Title
Leveraging Social Network Data to Ground Multilingual Background Measures: The Case of General and Socially Based Language Entropy

Authors
Iniesta, Antonio
Yang, Michelle
Beatty-Martínez, Anne L
et al.

Publication Date
2024

DOI
10.1037/cep0000352

Peer reviewed

Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale

Leveraging Social Network Data to Ground Multilingual Background Measures: The Case of General and Socially Based Language Entropy

Antonio Iniesta, Michelle Yang, Anne L. Beatty-Martínez, Inbal Itzhak, Jason W. Gullifer, and Debra Titone

Online First Publication, December 30, 2024. https://dx.doi.org/10.1037/cep0000352

CITATION

Iniesta, A., Yang, M., Beatty-Martínez, A. L., Itzhak, I., Gullifer, J. W., & Titone, D. (2024). Leveraging social network data to ground multilingual background measures: The case of general and socially based language entropy. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*. Advance online publication. https://dx.doi.org/10.1037/cep0000352

CANADIAN PSYCHOLOGICAL ASSOCIATION
SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE

Canadian Journal of Experimental Psychology /
Revue canadienne de psychologie expérimentale

ISSN: 1196-1961

https://doi.org/10.1037/cep0000352

Leveraging Social Network Data to Ground Multilingual Background Measures: The Case of General and Socially Based Language Entropy

Antonio Iniesta[1, 2, 3], Michelle Yang[1, 2, 3], Anne L. Beatty-Martínez[4], Inbal Itzhak[5], Jason W. Gullifer[6], and Debra Titone[1, 2, 3]

[1] Department of Psychology, McGill University
[2] Montréal Bilingualism Initiative, Montreal, Québec, Canada
[3] Centre for Research on Brain, Language, and Music, McGill University
[4] Department of Cognitive Science, University of California, San Diego
[5] Canadian Consortium on Neurodegeneration in Aging
[6] Department of Computer Science, Marianopolis College

OPEN DATA
OPEN MATERIALS

Recent research on multilingualism highlights the role of language diversity in modulating the cognitive capacities of communication and suggests a gap in available measures for quantifying socially realistic language experience. One questionnaire-based measure that potentially fills this gap is Language Entropy (e.g., Gullifer & Titone, 2018, 2020), which quantifies the balance between compartmentalised and integrated language use. However, an open question is whether questionnaire-based Language Entropy is a valid reflection of socially realistic language behaviours. To address this question, we grounded questionnaire-based Language Entropy using personal social network data for a linguistically diverse sample of speakers of French and English in the city of Montréal ($n = 95$). Specifically, we used exploratory factor analysis to characterise the factor structures resulting from questionnaire-based and social network-based Entropy. In addition, we examined the generalisability and stability of the relationship between both entropies across three bilingual groups with different social network compositions: simultaneous, English-dominant, and French-dominant. Our findings indicated that both questionnaire-based and social network-based entropies loaded onto the same factors and that the relationship between them was not affected by group differences in social network composition or by context. This suggests that questionnaire-based Language Entropy aligns well with social network-based Entropy and that this relationship is stable across different sociolinguistic realities, validating Language Entropy as a useful tool for quantifying language diversity.

Public Significance Statement
A recent development in multilingualism is the introduction of Language Entropy as a psychometric approach to understanding realistic language experiences. Our study confirmed that Language Entropy, which is a questionnaire-based measure, accurately reflects real-world language behaviour. Questionnaire-based Language Entropy aligns well with social network-based Entropy and that this relationship is stable across different sociolinguistic realities, showing that entropy is a reliable tool for understanding language diversity in various communities, helping to improve theories and methods in language research.

Keywords: multilingualism, language entropy, language diversity, personal social network, external validity

Supplemental materials: https://doi.org/10.1037/cep0000352.supp

Michael Vitevitch served as action editor.
Antonio Iniesta https://orcid.org/0000-0001-5122-5566
Debra Titone https://orcid.org/0000-0001-9060-9896
The data and scripts to reproduce the analyses presented in the article are available at https://osf.io/9arwc/. The authors disclose that Debra Titone is also the chief editor of this journal.

This research was supported by different sources: the Natural Sciences and Engineering Research Council Discovery Grant (Debra Titone), Canada Research Chairs (Debra Titone), the Social Sciences and Humanities Research Council Insight and Insight Development Grants (Debra Titone, Jason W. Gullifer), and the National Institutes of Health (Grant F32-AG064810 awarded to Anne L. Beatty-Martínez and Grant F32-HD082983 awarded to Jason W. Gullifer).

The data are available at https://osf.io/9arwc/?view_only=b5daa2f69f3143fc84f0a7d50f38839f.

The experimental materials are available at https://osf.io/9arwc/?view_only=b5daa2f69f3143fc84f0a7d50f38839f.

Correspondence concerning this article should be addressed to Antonio Iniesta or Debra Titone, Department of Psychology, McGill University, 2001 McGill College Avenue, Montréal, QC H3A 1G1, Canada. Email: antonio.martineziniesta@mcgill.ca or debra.titone@mcgill.ca

There is now a long history of assessing the impact of bilingual experience on language processing and learning (reviewed in D. A. Titone & Tiv, 2023), which is often based on information gathered through self-report evaluation instruments. This is not surprising given the multifaceted experiences that comprise multilingualism (e.g., DeLuca et al., 2019) and the high variability of linguistic experiences that people experience over the lifespan (e.g., Anderson et al., 2020; Baum & Titone, 2014). Measures such as second language (L2) age of acquisition (AoA), language proficiency, and the amount of L2 exposure have been widely employed to study individual differences in the continuum of bilingual experience (Luk & Bialystok, 2013; Surrain & Luk, 2019), which drive linguistic performance and have executive control implications (Gullifer et al., 2018; Hartanto & Yang, 2016; Kousaie et al., 2017; Pivneva et al., 2014; Subramaniapillai et al., 2019; D. Titone et al., 2017).

However, such tried and true measures, such as L2 AoA, proficiency, or L2 exposure, may be limited in capturing valuable information about the bilingual experience. To illustrate, let us consider two hypothetical people, Marco and Karla, who both learned English from birth (native language—L1) and French in school (i.e., 4 years old, second language—L2). However, as adults, they use their languages very differently in their daily lives. Marco lives in Ottawa, where the predominant language of the city is English, but teaches at a French school. Conversely, Karla lives in Montréal, where she uses both languages in school and social settings to varying degrees. Thus, while both people are highly proficient in French and English, Marco tends to compartmentalise his language usage overall. In contrast, Karla experiences a highly integrated use of English and French throughout her day across all settings. In this case, measures such as L2 AoA, proficiency, or L2 exposure would not accurately reflect the different patterns of usage for these two speakers. Thus, bilingual and multilingual individuals can vary dramatically in their exposure to and usage of each language, and this variation can be context- and location-dependent.

