing us to minimize undue influence from outliers while maintaining the continuous nature of the L2 AoA measure without removing participants from the sample. The general pattern of results did not change under robust linear regression, suggesting that the patterns are stable.

L2 accentedness. For the L2 accentedness model, the addition of entropy components improved model fit ($\chi^2(2) = 56.62$, $p < 0.05$), and the addition of two-way interactions further improved model fit ($\chi^2(3) = 30.47$, $p < 0.05$). Thus, the two entropy components explained unique variance relative to L2 AoA and L2 exposure[2], and there were significant interactions between L2 AoA and other predictors (see Table 2 for model outputs).

<Insert Table 2 about here>

Inspection of the interaction model (adjusted $R^2 = 0.30$, Intercept = 3.627, *SEM* = 0.067, $t(499) = 54.239$, $p < 0.05$, 95% CI [3.496, 3.758]) showed a main effect of L2 AoA ($b = 0.695$, *SEM* = 0.068, $t(499) = 10.184$, $p < 0.05$, 95% CI [0.562, 0.829]), indicating that later L2 AoA was associated with higher L2 accentedness ratings. There was a main effect of L2 exposure ($b = -0.445$, *SEM* = 0.072, $t(499) = 6.138$, $p < 0.05$, 95% CI [-0.587, -0.303]), indicating more L2 exposure was associated with lower L2 accentedness ratings. There was a main effect of non-work entropy ($b = -0.323$, *SEM* = 0.077, t(499) = 4.222, $p < 0.05$, 95% CI [-0.473, -0.173]), indicating that higher non-work entropy (i.e., more integration) was associated with lower L2 accentedness ratings. There was no significant main effect of work entropy ($b = -0.078$, SEM = 0.071, $t(499) = 1.112$, $p > 0.05$, 95% CI [-0.217, 0.060]). There was an interaction between L2 AoA and L2 exposure ($b = 0.142$, *SEM* = 0.061, $t(499) = 2.338$, $p < 0.05$, 95% CI [0.023, 0.261]), indicating that exposure effects were of greater magnitude at earlier L2 AoA. There was an interaction between L2 AoA and non-work entropy ($b = 0.169$, *SEM* = 0.072, $t(499) = 2.339$, $p < 0.05$, 95% CI [0.027, 0.310]), indicating that non-work entropy effects were of greater magnitude at earlier L2 AoA. There was an interaction between L2 AoA and work entropy ($b = -0.186$, *SEM* = 0.076, $t(499) = 2.437$, $p < 0.05$, 95% CI [-0.336, -0.036]), indicating higher work entropy was associated with lower ratings of L2 accentedness at later L2 AoA. See Figure 2 for the partial effects plots from this model.

<Insert Figure 2 about here>

L2 abilities. For the L2 abilities model, the addition of entropy components improved model fit ($\chi^2(2) = 36.52$, $p < 0.05$), but the addition of two-way interactions did not further improve model fit ($\chi^2(3) = 0.83$, $p > 0.05$). Thus, the two entropy components explained unique variance relative to L2 AoA and L2 exposure[2] and there was no evidence for interactions with L2 AoA (see Table 3 for model outputs).

<Insert Table 3 about here>

Inspection of the additive model (adjusted $R^2 = 0.40$, Intercept = 0.000, *SEM* = 0.035, $t(502) = 0.000$, $p = 1.00$, 95% CI [-0.067, 0.067]) showed that there was a main effect of L2 AoA ($b = -0.336$, *SEM* = 0.035, $t(502) = 9.681$, $p < 0.05$, 95% CI [-0.404, -0.268]), indicating that later L2 AoA was associated with lower L2 ability ratings. There was a main effect of L2 exposure ($b = 0.359$, *SEM* = 0.037, $t(502) = 9.651$, $p < 0.05$, 95% CI [0.286, 0.432]), indicating more L2 exposure was associated with higher L2 ability ratings. There was a main effect of non-work entropy ($b = 0.257$, *SEM* = 0.039, $t(502) = 6.525$, $p < 0.05$, 95% CI [0.180, 0.334]), indicating that higher non-work entropy (i.e., more integration) was associated with higher L2 ability ratings. There was a main effect of work entropy ($b = 0.075$, *SEM* = 0.036, $t(502) = 2.061$, $p < 0.05$, 95% CI [0.004, 0.146]), indicating that higher work entropy (i.e., more integration) was associated with higher L2 ability ratings. See Figure 3 for the partial effects plots from this model.

<Insert Figure 3 about here>

Discussion

Here, we applied the concept of entropy to formalize an index of social diversity of language use that would serve as a novel indicator of individual differences in current language experience. This measure, language entropy, allows for the continuous estimation of compartmentalized to integrated language use within and across bilingual communicative contexts, with higher entropy values corresponding to more integrated language usage. We computed language entropy for various contexts for a large sample (N = 507) of bilingual/multilingual speakers of French and English drawn from the population of Montréal. Our first, qualitative analysis provides support for the idea that bilingual language use in Montréal is highly diverse, particularly in social and work contexts, in line with common knowledge and available data the Canadian census in Montréal (Statistics Canada, 2017). Next, we showed that language entropy is an important predictor for self-perceived markers of L2 success, namely self-perceived L2 accentedness and L2 abilities, in line with recent theoretical perspectives on the neurocognition of bilingualism which state that the context of language usage is an important determinant of the way in which bilinguals represent, access, and control the two languages (Abutalebi & Green, 2016; Green & Abutalebi, 2013).

Our qualitative analysis suggests that while our sample was homogeneous in the sense that all individuals learned their L2 during early childhood or adolescence (the majority consisting of relatively early bilinguals who acquired the two languages before the age of 15), the population was quite heterogeneous in terms of daily usage of their languages across communicative contexts. Overall, entropy values indicated a general pattern of integrated language use, but there was variation across individuals within communicative contexts, suggesting that the people in our sample differed in whether they used their languages in a compartmentalized or integrated manner. Moreover, communicative contexts were also variable in terms of language entropy, with social and work contexts exhibiting higher language entropy (i.e., diversity), indicative of more integrated language use than the other contexts, notably the home context. This finding is compatible with data from the Canadian census in Montréal (Figure 1), showing lower language diversity in home contexts than in work contexts. The principal component analysis on entropy measures for each of the five communicative contexts further suggested that there may be two components underlying the social diversity of language use (at least for our sample drawn from Montréal): first, a more global language diversity (the non-work entropy component), and second, language diversity in the work context (the work entropy component).