Moreover, the cognitive demands inherent in distinct interactional language contexts have a strong potential to influence multilingualism across individual speakers who have unique linguistic experiences (Abutalebi & Green, 2016; Beatty-Martínez et al., 2020; Green & Abutalebi, 2013; Ooi et al., 2018; Pot et al., 2018; Tiv et al., 2023). This body of work suggests that language diversity across social contexts is crucial for regulating how languages are represented, accessed, and controlled, going beyond basic measures such as L2 age of language acquisition and self-reported ability in each language (e.g., Gullifer & Titone, 2020; Tiv et al., 2021; D. A. Titone & Tiv, 2023).

To address this, Gullifer and Titone (2018, 2020) introduced a new psychometric approach to characterise individual language experiences by operationalising language diversity as a continuum (see Baum & Titone, 2014). They introduced Language Entropy, which indexes the relative balance or diversity in the use of two or more languages and has been utilised to distinguish compartmentalised from integrated language use. Language Entropy is computed from proportional information about people's language use. This information is elicited by language history questionnaires (LHQs) that are standard in the field (Birdsong et al., 2012; Dunn & Fox Tree, 2009; P. Li et al., 2014; Marian et al., 2007).

Crucially, Language Entropy can be applied flexibly, both in a general manner (i.e., using global language usage proportions) and in a socially specific manner (i.e., by averaging language use across different social spheres, such as at home, at work, and in social settings). Accordingly, when Language Entropy is low, language use is predictable because, for example, only one language is used within a given social setting. As Language Entropy increases, language use becomes less predictable, in that any language is equally possible within a social setting. Thus, lower Language Entropy indexes low language diversity and high compartmentalisation, whereas greater Language Entropy indexes high language diversity and high integration.

Relevant here, an open question is whether Language Entropy, as typically derived from basic questionnaires, patterns with more ecologically valid measures that more directly quantify socially realistic language use. Therefore, the primary objective of this study was to evaluate the external validity of Language Entropy in reflecting more specific and detailed distributions of language use in socially realistic situations. Since underlying constructs can only be measured indirectly (MacCorquodale & Meehl, 1948), it is necessary to evaluate directly whether Language Entropy is a valid estimate of the construct in question. To achieve this, we turned to personal social network data, which has relatively higher ecological validity but is more challenging to acquire in practice, to ground questionnaire-based Language Entropy.

Personal social network analysis offers granular insight into the compositional and structural characteristics of an individual's social environment and the people in it (Borgatti et al., 2009; McCarty et al., 2019; Scott, 2017; see Cuartero et al., 2023, for a review). In an egocentric social network, an individual, referred to as an ego, provides data about the network members, referred to as alters. Typically, people show a high degree of accuracy in recalling their network configuration (Parkinson et al., 2017). A wealth of mathematically complex properties or measures can be extracted from networks to provide insights into the structure or composition governing the different relationships within the network (see Vitevitch, 2019). Network structure is defined as the quantitative description of the arrangement of social ties and the extent to which a person connects with others who are otherwise unconnected (e.g., density, betweenness, and centrality; see Arnaboldi et al., 2013; Burt, 2015). Network composition is defined as the quantitative description that summarises the characteristics of network members (see Dhand et al., 2021; e.g., network size, the relational closeness between ego-alters, the similarity or homophily that exists between the ego and the alters within a particular network—such as the same hobbies, city, and gender; Fu et al., 2012; Hegde & Tumlinson, 2014; McPherson et al., 2001; Venturelli et al., 2020; see Vacca et al., 2021).

Socioecological models of language use such as the systems framework of bilingualism (D. A. Titone & Tiv, 2023; Tiv, Kutlu, Gullifer, et al., 2022; see also Atkinson et al., 2016; De Bot et al., 2007; Edwards, 2012; Steffensen & Fill, 2014) suggest that language use is intimately linked to both individual attributes and environmental factors, such as person-to-person interactions across different contexts of daily life. Importantly here, such interpersonal dynamics may be quantified using language-tagged social networks as a proxy to explore language usage and language diversity in multilinguals (Navarro & Rossi, 2023; Tiv, Kutlu, Gullifer, et al., 2022; Tiv, Kutlu, O'Regan, & Titone, 2022; D. A. Titone & Tiv, 2023). Such an approach casts bilingual language experience as a type of complex system and has been used to illuminate important behavioural phenomena such as language brokering (Kim et al., 2014), second language acquisition (Paradowski et al., 2022), global L2 usage patterns (Tiv et al., 2020), and attitudes (Feng et al., 2023). We used language-tagged social networks to calculate Language Entropy in socially realistic situations, aiming for higher ecological validity through more specific and detailed reports. This approach was intended to explore the extent to which it aligns with the Language Entropy derived from basic questionnaires in an intrasubject design. In other words, we examined the same participants to assess the congruence between questionnaire-based Language Entropy and network-based Entropy using different methods and analyses.

The Present Study

To assess the external validity of Language Entropy in capturing socially realistic language use, we examined the links between Language Entropy (both general and socially specific) and more detailed social network measures. External validation considers how the measure relates to others and fits within existing theories and frameworks (see Loevinger, 1957). The idea was to verify whether Language Entropy extracted from proportional information about self-reported people's language use in a general LHQ aligned with Language Entropy extracted from the people's language use as reported in their personal social network. For our whole sample, we tried to answer this question by performing an exploratory factor analysis (EFA) to determine whether the constructs measured by the questionnaire-based Entropy are consistent with those measured by social network-based Entropy. EFA is a statistical method used to uncover the underlying structure of the data. By identifying the common factors that explain the correlations among observed measures, EFA helps to reveal the latent constructs they represent. We expected to find that questionnaire-based Language Entropy would tap into the same underlying construct as the social network-based Entropy and that this solution would apply for both general and socially specific manners to compute language use (see Figure 1).

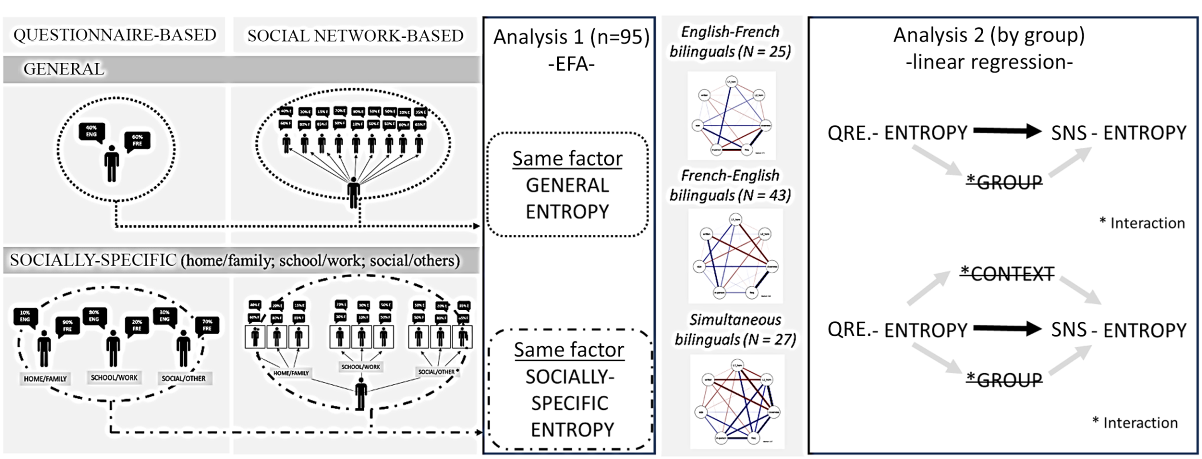
Such findings would reinforce the utility of the questionnaire-based Language Entropy as a practical and effective method for capturing the complexity of language use in socially realistic situations, thereby supporting its application in research and practical practice.