Interestingly, the qualitative analysis also showed that language entropy values were not highly correlated with L2 AoA. This finding suggests that regardless of whether an individual acquired their second language early or later in young adulthood, they could come to use their languages in either a compartmentalized or integrated manner. More research is needed to determine whether this is also the case for individuals who acquire a second language later in life.

Crucially, we demonstrated the utility of the social diversity of language use as estimated by language entropy as a predictor by testing whether individual differences in the two entropy components (i.e., work and non-work) predicted self-report data for L2 accentedness and L2 abilities over and above classic predictors of language experience: L2 AoA and L2 exposure. Our models indicated that early L2 AoA and high L2 exposure were associated with decreased L2 accentedness and increased L2 abilities ratings, consistent with previous work suggesting that the amount of language exposure (whether early or otherwise) and is an important predictor of L2 success (e.g., Flege et al., 1995). Importantly, we also found that both language entropy components improved model fit over and above L2 AoA and L2 exposure.

The addition of the entropy components explained unique variance related to L2 accentedness and L2 abilities. Although the R-squared improvement was modest, the effects were significant, and they patterned in similar and sensible directions across the two models. Specifically, higher scores on the entropy components related to decreased L2 accentedness ratings and increased L2 abilities ratings, suggesting that integrated language use is associated with self-report outcome measures traditionally thought to reflect L2 success (but see below for limitations) apart from the classic measures of L2 AoA and L2 exposure. For accentedness ratings, there were further interactions between static historical language experience (L2 AoA) and indices of current experience (L2 exposure and the entropy components), suggesting that the greatest impact of ongoing language experience may depend on the specific communicative context in which the language is acquired or most frequently used. For example, early bilinguals tend to acquire both languages in the home or in school, which were the environments that were most predictive for early bilinguals in our sample (represented by scores on our general entropy component). In contrast, late bilinguals may tend to acquire and use the L2 with associates at work, a context that tends to be more linguistically diverse overall in Montréal and that was the most predictive for late bilinguals in our sample. Overall, these results provide a first step in validating language entropy as a measure that is predictive of variance related to bilingual language representation.

One might argue that although the language entropy components explain additional variance in the models presented here, the associations between the components and the outcome measures were highly similar to that of L2 exposure (i.e., more language exposure is associated with better language outcomes), obviating the calculation of more complex measures such as language entropy. In our view, the fact that language entropy patterns well with L2 exposure but explains unique variance for a large sample of participants functions to validate the measure psychometrically and in relation to other known constructs. In future studies of bilingualism and bilingual language processing that employ different methodologies (e.g., online language processing tasks, executive control tasks, etc.), language entropy may pattern differently than global L2 exposure (e.g., see predicted differences between interactional contexts in Abutalebi & Green, 2016; Green & Abutalebi, 2013). To the extent that language entropy is an important predictor in these future studies, this work will be foundational in providing confidence that it is the construct of relative language balance behind language entropy explaining variance and not some other confounded variable. Furthermore, although entropy required more up-front feature engineering compared to simply using L2 exposure, it may actually yield simpler model specification and interpretation. For example, in the case that language balance is an important predictor theoretically (e.g., if some experimental effect peaks at 50-50 balanced bilingualism), one could model this effect by adding an additional nonlinear effect of L2 exposure, using more degrees of freedom and potentially making interpretation less straightforward. The case becomes even more complex if the researcher is interested in language diversity among multilinguals as opposed to bilinguals. In sum, language entropy can be an efficient way to model language balance and diversity among bilinguals and multilinguals.

Of note, we wish to be careful in associating either of the outcome measures here (L2 accentedness and L2 abilities) strongly with objective or true language proficiency for at least two reasons. First, the way in which we assessed L2 accentedness and abilities was through self-report, a method that has been scrutinized recently because self-perceptions can be tainted by often stigmatized aspects of language usage (Gollan et al., 2012; Surrain & Luk, 2017; Tomoschuk et al., 2018). Second, even when these factors are measured objectively, through for example a language production task, they can be subject to a variety of psycholinguistic and sociolinguistic influences that may be unrelated to language proficiency per se. For example, phonetic productions are subject to cross-language competition (e.g., Goldrick, Runnqvist, & Costa, 2014) and phonetic convergence (e.g., Pardo, 2006). That being said, the fact that the results pattern well with studies with arguably more objective data (i.e., accent ratings made on production data by research assistants; e.g., Flege, et al., 1995) suggests some validity in the self-report measure in tapping into the construct with our sample of speakers. Crucially, we would argue that the language experience variables that we included as predictors here (e.g., L2 AoA, L2 exposure, and language use in various contexts) are less likely to be subject to the limitations

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

associated with self-report, as they constitute self-report measures about relatively objective facts about language use, such onset and amount of language use, without tapping into subjective, self-evaluative feelings about one's own language ability relative to other speakers. Moreover, we have recently demonstrated that language entropy is related to functional brain organization and executive control abilities within a similar sample of bilingual speakers (Gullifer et al. 2018), providing further support for the efficacy of these variables as predictors of objective outcome measures.