External validation also necessitates establishing construct generalisability, which involves assessing the extent to which a construct applies to different populations and various sociolinguistic realities. Given that we conducted this investigation in the diverse and multicultural city of Montréal, it was imperative to consider that the historical, political, and cultural context surrounding English and French in Québec is sociolinguistically rich (Kircher, 2014; Leimgruber & Fernández-Mallat, 2021) and can shape language use patterns. While Québec's provincial government's official language is French, at the national level, Canada has two official languages: English and French. French speakers are concentrated in the province of Québec (~ 94%; Statistics Canada, 2016), where the maintenance of French dominance is largely attributed to sustained language planning efforts from the provincial government, reinforcing French in the workplace, education, and public sphere (Kircher, 2014). However, in the city of Montréal, the English-speaking minority occupies a large proportion of high-paying and high-status positions (Hamers & Hummel, 1994; Kircher, 2014). In total, 59.9% of residents in the metropolitan area of Montréal reported French as their mother tongue, compared to 11.2% of residents who reported English as their mother tongue (Statistics Canada, 2021). However, 56.4% of residents report that they can converse in both English and French, compared to just 18.0% in the rest of Canada and 46.4% in Québec as a whole (Statistics Canada, 2021). Thus, for an individual living in Montréal, there is considerable variation in how they distribute their language across people and across various social contexts (e.g., Duval et al., 2024).

To provide a more rigorous and direct analysis to confirm and support the EFA findings, we assessed the generalisability of Language Entropy across different sociolinguistic realities. The idea is that if the relationship between questionnaire-based Language

Figure 1

Study Design and Data Analysis Plan to Explore the External Validity of Language Entropy in Capturing Socially Realistic Language Use



Note. Analysis 1 involves performing an exploratory factor analysis (EFA) on the entire sample. Analysis 2 investigates the consistency and stability of Language Entropy across different bilingual groups using linear regression models. SNS = social network entropy; QRE. = questionnaire. See the online article for the color version of this figure.

Entropy and network-based Entropy remains consistent across various bilingual groups with different network compositions, we would demonstrate that Language Entropy is a generalisable and useful measure across different sociolinguistic realities. Using linear regression models, we investigated the consistency of Language Entropy (derived from basic questionnaires and social network measures) among different sociolinguistic groups of multilingual adults living in Montréal.

To demonstrate that Language Entropy is generalisable and stable across individuals, we aimed to find no significant interactions in models that included the group factor (i.e., different bilingual groups in the analysis of general entropy measures) or in models that included both group and context factors (i.e., different bilingual groups and social contexts in the analysis of socially specific entropy measures). Such results would provide evidence of stability and consistency in the relationship between questionnaire-based and social network-based Language Entropy across various sociolinguistic realities (see Figure 1).

Method

Participants

We recruited 95 participants aged 18–45 years ($M = 21.42$, $SD = 3.57$; $n_{\text{females}} = 78$) through the McGill University Psychology participant pool in Montréal, Canada. The participants were compensated with course credit. The language of instruction and administration at McGill University is English. Most participants were born in Canada (64.21%) or France (21.05%) and grew up in a highly educated household, as 77.89% of parents/caretakers of our participants had a university degree, and 81.05% of them were native speakers of English or French, and for the rest of them, it was reported that they knew English or French. Participants reported that during their schooling, they received education in English for 41.47% of the time, in French for 41.95% of the time, and had bilingual education (both English and French) for 16.58% of the time.

Through a LHQ (see the Materials and Procedure section), participants were asked to report the languages they spoke in order of fluency and the age at which they began to learn each language. All participants reported English and French as their first or second languages. In total, 58.94% of the participants reported a third language, including Arabic (2), Armenian (1), Creole (3), German (13), Greek (1), Hebrew (1), Hindi (1), Italian (2), Québec Sign Language (1), Mandarin (2), Russian (1), Spanish (26), Tagalog (1), and Tamil (1).

We created a grouping factor for the participants based on the age of acquisition according to the following criteria: (a) If English was reported with the earliest AoA, and French AoA was reported after 3 years old, participants were grouped as English–French bilinguals ($n = 25$); (b) if French was reported with the earliest AoA, and English AoA was reported after 3 years old, participants were grouped as French–English bilinguals ($n = 43$); (c) if English and French were reported to be learned before 3 years old, or at the same age, participants were grouped as simultaneous bilinguals ($n = 27$). Participants also reported the extent to which they used the L1 ($M = 70.23$, $SD = 20.71$), L2 ($M = 27.97$, $SD = 20.35$), and L3 ($M = 1.79$, $SD = 4.85$) in general (i.e., percentage of usage of each language) and also in a variety of communicative contexts including home, family, at work, at school, in social settings, and other situations (i.e., percentage of usage of language in each setting).

Materials and Procedure

All materials and procedures were approved by the McGill University Research Ethics Board (REB No. 96-0816). After reading and signing the informed consent form, participants were asked to fill out two computer-based questionnaires in the lab. Both questionnaires were administered in person in English to ensure that all participants experienced the same treatment by researchers. The experimental session lasted approximately 50 min.

Firstly, they completed an LHQ based on the LHQ 2.0 (P. Li et al., 2014), in which participants answered general and language-related demographic questions. We extracted classic measures such as L2 AoA (based on the onset of learning) and the overall self-report competence, based on scores from 1 to 7 for speaking, reading, writing, and listening. Crucially, we extracted language usage in a general manner (i.e., using global language usage percentages) or in a socially specific manner (i.e., language use across different social spheres, i.e., at home, at work, in social settings).

Secondly, participants completed a social network survey (an earlier and simplified version of Tiv, Kutlu, O'Regan, & Titone, 2022). The survey was based on an egocentric network approach in which each respondent (also referred to as the "ego") reported on individuals they know (referred to as "alters"), thus allowing us to collect ego and ego-alter language-tagged data. Participants (i.e., egos) were asked to nominate a maximum of five people (i.e., alters) with whom they regularly interacted across various communicative contexts (i.e., family, home, work, school, social, and other) in order of importance. The difference between home and family was specified in terms of housing status; home included the top five people that participants interact with in the place where they were living (i.e., the place you currently live in), and family included the top five people that participants interact with among family members not sharing home (i.e., the family members you do not live with). The social context included the top five friends participants interact with in social settings. The category of other was included to cover other environments important to the participants and not considered in the other ones (i.e., hobby environment, religious community, neighbourhood, volunteer setting).

Finally, participants answered basic demographic questions about each person in their social network, along with information related to social network composition. This included the frequency of communication (rated from 1 to 4, where 1 means daily and 4 means yearly), the relational closeness between the participant (ego) and each alter (rated on a 1–7 Likert scale, where 1 means *extremely close* and 7 means *stranger-like*), and the mode of communication (yes or no) for in person and written communication (see Dhand et al., 2021; Vacca et al., 2021). Most relevant to our research questions, participants also indicated which language or languages they used with each alter, providing an estimate of the percentage of time each language was used with each person.

Computing Language Entropy: Questionnaire-Based and Social Network

We computed Shannon entropy (H) using the following equation and the methods available in the languageEntropy R package

(Gullifer & Titone, 2018), which considers the proportion of usage for a particular language in a set of languages, providing a continuous index of language usage. The process can be computed in a general manner (i.e., using global language usage proportions) and in a socially specific manner (i.e., by averaging language use in the different spheres, such as at home, at work, and in social settings). Thus, using the situation described earlier (see introduction), our hypothetical participants, Marco and Karla, provided a percentage of use for French (L2) and English (L1) in their daily lives (which should add up to 100; Marco used English 90% of the time and French 10%, while Karla used English 60% of the time and French 40%). This information was used to calculate general entropy. Additionally, participants provided the percentage of language use in work and in social settings (which should add up to 100 in each context). In the work setting, Marco used 80% French and 20% English, while Karla used 40% French and 60% English. In the social setting, Marco used 85% English and 15% French, while Karla used 70% English and 30% French. This allowed us to calculate an entropy index specific to each participant and social context.

We calculated entropy scores from the self-reported language use data (i.e., LHQ 2.0) for the L1, L2, and L3 in general and in each communicative context (home, family, work, school, social, and others). For speaking, we converted the percentages to proportions, which were then used to calculate the entropy for each participant (Gullifer et al., 2018, 2021; Gullifer & Titone, 2018, 2020). Importantly, we also calculated the entropy for each alter in two ways: first by averaging them all to calculate general entropy, and second by averaging the alters in each context to calculate socially specific entropy. Our procedure resulted in 14 entropy scores for each participant, seven of them based on questionnaires and the other seven on social networks (including the general entropy and the entropy for the six communicative contexts; see Table 1). Language Entropy provided a continuous index with a range from 0 to some maximum value. Language Entropy was at its minimum ($H = 0$) when one language in a set was used all the time in that context (i.e., 100% of the time), and the other languages never occurred, representing a completely compartmentalised usage. Language Entropy was at its maximum when the percentage of usage for two or more languages was equal within a communicative context (i.e., $H = 1$ for a 50%–50% for a bilingual individual; $H = 1.585$ for a 33%–33%–33% for a trilingual individual), representing a completely integrated usage.

Computing Social Network Composition Measures

In our study, we calculated various social network composition measures by averaging the data provided by participants. These measures included *network size*, which refers to the total number of individuals (alters) in each participant's social network. We also assessed the *frequency* of communication with alters, capturing how often participants interacted with each member of their network. Additionally, we measured the *closeness* in the relationship between the ego (participant) and each alter, which reflects the strength of their social ties. Furthermore, we examined the mode of communication, recording whether participants engaged in *in-person* and *written* communication with their alters. By averaging these data points, we obtained comprehensive and detailed insights into the social network compositions of our multilingual adult participants in Montréal.

In addition, and considering that *homophily* has been described as the similarity that exists between the ego and the alters within a particular network (Fu et al., 2012; Hegde & Tumlinson, 2014; Huber & Malhotra, 2017; McPherson et al., 2001, 2001; Parkinson et al., 2018; Venturelli et al., 2020), we classified each of the nominated alters in terms of the languages that the ego used to communicate with them, as the first step to compute the ego-alter language homophily. In the social network survey, each ego nominated alters from different social spheres and specified the language or languages used with each alter. Homophily was coded as a binary variable (TRUE/FALSE). Homophily was TRUE for

Table 1

Descriptive Statistics for Questionnaire-Based and Social Network-Based Entropy for General and Socially Specific Manners to Compute Language Entropy

		Socially specific entropy					
Instrument	General entropy	Home	Family	Work	School	Social	Others
$N = 95$							
LHQ	0.518 (0.217)	.442 (.388)	.437 (.373)	.595 (.380)	.472 (.364)	.683 (.386)	.459 (.439)
SNS	0.269 (0.185)	.277 (.262)	.266 (.281)	.271 (.266)	.254 (.266)	.279 (.238)	.245 (.287)
N° alters	*16.863 (5.099)*	*2.473 (1.681)*	*2.863 (1.848)*	*1.505 (2.158)*	*3.979 (1.637)*	*4.442 (1.315)*	*1.937 (2.231)*
NAs distribution							
LHQ	0	11	20	61	13	6	49
SNS	0	13	18	62	9	6	50
English–French bilinguals ($N = 25$)							
LHQ	0.421 (0.243)	.354 (.452)	.311 (.308)	.707 (.371)	.331 (.311)	.500 (.451)	.384 (.452)
SNS	0.247 (0.208)	.226 (.283)	.308 (.361)	.401 (.210)	.211 (.252)	.242 (.289)	.296 (.394)
French–English bilinguals ($N = 43$)							
LHQ	0.568 (0.206)	.433 (.342)	.422 (.406)	.496 (.450)	.562 (.42)	.805 (.353)	.591 (.443)
SNS	0.292 (0.196)	.285 (.253)	.259 (.29)	.220 (.287)	.307 (.280)	.291 (.244)	.284 (.267)
Simultaneous bilinguals ($N = 27$)							
LHQ	0.528 (0.186)	.539 (.391)	.538 (.339)	.555 (.35)	.447 (.269)	.628 (.326)	.354 (.411)
SNS	0.251 (0.140)	.31 (.262)	.25 (.215)	.211 (.274)	.203 (.248)	.288 (.184)	.157 (.197)

Note. Values in italics represent that N° alters is nested and related to SNS and not to LHQ. LHQ = language history questionnaire; SNS = social network; NA = not available.

each of the languages that ego-alter shared, regardless of the percentage of usage. For instance, if the ego's L1 was French and the ego used only French with a specific alter (e.g., a family member), this alter was considered part of the L1 homophily. If the ego communicated with another alter (e.g., a coworker) in both French and English, this alter was part of both the L1 homophily and the L2 homophily. In this scenario, the same alter contributed to both L1 and L2 homophily. We then computed proportion L1 and L2 homophily for each ego, representing the proportion of the network that uses the L1 or/and the L2 during communication (i.e., the amount of representation for each of the language in the network). This variable ranged from 0 to 1.0, with each end representing low homophily and high homophily, respectively. This is calculated for each ego by summing the number of TRUE alters for each language and dividing by the total number of alters in that category (Marsden, 1988; McPherson et al., 2006).

Results

To assess the external validity of Language Entropy in capturing socially realistic language use distributions, we took a bottom–up, data-driven approach by conducting an EFA to identify underlying latent factors for questionnaire-based Language Entropy and social network-based Entropy for general and socially specific manners to compute the measures (see also Gullifer et al., 2021; Siegelman et al., 2024).