A limitation of the present results that influences generalizability is related to the information that we did not elicit on the language history questionnaire, the lack of which could bias our estimates of language entropy. For example, estimates of entropy here could be biased depending on the amount of time an individual spends in each communicative context. if a hypothetical participant spends most of their daily hours in one integrated context and a minority of their time in several other compartmentalized contexts, then an aggregate measure of entropy for that participant (e.g., their component score across multiple contexts) would underestimate the extent to which that participants entropy score reflects their true integratedness. When the amount of time spent in each communicative context is available, this information could be used to weight entropy scores accordingly. Moreover, we did not elicit information about code-switching behavior, another interactional context predicted to be of critical importance by instantiations of the adaptive control hypothesis (Abutalebi & Green, 2016; Green & Abutalebi, 2013; Green & Wei, 2014). Nor did we elicit information about translation experience, language brokering, or other important communicative contexts that may come to influence language and executive control (e.g., Dong & Liu, 2016; Dong & Zhong, 2017; López & Vaid, 2018). Thus, at this stage more work is needed to determine the generalizability this measure is to other populations of interest. Crucially, entropy is a general measure that provides a measure of information related to states in a system and it could be easily adapted to apply to other communicative contexts / language usage environments where questions of diversity are at issue. Our recommendation is that future work should investigate these issues by eliciting information about contexts that are relevant to the population(s) of interest (e.g., language brokering for heritage populations) to ensure accurate estimates for that population.

An alternative explanation for the importance of the entropy as a predictor is that it accounts for unique variance simply because it incorporates information about L3 usage (a substantial portion of the sample reported knowledge of an L3), whereas L2 exposure and L2 AoA do not. To test this hypothesis, we conducted additional model comparisons on a set of models that included a measure of non-L1 exposure (i.e., the sum of L2 and L3 exposure), and the pattern of results remained highly similar to those reported above. Specifically, for the L2 accentedness model, the model with the interactions remained the best-fitting model ($\chi^2(3) = 29.63$, $p < 0.05$) and the direction and significance of the estimates of that model were the same as reported above. For the L2 abilities model, the model with additive effects of the entropy components remained the best-fitting model ($\chi^2(2) = 26.17$, $p < 0.05$). The direction of all estimates were the same as reported above, though slope of the work entropy component was no longer significantly different from 0 ($b = 0.053$, $SEM = 0.037$, $t(502) = 1.417$, $p = 0.16$, 95% CI [-0.020, 0.125]), suggesting that some variance in this component was accounted for by the inclusion of L3 exposure for this model. In sum, the importance of the non-work entropy component in both models and the work entropy component in the L2 accentedness model cannot be completely attributed to the mere fact that they incorporate additional information about L3 usage.

Instead, language entropy provides information about the relative balance in language usage. Alternatively, one could think of language entropy as providing information related to the number of different "language states", or specific configurations of language use, that a bilingual might find themselves in over the course of daily life. Compartmentalized bilinguals with low language entropy across their communicative contexts experience few unique language states and may be relatively certain about which particular language will occur given the communicative context. In contrast, integrated bilinguals with high language entropy across their communicative contexts experience a greater number of language states and may have decreased certainty about when a particular language will occur at any given time.

Thus, throughout daily life, compartmentalized and integrated bilinguals (and those in between) may experience different rates of competition between their languages and may come to control that competition languages in different ways. For example, while compartmentalized bilinguals may experience competition between their languages, they may become adept at inhibitory control to suppress competition from the irrelevant language. In contrast, highly integrated bilinguals may actually benefit from cross-language activation, allowing them to flexibly switch between their different languages as the context demands, and indeed suppressing a language may lead to disfluency if that language is required. These are general predictions made by the Adaptive Control Hypothesis (Abutalebi & Green, 2016; Green & Abutalebi, 2013) which posits that an individual's cognitive system in charge of language representation and control adapts to meet the demands of one's interactional context (i.e. situations which reflect particular patterns of use such as contexts that involve the use of two languages vs. one language). Thus, the social diversity of language use, characterized using entropy, appears to be a fruitful way to measure interactional context among individuals.

The present data were collected on a large sample of bilingual/multilingual individuals living in a diverse linguistic environment. An important question is the extent to which language entropy generalizes to other populations of bilinguals and monolinguals. In our view language

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

entropy is a general measure of diversity that should nicely characterize speakers in other locales as well to the extent that critical information about language usage in relevant communicative contexts is available for or can be elicited from these individuals in a non-stigmatized manner (e.g., for individuals living in locations where bilingualism is not valued). Thus, future research should investigate the extent to which groups of speakers who live in locales with different interactional contexts (e.g., primarily single language contexts vs. dual language contexts) exhibit differences in language entropy.

In order to facilitate future research and compute language entropy, researchers can use the methods available in the languageEntropy package (Gullifer & Titone, 2018). The package includes instructions, examples, and help files for each of the provided functions. We offer a final pragmatic note to researchers who might try using language entropy to index language diversity. It would be fruitful to investigate the correlations between entropy measures for different contexts for the sample in question. In our sample, the majority of the entropy measures in each context patterned well together with the exception of work entropy, yielding two principal components. However, this may not be the case with every sample, and we would caution researchers against applying the same criteria on a different sample from a different location.

Characterizing and quantifying individual differences related to language entropy may have implications for debates within the field of bilingualism, such as whether bilingual (relative to monolingual) experience leads to changes compared to monolinguals in general cognitive capacities (Bialystok, Craik, Klein, & Viswanathan, 2004; de Bruin, Treccani, & Sala, 2015; Paap, Johnson, & Sawi, 2015; Takahesu Tabori, Mech, & Atagi, 2018; Titone et al., 2017). Many studies weighing in on either side of the debate rely upon comparisons between groups of speakers, and often fail to acknowledge the ways in which the bilingual populations sampled from around the globe differ from one another, including along the dimension of social diversity of language use (e.g., Baum & Titone, 2014; Gullifer et al., 2018; Titone et al., 2017). Thus, language entropy may provide a means to more accurately characterize bilingual populations when used in addition to other classic measures like L2 AoA and L2 exposure. A final point is that even less attention is dedicated to the ways in which monolinguals may differ from one another. Monolinguals can also exhibit diversity in language usage and exposure in the form of register switching or to the extent that they are ambiently hear other languages that they do not speak. Language entropy may be fruitful in investigating diversity within monolingual populations as well. Thus, social diversity measures, such as entropy, that characterize language experience in more nuanced and realistic ways may be crucial for clarifying inconsistent evidence pertinent to such ongoing debates.