To prepare the data for EFA, all the variables underwent a logarithmic transformation to stabilise variance and reduce skewness, followed by scaling for normalisation. We assessed the multivariate normality of the standardised variables using Mardia's test, which evaluates both skewness and kurtosis to determine if the data follow a multivariate normal distribution. To further assess the factorability of the data, we conducted Bartlett's test of sphericity and calculated the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy. To determine the appropriate number of factors for each of the EFAs, we used two methods: parallel analysis and Horn's parallel analysis. In parallel analysis, we used the *fa.parallel* function from the *psych* R package (Revelle, 2018). This method compares the eigenvalues from the actual data with those from a randomly generated data set of the same size and structure. Horn's parallel analysis, performed using the *paran* R package, provided further validation by running multiple iterations and comparing the observed eigenvalues with those from randomly generated data, focussing on the 95th percentile. This dual approach ensured a robust determination of the number of factors to retain.

Finally, we conducted the EFA to identify the underlying factor structure of our standardised variables using the maximum likelihood (ML) method with oblimin rotation. The maximum likelihood method was chosen for factor extraction because it offers several advantages in statistical estimation. Specifically, this method assumes that the data are multivariate normally distributed and seeks to find the factor solution that best fits the observed data by maximising the likelihood function. We applied oblimin rotation, a type of oblique rotation, to the factor solution because it allows the factors to be correlated. This rotation technique helps achieve a simpler and more interpretable factor structure by maximising high loadings and minimising low loadings on each factor, thus clarifying the relationships between the observed variables and the underlying factors.

Although we initially had six contexts in which participants indicated the percentage of use of each of their languages (questionnaire-based) or nominated five alters in each of these contexts (social network-based), we found many participants without some contexts, which complicated the implementation of the EFA for socially specific manners to compute the measures (see Table 1). For this reason, and considering previous studies that showed an external versus internal distribution of socially specific entropy (see Duval et al., 2024; Gullifer & Titone, 2020), we decided to group the contexts into three larger categories: (1) family–home; (2) work–school; and (3) social–others, reducing drastically the degree of missing data (family–home: zero in LHQ and one in social network entropy [SNS]; work–school: six in LHQ and four in SNS; social–others: three in LHQ and four in SNS).

To provide a more rigorous and direct analysis to confirm and support the EFA findings, we assessed the generalisability and the stability of Language Entropy across different sociolinguistic realities by comparing across three distinct groups of multilinguals living in Montréal: English–French, French–English, and simultaneous bilinguals. Firstly, we investigated the differences between groups in each of the social network composition measures (i.e., size, frequency, closeness, in-person communication, written communication, L1 homophily, and L2 homophily). Also, we explored the correlation between the measures within each group as additional evidence of the compositional network differences between them.

Two different linear regression models (i.e., general and socially specific manners to compute Language Entropy) were run in R (RStudio Team, 2020). These models predicted questionnaire-based Language Entropy using social network-based Entropy. Both variables were log-transformed and scaled. The general model included an interaction with group, while the socially specific model included interactions with both group and context. In both models, group (with three levels: English–French, French–English, and simultaneous bilingual adults) and context (with three levels: family–home, work–school, and social–other) factors were treatment-coded, with the English–French group and family–home serving as the reference levels (coded as 0). We report the output of each model using *model_parameters* function. The idea was that if Language Entropy is generalisable and stable across different sociolinguistic realities, the differences in social network composition between groups should not affect the relationship between questionnaire-based and social network-based Language Entropy. To confirm this, we should not find significant interactions with group or with group and context factors. This would indicate that SNS predicts questionnaire-based Entropy consistently across different circumstances, demonstrating that the core relationship is robust enough to be stable regardless of the social composition of the participants (see Figure 1).

The data and scripts to reproduce the analyses presented below are available at https://osf.io/9arwc/ (Iniesta et al., 2024).

Questionnaire-Based and Social Network-Based Language Entropy: EFA Results

Regarding the general entropy, the log-transformed and standardised data showed skewness and kurtosis values between −1 and 1 for both questionnaire-based and social network-based Language Entropy (see Table 2A). Mardia's kurtosis coefficient was 0.234 (p = .815), indicating no significant deviation from multivariate

Table 2

Data Normality and Exploratory Factor Analyses for Questionnaire-Based and Social Network-Based Entropy for General (A) and Socially Specific Manners (B) to Compute Language Entropy

Language entropy	Skewness	Kurtosis	ML1	ML2	Variance	Complexity	Uniqueness
(A) General							
LHQ-entropy	−0.396	−0.400	0.78		61.07%	1.00	0.39
SNS-entropy	0.534	−0.379	0.78			1.00	0.39
(B) Socially specific							
LHQ-entropy work–school	−0.117	−0.778		0.37	24.77%	1.55	0.89
SNS-entropy work–school	0.675	−0.298		0.58		1.24	0.72
LHQ-entropy social–others	−0.392	−0.845		0.64		1.06	0.53
SNS-entropy social–others	0.554	−0.549		0.80		1.01	0.32
LHQ-entropy family–home	−0.106	−0.611	0.74		24.26%	1.00	0.45
SNS-entropy family–home	0.686	−0.055	0.91			1.00	0.15

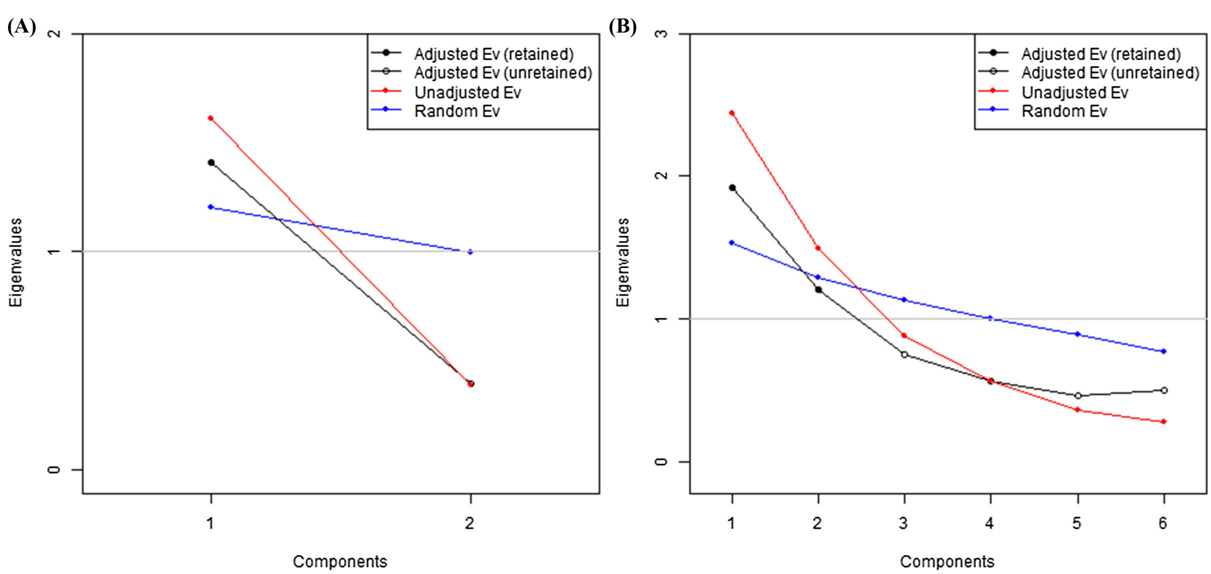
Note. LHQ = language history questionnaire; SNS = social network; ML = maximum likelihood.