References

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*(4), 689–698. http://doi.org/10.1017/S1366728916000225

Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods, 50*(1), 250-263.

Bahrick, H. P., Hall, L. K., Goggin, J. P., Bahrick, L. E., & Berger, S. A. (1994). Fifty years of language maintenance and language dominance in bilingual hispanic immigrants. *Journal of Experimental Psychology: General*, *123*(3), 264–283. http://doi.org/10.1037/0096-3445.123.3.264

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, *35*(5), 857–894. http://doi.org/10.1017/S0142716414000174

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173–189. http://doi.org/10.1016/j.jml.2017.04.002

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. http://doi.org/10.1037/0882-7974.19.2.290

Birdsong, D., Gertken, L.M., & Amengual, M. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin. Web. 20 Jan. 2012. <https://sites.la.utexas.edu/bilingual/>.

de Bruin, A., Treccani, B., & Sala, S. D. (2015). Cognitive advantage in bilingualism an example of publication bias? *Psychological Science*, 1–9. http://doi.org/10.1177/0956797614557866

del Prado Martín, F. M., Kostić, A., & Baayen, R. H. (2004). Putting the bits together: an information theoretical perspective on morphological processing. *Cognition*, *94*(1), 1–18. http://doi.org/10.1016/j.cognition.2003.10.015

Dong, Y., & Liu, Y. (2016). Classes in translating and interpreting produce differential gains in switching and updating. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01297

Dong, Y., & Zhong, F. (2017). Interpreting experience enhances early attentional processing, conflict monitoring and interference suppression along the time course of processing. Neuropsychologia, 95, 193–203. https://doi.org/10.1016/j.neuropsychologia.2016.12.007

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient bilingual dominance scale. *Bilingualism: Language and Cognition, 12*(3), 273-289.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97*(5), 3125–3134. http://doi.org/10.1121/1.413041

Fricke, M., Zirnstein, M., Navarro-Torres, C., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition, 22*(1), 200-207.

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science, 25*(4), 1031-1036.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594–615. http://doi.org/10.1017/S1366728911000332

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. http://doi.org/10.1080/20445911.2013.796377

Green, D. W., & Wei, L. (2014). A control process model of code-switching. Language, *Cognition and Neuroscience, 29*(4), 499-511.

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1), 163–187. http://doi.org/10.1075/intp.2.1-2.07gro

Grosjean, F. (2001). The bilingual's language modes. In J. Nichol (Ed.), *One Mind, Two Languages Bilingual Language Processing* (pp. 1–22). Malden, MA: Blackwell Publishers.

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (2018). Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia*. http://doi.org/10.1016/j.neuropsychologia.2018.04.037

Gullifer, J. W., & Titone, D. (2018). Compute language entropy with {languageEntropy}. Retrieved from https://github.com/jasongullifer/languageEntropy. doi:10.5281/zenodo.1406201

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32*(2), 101–123. http://doi.org/10.1023/A:1022492123056

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs, *150*, 10–19. http://doi.org/10.1016/j.cognition.2016.01.016

Heller, M. (1982). *Language, Ethnicity and Politics in Quebec*. Berkeley.

Higgins, R. (2004). French, English, and the idea of gay language in Montreal. In W. Leap & T. Boellstorff (Eds.), *Speaking in Queer Tongues*. University of Illinois Press.

Hofweber, J., Marinis, T., & Treffers-Daller, J. (2016). Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism*, *6*(5), 648–668. http://doi.org/10.1075/lab.15052.hof

Jylkkä, J., Soveri, A., Wahlström, J., Lehtonen, M., Rodriguez-Fornells, A., & Laine, M. (2017). Relationship between language switching experience and executive functions in bilinguals: an Internet-based study. *Journal of Cognitive Psychology*, *29*(4), 404–419. http://doi.org/10.1080/20445911.2017.1282489

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Kousaie, S., Chai, X. J., Sander, K. M., & Klein, D. (2017). Simultaneous learning of two languages from birth positively impacts intrinsic functional connectivity and cognitive control. *Brain and Cognition*, *117*, 49–56. http://doi.org/10.1016/j.bandc.2017.06.003

Kroll, J. F., Dussias, P. E., & Bajo, M. T. (2018). Language Use Across International Contexts: Shaping the Minds of L2 Speakers. *Annual Review of Applied Linguistics, 38*, 60-79.

Lamarre, P. (2002). Multilingual Montreal: Listening in on the language practices of young Montrealers. *Canadian Ethnic Studies*, *34*(3).

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. http://doi.org/10.1016/j.cognition.2007.05.006

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods*, *38*(2), 202–210. http://doi.org/10.3758/BF03192770

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. http://doi.org/10.1017/S1366728913000606

López, B. G., & Vaid, J. (2018). Fácil or A piece of cake: Does variability in bilingual language brokering experience affect idiom comprehension? *Bilingualism: Language and Cognition*, *21*(02), 340–354. https://doi.org/10.1017/S1366728917000086

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. http://doi.org/10.1037/1082-989X.7.1.19

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. http://doi.org/10.1044/1092-4388(2007/067)

McLaughlin, B. (1977). Second-language learning in children. *Psychological Bulletin*, *84*(3), 438–459. http://doi.org/10.1037/0033-2909.84.3.438

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. http://doi.org/10.1016/j.cortex.2015.04.014

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382-2393.

Piske, T., MacKay, I., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*, 191–215. http://doi.org/10.006/jpho.2001.0134

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 787–796. http://doi.org/10.1037/a0035583

Poarch, G. J., Vanhove, J., & Berthele, R. (2018). The effect of bidialectalism on executive function. *International Journal of Bilingualism*. Advance online publication. doi:10.1177/1367006918763132

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from spanish-english and mandarin-english bilinguals. *Journal of the International Neuropsychological Society : JINS*, *17*(4), 682–691. http://doi.org/10.1017/S1355617711000580

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria.

Shannon, C. E. (1948). The mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423.