normality. Bartlett's test of sphericity was highly significant, $\chi^2(1) = 43.16$, $p < .001$, indicating significant correlations among the variables, which supports their suitability for factor analysis. The overall KMO measure was 0.50, suggesting that the data are appropriate for factor analysis. The parallel analysis suggested retaining one factor. Horn's parallel analysis confirmed the retention of one component with an adjusted eigenvalue of 1.40 (i.e., above the threshold of 1). The elbow plot comparing retained versus random eigenvalues further supported retaining one factor (see Figure 2a). Based on our earlier determination, we specified a single-factor model. The results revealed that both variables had high loadings on the single factor, with loading values of 0.78 for both questionnaire-based and social network-based entropies. The complexity for each variable was 1.00, and the uniqueness values were 0.39. This single latent factor accounted for 61.07% of the total variance in the original data. These findings suggest a strong common underlying factor influencing both entropy measures, providing a robust representation of the data's structure (see Table 2A).

Regarding the context-specific entropy, the log-transformed and standardised data showed skewness and kurtosis values between −1 and 1 for questionnaire-based and social network-based entropies in family–home, work–school, and social–other contexts (see Table 2B). Mardia's kurtosis coefficient was 1.493 ($p = .135$), indicating no significant deviation from multivariate normality. Bartlett's test of sphericity was highly significant, $\chi^2(1) = 143.16$, $p < .001$, indicating significant correlations among the variables, which supports their suitability for factor analysis. The overall KMO measure was 0.60, suggesting that the data are appropriate for factor analysis. The parallel analysis suggested retaining two factors. Horn's parallel analysis confirmed the retention of two components with an adjusted eigenvalue of 1.92 and 1.20 (i.e., both above the threshold of 1). The elbow plot comparing retained

Figure 2

Elbow Plots for the Parallel Analysis Comparing Retained Versus Random Eigenvalues to Confirm the Number of Factors for the General (A) and Socially Specific (B) EFAs Including Questionnaire-Based and Social Network-Based Entropy



Note. EFA = exploratory factor analysis; Ev = Eigenvalues. See the online article for the color version of this figure.

versus random eigenvalues further supported retaining two factors (see Figure 2b).

Based on our earlier determination, we specified a two-factor model. For the first factor (ML2), questionnaire-based and social network-based Language Entropy for the social–others context loaded together, with loading values of 0.80 and 0.64, respectively, with complexity values of 1.01 and 1.06 and uniqueness values of 0.32 and 0.53. Additionally, the questionnaire-based and social network-based entropies for work–school context loaded also into the first factor, with loading values of 0.58 and 0.37, respectively, with complexity values of 1.24 and 1.55 and uniqueness values of 0.72 and 0.89. For the second factor (ML1), questionnaire-based and social network-based entropies for family–home context loaded together, with loading values of 0.91 and 0.74, respectively, with complexity values of 1.00 and uniqueness values of 0.15 and 0.45. The two latent factors accounted for a total of 49.03% of the variance in the original data, with the first factor (ML2) explaining 24.77% and the second factor (ML1) explaining 24.26% (see Table 2B).

Generalisability of Entropy Across Different Groups: Linear Regression Results

The three groups (i.e., English–French bilinguals, French–English bilinguals, simultaneous bilinguals) showed some differences in terms of social network composition (see Supplemental Figure S1a). Specifically, French–English bilinguals showed differences, or tended to differ, from English–French bilinguals in terms of L1 homophily ($\beta = -.48$; $t = -1.95$; $p = .054$), L2 homophily ($\beta = .47$; $t = 1.88$; $p = .063$), frequency of communication ($\beta = .54$; $t = 2.18$; $p = .032$), and in-person communication ($\beta = .68$; $t = 2.83$; $p = .006$). Both groups did not show differences in terms of network size ($\beta = .30$; $t = 1.21$; $p = .230$), closeness ($\beta = .05$; $t = .22$; $p = .829$), or written communication ($\beta = .11$; $t = .45$; $p = .654$). In addition, simultaneous bilinguals showed differences, or tended to differ, from English–French bilinguals in terms of L1 homophily ($\beta = -.54$; $t = -1.98$; $p = .051$) and frequency of communication ($\beta = .53$; $t = 1.94$; $p = .055$). Both groups did not show differences in terms of L2 homophily ($\beta = .44$; $t = 1.61$; $p = .111$), network size ($\beta = .36$; $t = 1.30$; $p = .196$), closeness ($\beta = .16$; $t = .59$; $p = .557$), in-person communication ($\beta = -.03$; $t = -.11$; $p = .915$), or written communication ($\beta = .26$; $t = .94$; $p = .351$).

A refit model with French–English bilinguals as the reference level showed that simultaneous bilinguals differed from French–English bilinguals in terms of in-person communication ($\beta = .65$; $t = 2.78$; $p = .007$). There were no differences between simultaneous and French–English bilinguals in any other network composition measures (all $ps > .05$; L1 homophily: $\beta = -.06$; $t = -.24$; $p = .812$; L2 homophily: $\beta = -.03$; $t = -.11$; $p = .915$; network size: $\beta = .06$; $t = .24$; $p = .814$; closeness: $\beta = .11$; $t = .44$; $p = .658$; frequency of communication: $\beta = -.01$; $t = -.03$; $p = .973$; written communication: $\beta = .15$; $t = .60$; $p = .550$).

Additionally, we conducted a network analysis using JASP to descriptively explore the differential relationships between network composition measures across the three bilingual groups. We included the COR (correlations) estimator to assess the relationships between variables within each group. For each group, we extracted the weights matrix (see Supplemental Table S1), which represents the strength of the connections between variables. We generated network plots for each group, applying a fixed ratio and arranging the layout in a circular format just to introduce a visual representation of the correlation between measures (see Supplemental Figure S1b).

Now that we have demonstrated the groups have different social configurations, we aimed to determine if the grouping factor modulates the relationship between questionnaire-based and social network-based Language Entropy. We found that social network-based Entropy predicted questionnaire-based Entropy in the model, including general entropy ($\beta = .65$; $t = 4.66$; $p < .001$). There were differences between groups in entropy, with English–French bilinguals showing the lowest entropy (see Supplemental Figure S2). Crucially, we did not find significant interactions between social network-based Entropy and group. This indicates that social network-based Entropy predicted questionnaire-based Entropy to the same extent in English–French bilinguals as in French–English bilinguals ($\beta = -.15$; $t = -.86$; $p = .393$) and simultaneous bilinguals ($\beta = .10$; $t = .44$; $p = .663$; see Supplemental Table S2a for the full model output). A refit model with French–English bilinguals as the reference level showed no differences between French–English and simultaneous bilinguals ($\beta = .26$; $t = 1.15$; $p = .253$). The relationship between social network-based and questionnaire-based Entropy was stable across groups (see Figure 3).