Statistics Canada. (2017). Montréal, V [Census subdivision], Quebec and Canada [Country] (table). Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Retrieved from https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

Subramaniapillai, S., Rajah, M. N., Pasvanis, S., & Titone, D. (2018). Bilingual experience and executive control over the adult lifespan: The role of biological sex. *Bilingualism: Language and Cognition*, *8*, 1–19. http://doi.org/10.1017/S1366728918000317

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015*. *Bilingualism: Language and Cognition*, *11*, 1–15. http://doi.org/10.1017/S1366728917000682

Takahesu Tabori, A. A., Mech, E. N., & Atagi, N. (2018). Exploiting Language Variation to Better Understand the Cognitive Consequences of Bilingualism. *Frontiers in Psychology*, *9*. http://doi.org/10.3389/fpsyg.2018.01686

Titone, D., Gullifer, J., Subramaniapillai, S., Rajah, N., & Baum, S. (2017). History-inspired reflections on the bilingual advantages hypothesis. In E. Bialystok (Ed.), *Growing Old with Two Languages Effects of Bilingualism on Cognitive Aging*. Amsterdam, The Netherlands.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1412–1431. http://doi.org/10.1037/a0024492

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, *15*, 1–21. http://doi.org/10.1017/S1366728918000421

Wickham, H. (2017). *tidyverse: Easily install and load "tidyverse" packages*. http://tidyverse.tidyverse.org/

Yang, H., Hartanto, A., & Yang, S. (2016). The complex nature of bilinguals' language usage modulates task-switching outcomes. *Frontiers in Psychology*, *7*(APR). http://doi.org/10.3389/fpsyg.2016.00560

[1] Importantly, although the entropy value can often fall between 0 and 1, entropy should not be interpreted as a proportion.

[2] We further tested whether the addition of L2 exposure and L2 AoA improved model fit when entropy measures were used instead in the base model. These model comparisons were also significant, indicating that all of the measures accounted for unique variance.

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Table 1. Participant language history information.

	Sample (N=507)	
	M	SD
Basic demographics		
Age (years)	22.85	3.73
L2 AoA (years)	6.76	4.30
Years bilingual	16.09	5.34
L1 Exposure (percentage)	63.80	20.67
L2 Exposure (percentage)	33.82	20.10
L3 Exposure (percentage)	2.38	5.58
Language entropy		
Reading	0.60	0.41
Speaking	0.70	0.41
Home	0.61	0.46
Work	0.76	0.37
Social	0.94	0.28
Principal component: Language entropy - non-work	0.00	1.00
Principal component: Language entropy - work	0.00	1.00
Accent perception in L2		
Self (1-7)	3.39	1.75
Others (1-7)	3.85	2.05

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Mean L2 accent perception (1-7)	3.62	1.79
Language abilities		
L2 speaking (1-10)	7.43	2.08
L2 reading (1-10)	7.97	1.92
L2 writing (1-10)	7.06	2.27
L2 translating (1-10)	7.14	2.13
L2 listening (1-10)	8.25	1.90
L2 pronunciation (1-10)	7.03	2.24
L2 fluency (1-10)	7.23	2.22
L2 vocabulary (1-10)	6.84	2.08
L2 grammatical (1-10)	6.88	2.35
L2 overall competence (1-10)	7.45	1.95
L1 speaking (1-10)	9.92	0.41
L1 reading (1-10)	9.94	0.36
L1 writing (1-10)	9.76	0.83
L1 translating (1-10)	9.62	0.95
L1 listening (1-10)	9.96	0.30
L1 pronunciation (1-10)	9.89	0.50
L1 fluency (1-10)	9.89	0.57
L1 vocabulary (1-10)	9.69	0.84
L1 grammatical (1-10)	9.62	1.04
L1 overall competence (1-10)	9.86	0.52
Principal component: Abilities - L2	0.00	1.00

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Principal component: Abilities - L1-1	0.00	1.00
Principal component: Abilities - L1-2	0.00	1.00

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Table 2. Model outputs for the three nested models predicting mean L2 accentedness ratings. Model comparisons indicate that model 3 (interaction model) was the best-fitting model. We report 95% confidence intervals for each point estimate in parentheses.

	Dependent variable:		
	Mean L2 accentedness ratings		
	(1)	(2)	(3)
Intercept	3.621^{***}	3.621^{***}	3.627^{***}
	(3.487, 3.756)	(3.490, 3.753)	(3.496, 3.758)
L2 AoA (scaled)	0.729^{***}	0.681^{***}	0.695^{***}
	(0.594, 0.863)	(0.548, 0.814)	(0.562, 0.829)
L2_exposure (scaled)	-0.587^{***}	-0.452^{***}	-0.445^{***}
	(-0.722, -0.452)	(-0.595, -0.310)	(-0.587, -0.303)
Non-work entropy component		-0.341^{***}	-0.323^{***}
		(-0.492, -0.190)	(-0.473, -0.173)
Work entropy component		-0.056	-0.078
		(-0.195, 0.084)	(-0.217, 0.060)
L2 AoA * L2_exposure			0.142^{**}
			(0.023, 0.261)
L2 AoA * Non-work entropy			0.169^{**}
			(0.027, 0.310)

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

L2 AoA * Work entropy component			-0.186**
			(-0.336, -0.036)
Observations	507	507	507
R^2	0.260	0.295	0.314
Adjusted R^2	0.257	0.289	0.304
Residual Std. Error	1.542 (df = 504)	1.509 (df = 502)	1.493 (df = 499)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3. Model outputs for the three nested models predicting scores on the L2 abilities component. Model comparisons indicate that model 2 (additive model) was the best-fitting model. We report 95% confidence intervals for each point estimate in parentheses.

	Dependent variable:		
	L2 abilities component		
	(1)	(2)	(3)
Intercept	-0.000	-0.000	0.0004
	(-0.071, 0.071)	(-0.067, 0.067)	(-0.067, 0.068)
L2 AoA (scaled)	-0.373***	-0.336***	-0.340***
	(-0.444, -0.302)	(-0.404, -0.268)	(-0.409, -0.271)
L2_exposure (scaled)	0.464***	0.359***	0.357***
	(0.392, 0.535)	(0.286, 0.432)	(0.283, 0.430)
Non-work entropy component		0.257***	0.254***
		(0.180, 0.334)	(0.176, 0.332)
Work entropy component		0.075**	0.077**
		(0.004, 0.146)	(0.006, 0.149)
L2 AoA * L2_exposure			-0.011
			(-0.072, 0.051)
L2 AoA * Non-work entropy			-0.015
			(-0.088, 0.059)

L2 AoA * Work entropy component			0.046
			(-0.031, 0.124)
Observations	507	507	507
R^2	0.337	0.409	0.411
Adjusted R^2	0.334	0.404	0.402
Residual Std. Error	0.816 (df = 504)	0.772 (df = 502)	0.773 (df = 499)
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Figure 1. Plot of census data related to most common language used across available communicative contexts for the following geographic locations: Montréal (city), Québec (province), and Canada (country). We obtained count data from Statistics Canada (2016) for the following questions: *What is the language that this person first learned at home in childhood and still understands?* (i.e., mother tongue on the horizontal axis), *What language does this person speak most often at home?* (i.e., home language on the horizontal axis), *In this job, what language did this person use most often?* (i.e., work language on the horizontal axis). We transformed these data to proportions (illustrated on Panel A) and computed language entropy (illustrated on Panel B). Of note, for each question, respondents could report the common use of multiple languages. In these cases, we condensed the categories to "Other – Two languages" and "Other – Three languages."

Figure 2. Partial effects plots for the L2 accentedness model. A. Plot of the interaction between L2 AoA and L2 exposure. B. Plot of the interaction between L2 AoA and the general entropy component. C. Plot of the interaction between L2 AoA and the work entropy component. Confidence intervals illustrate 1 *SEM*.

Figure 3. Partial effects plots for the L2 abilities model. A. Plot effect of L2 AoA. B. Plot of the effect of L2 exposure. C. Plot of the effect of the general entropy component D. Plot of the effect of the work entropy component. Confidence intervals illustrate 1 *SEM*.

Figure S1. Plot of the bivariate correlations between variables related to language experience, including language entropy component scores. (For ease of interpretation related to color in this figure legend, the reader is referred to the web version of this article.)