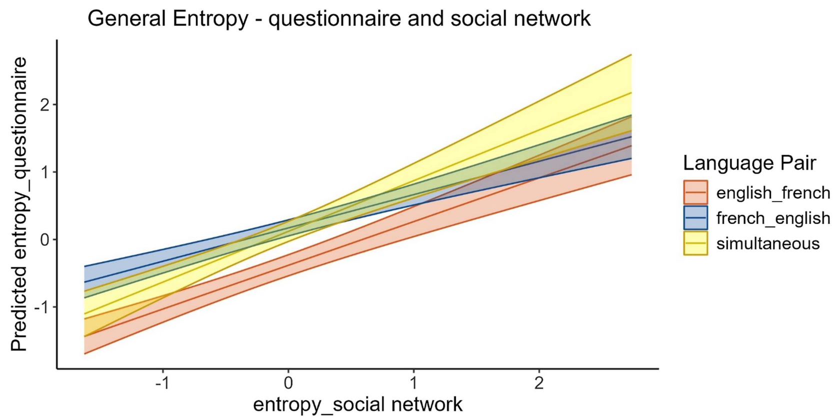
In the same direction, we found that social network-based Entropy predicted questionnaire-based Entropy in the model, including socially specific entropy ($\beta = .60$; $t = 5.04$; $p < .001$). There were differences between groups in entropy, with English–French bilinguals showing the lowest entropy. In addition, there were differences between contexts in entropy, with the family–home the context having the lowest entropy (see Supplemental Figure S3). Crucially, we did not find significant interactions between social network-based Entropy and group, indicating that social network-based Entropy predicted questionnaire-based Entropy to the same extent in English–French bilinguals as in French–English bilinguals ($\beta = -.03$; $t = -.16$; $p = .876$) and simultaneous bilinguals ($\beta = .06$; $t = .26$; $p = .795$). We did not find significant interactions between social network-based Entropy and contexts, indicating that social network-based Entropy predicted questionnaire-based Entropy to the same extent in family–home as in social–others ($\beta = .15$; $t = .82$; $p = .415$) and work–school ($\beta = -.21$; $t = -.98$; $p = .329$) contexts. We did not find any three-way interaction (see Supplemental Table S2b for the full model output). A refit model with French–English bilinguals or social–others as the reference levels showed neither differences between French–English and simultaneous bilinguals ($\beta = .09$; $t = .38$; $p = .707$) nor between social–others and work–school contexts ($\beta = -.36$; $t = -1.59$; $p = .114$). The relationship between social network-based and questionnaire-based Entropy was stable across groups and contexts (see Figure 4).

Discussion

A relatively novel advance in the field of multilingualism was the introduction of Language Entropy as a psychometric approach to characterising socially realistic language experiences (Gullifer & Titone, 2018, 2020). Considering the multifaceted experiences that comprise multilingualism (DeLuca et al., 2019) and the high variability of linguistic experiences across the lifespan (Anderson et al., 2020), entropy indexes the relative balance or diversity in the use of two or more languages and has been utilised to distinguish compartmentalised from integrated language use in distinct

Figure 3

Social Network-Based Entropy Predicted Questionnaire-Based Entropy in the Model That Included General Entropy Without Any Significant Interactions With the Group Factor



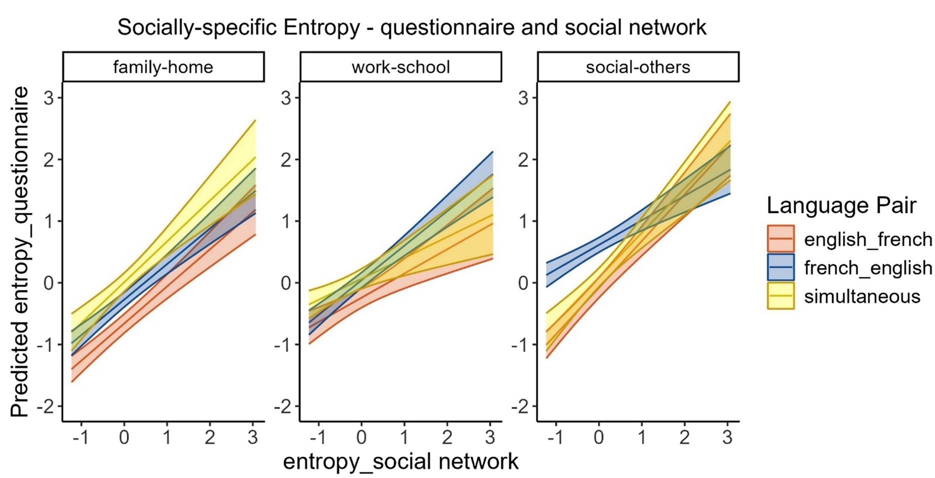
Note. See the online article for the color version of this figure.

interactional language contexts. Although individual differences in entropy have previously been linked with neurocognitive aspects of executive control and language proficiency (Gullifer et al., 2018, 2021; Gullifer & Titone, 2020; Kałamała et al., 2020; van den Berg et al., 2022), unclear was whether Language Entropy (derived from basic questionnaires) patterns with more ecologically valid measures that quantify socially realistic language use (Bruine de Bruin et al., 2020; Kałamała et al., 2020, 2022; J. Li et al., 2020; Mann & de Bruin, 2022). The aim of this study was to assess the external validity of Language Entropy in capturing socially realistic language use distributions.

Firstly, we leveraged personal social network data, which hold high ecological validity but are challenging to administer relative to the more accessible questionnaire-based Language Entropy measure. This served as a basis for validating entropy as a measure of diversity in language usage. We calculated entropy from self-reported language usage data (i.e., P. Li et al., 2014) in general and in each communicative context (home, family, work, school, social, and others). We expanded the level of analysis to individuals' language-tagged social networks (e.g., Feng et al., 2023; Tiv et al., 2020) to assess language usage for individuals (i.e., egos) between egos and their associates (i.e., ego-alter connections). We took a bottom–up, data-driven approach to assess underlying latent factors from questionnaire-based and social network-based Entropy measures (see also Gullifer et al., 2021; Siegelman et al., 2024). The results showed that questionnaire-based Language Entropy and social network-based Entropy patterned together and that this solution extended to both general and socially specific manners to compute entropy.

Figure 4

Social Network-Based Entropy Predicted Questionnaire-Based Entropy in the Model That Included Socially Specific Entropy Without Any Significant Interactions With Group or Context Factors



Note. See the online article for the color version of this figure.