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Figure 1

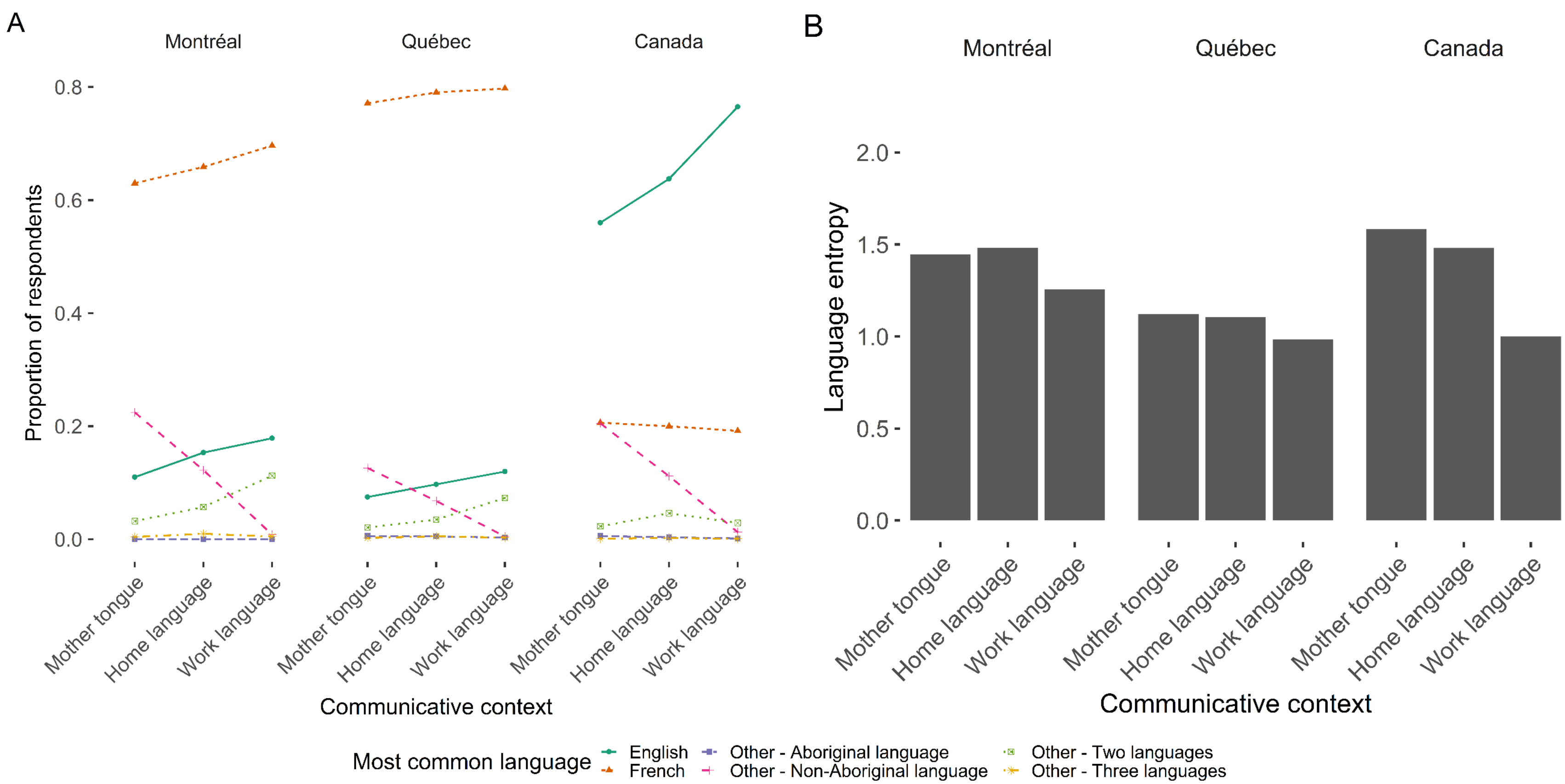
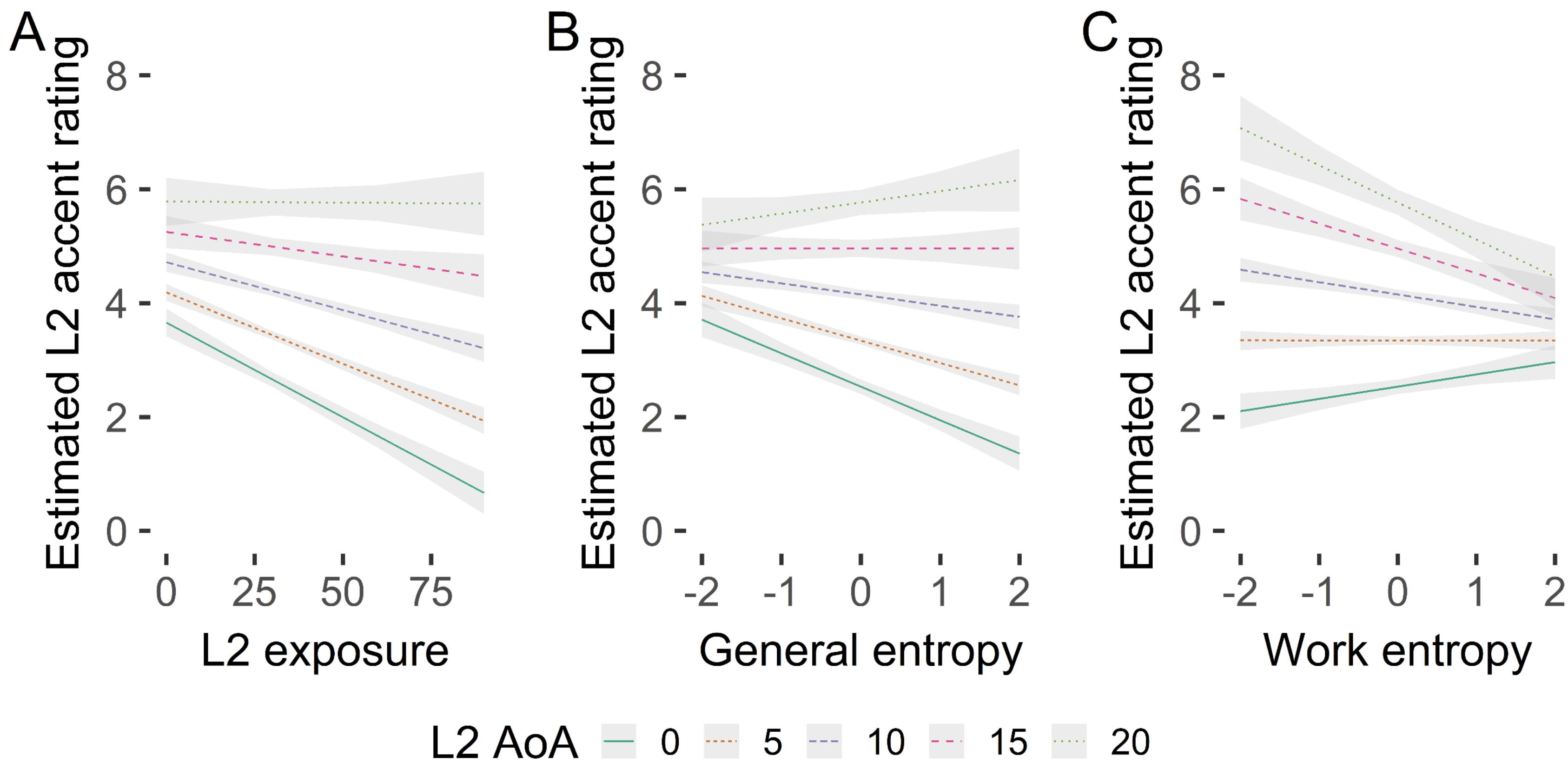
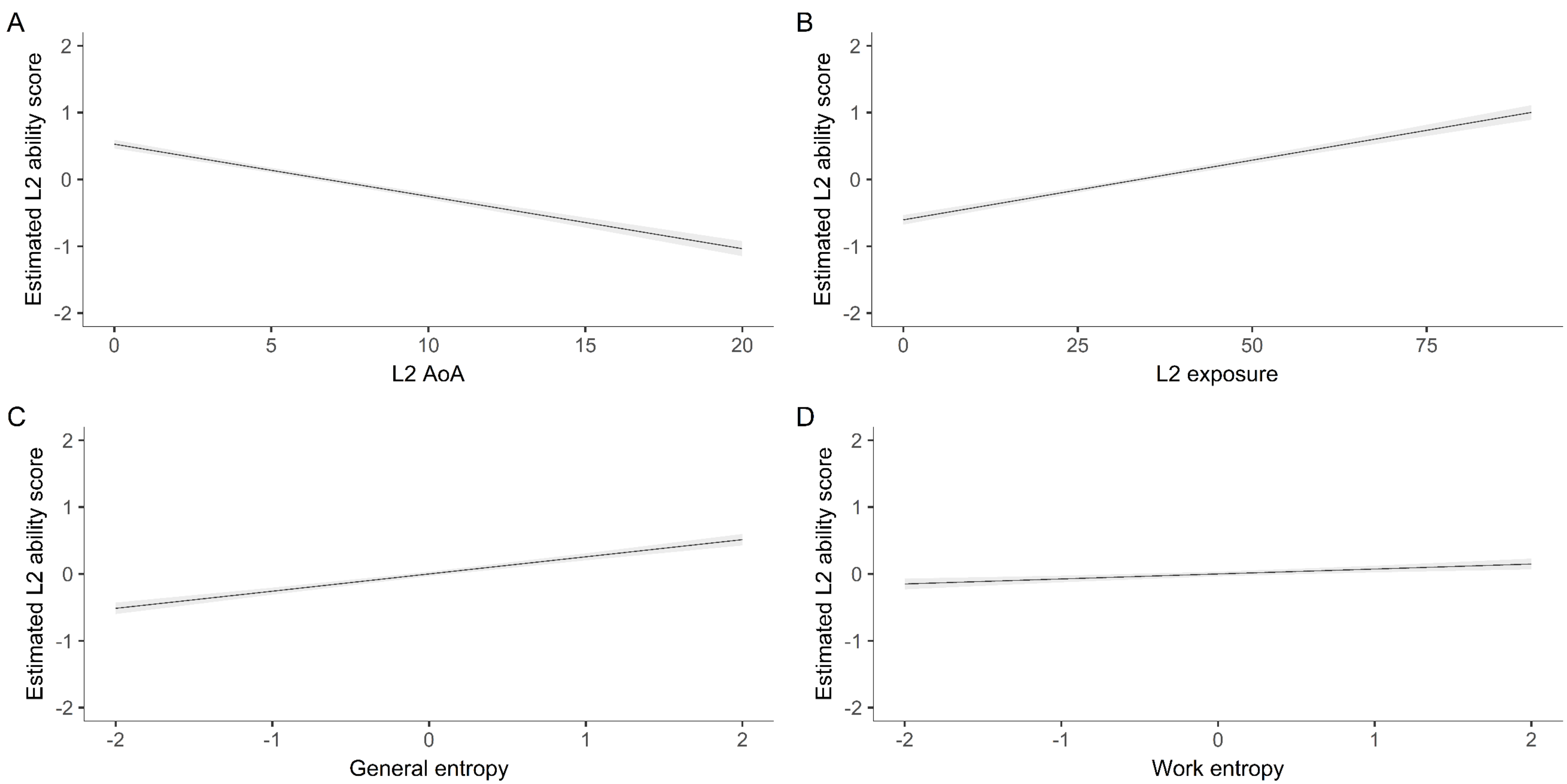

A
Montréal
Québec
Canada
0.8
0.6
0.4
0.2
0.0
Proportion of respondents
Mother tongue
Home language
Work language
Communicative context
B
Montréal
Québec
Canada
2.0
1.5
1.0
0.5
0.0
Language entropy
Mother tongue
Home language
Work language
Communicative context
Most common language
English
French
Other - Aboriginal language
Other - Non-Aboriginal language
Other - Two languages
Other - Three languages

Figure 2



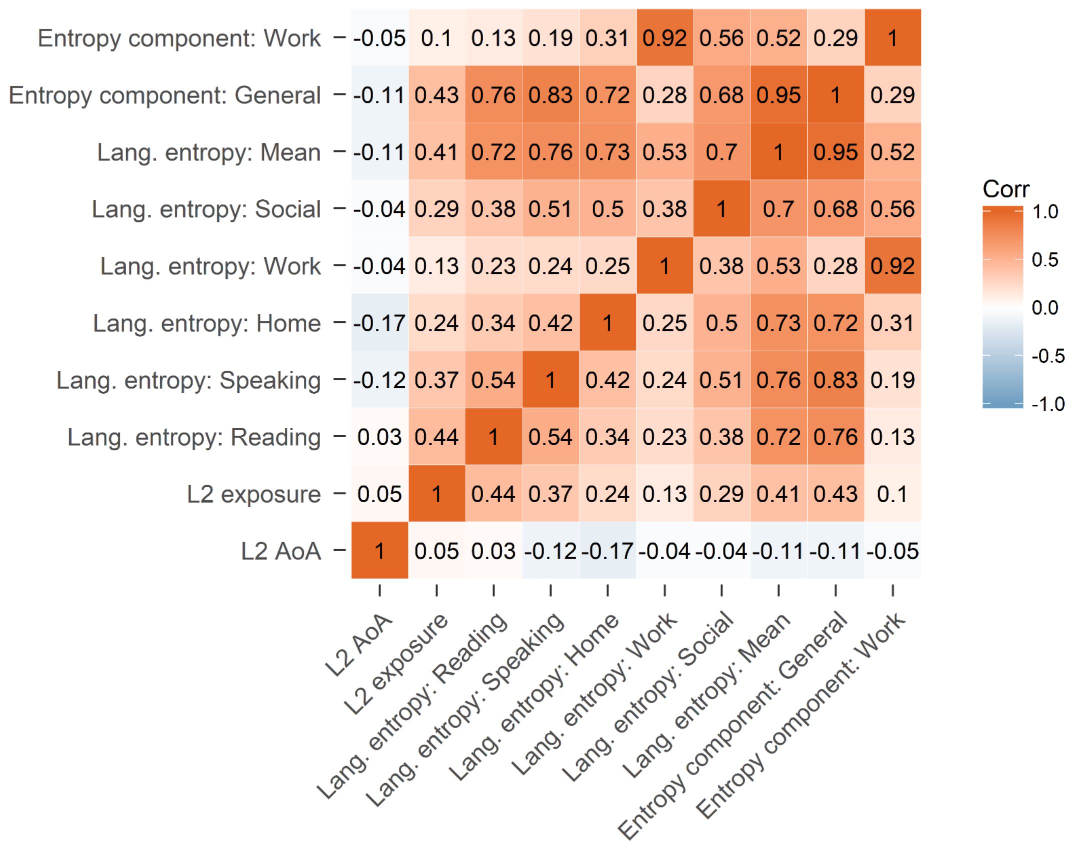
Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Figure 3



Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Figure S1



Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026

Highlights

- Bilinguals exhibit diversity in the socially distributed use of language
- We apply LANGUAGE ENTROPY to measure social diversity of language usage
- Language entropy predicts unique variance in self-report L2 accentedness
- Language entropy predicts unique variance in self-report L2 abilities

Preprint for: Gullifer, J. W., & Titone, D. (in press). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728919000026