The systems framework of bilingualism (Tiv, Kutlu, Gullifer, et al., 2022; D. A. Titone & Tiv, 2023) suggests that interpersonal dynamics involve person-to-person interactions in various contexts of daily life, broadening the perspective beyond the individual and emphasising the importance of their associates. Previous studies implemented language-tagged social networks as a proxy to explore language usage and language diversity in multilinguals (Navarro & Rossi, 2023; D. A. Titone & Tiv, 2023; Tiv, Kutlu, Gullifer, et al., 2022; Tiv, Kutlu, O'Regan, & Titone, 2022). This study serves as a cross-validation of questionnaire-based Language Entropy and social network-based Entropy, demonstrating that both indexes capture the same complex variability in diversity in language use. Importantly, we showed that social network analysis can be used methodologically to assess ecological external validation (i.e., extrapolate to more valid settings).

Secondly, we conducted this investigation in the diverse and multicultural city of Montréal and capitalised on this diversity to demonstrate the psychometric stability of Language Entropy. In this linguistically rich environment, each speaker's unique language background is shaped by historical, political, and sociocultural factors (Kircher, 2014; Leimgruber & Fernández-Mallat, 2021). External validation also necessitates establishing construct generalisability, which involves assessing the extent to which a construct applies to different populations and various sociolinguistic realities.

To provide a more rigorous and direct analysis to confirm and support the EFA findings, we assessed the generalisability of Language Entropy across different sociolinguistic realities. Our sample was composed of French–English, English–French, and simultaneous bilinguals from the McGill University community. Despite the three groups having different social compositions and varying historical, political, and cultural contexts surrounding English and French (Kircher, 2014; Leimgruber & Fernández-Mallat, 2021), we found that social network-based Entropy predicted questionnaire-based Entropy for both general and socially specific measures. Importantly, there were no interactions with group or context factors, indicating that social network-based Entropy predicted questionnaire-based Entropy consistently across the three bilingual groups and contexts. The absence of significant interactions suggests that differences in social network composition between groups and contexts did not affect the relationship between questionnaire-based and social network-based language Entropy. Thus, Language Entropy proved to be a generalisable and useful measure across diverse sociolinguistic realities.

Of crucial importance, these results should not be interpreted as a call to abandon the use of social network analysis in the field of bilingualism. Indeed, there is substantial richness in social network data that cannot be captured by more simple measures such as Language Entropy (Borgatti et al., 2009; McCarty et al., 2019; Scott, 2017; see Cuartero et al., 2023). Rather, the goal of this article was simply to establish simple questionnaire-based Language Entropy as a valid tool for investigating social language use and linguistic diversity when more complex social network data are neither necessary nor feasible. Thus, here we capitalised in the richness of social network data by showing that social network composition measures were capturing differences between groups. English–French bilinguals tended to show lower representation of the L2 (i.e., L2 homophily) and/or higher representation of the L1 (i.e., L1 homophily) in comparison with French–English and simultaneous bilinguals but with higher frequency of communication with the whole network in general. These differences between groups could be underlying the differences that we found between groups in Language Entropy (see Supplemental Figure S2), being English–French bilinguals the group with lower entropy, or lower linguistic diversity.

At McGill University, the participant sample for our project, the main instructional and administrative language is English. According to the 2021 student demographic survey, 62.4% of respondents have "professional/native speaker" proficiency in English, compared to 29.2% in French (McGill University, 2021), so English–French bilinguals can use more English in general, resulting in a lower linguistic diversity. Dominance of a single language, especially the first language, reduces the blending of languages in individual contexts. This observation holds true even in a highly bilingual city, where entropy has smaller or negligible effects on behaviour when the first language is extensively used across various contexts (as seen in the case of Toronto; Wagner et al., 2023).

As the sample analysed were mostly undergraduate students and some contexts did not have many observations, we grouped social contexts into three larger categories: (1) home–family, (2) work–school, and (3) social–others to reduce the number of missing data that are incompatible with the implementation of the EFA. Although other methods are available to handle missing data (see Nassiri et al., 2018; Weaver & Maxwell, 2014), deletion and imputation methods have been pointed out to be difficult to apply (Dray & Josse, 2015; Ilin & Raiko, 2010; Lorenzo-Seva & Van Ginkel, 2016), reducing the statistical power. Moreover, the identified missing data did not arise from a failure to collect responses or random nonresponses to specific questions. In this scenario, nonresponse carried a social communicative significance, reflecting the absence of a specific context and, consequently, the nonutilisation of any language within that particular context. While our solution has limitations, it demonstrates alignment with the contextual variability observed in previous studies. For example, Gullifer and colleagues (Gullifer et al., 2021) probed language usage and Language Entropy across 16 different communicative contexts or domains. Using factor analysis, they identified three latent domains of Language Entropy: entropy for internal and personal aspects of language, entropy for external or professional aspects of language, and entropy for the consumption of media and socialisation (see also Duval et al., 2024). Our context-specific EFA results further support this differentiation, demonstrating that family–home entropy grouped separately, while school–work and social–other entropies clustered together within the same factor.

Finally, even though our sample size exceeds the minimum of 50 participants established as a reasonable absolute minimum (de Winter et al., 2009), EFA is generally regarded as a technique for large sample sizes. However, collecting data from social networks can be costly and time-consuming. There are alternative solutions in the literature for exploring factor structures with small sample sizes. For example, unweighted least squares tend to recover the factor structure more accurately when few factors are retained (see Jung, 2013). Additionally, unrestricted factor analysis has been proposed to explore factor structure without constraints on the factor loadings, allowing them to vary in order to best fit the data (Steenkamp &

Maydeu-Olivares, 2023). With some adjustments to our initial R code and EFA procedure, we found that both methods resulted in the same reported factor structure with slightly different loadings (see Supplemental Tables S3 and S4).

With these caveats in mind, the results of this study reinforce a more holistic perspective on language diversity combining cutting-edge quantitative tools offered by entropy and network science. The results hold significant implications for the study of various aspects of psycholinguistic behaviour, cognition, and neuroplasticity in the context of multilingualism, as well as for the study of individual differences. They contribute to our understanding of the complex sources of sociolinguistic contexts that influence people's language use. Importantly, future studies with large sample sizes and conducted in diverse locations influenced by various historical, political, and sociocultural factors will aid in validating measures of bilingual language experience, extending entropy and social network sciences, and advancing socioecological models of language use (as reviewed by Tiv et al., 2021; Tiv, Kutlu, Gullifer, et al., 2022; see also Atkinson et al., 2016; De Bot et al., 2007; Edwards, 2012; Steffensen & Fill, 2014), both theoretically and methodologically.

Résumé

Des recherches récentes sur le multilinguisme mettent en évidence le rôle de la diversité linguistique dans la modulation des capacités cognitives de la communication et semblent indiquer la présence d'une lacune dans les mesures disponibles pour quantifier l'expérience linguistique socialement réaliste. L'entropie du langage (par exemple, Gullifer & Titone, 2018, 2020), qui quantifie l'équilibre entre l'utilisation compartimentée et intégrée des langues, est une mesure basée sur un questionnaire qui pourrait combler cette lacune. Toutefois, la question de savoir si l'entropie du langage basée sur les questionnaires est un reflet valable des comportements langagiers socialement réalistes reste ouverte. Pour répondre à cette question, nous avons mesuré l'entropie du langage basée sur les questionnaires en utilisant des données personnelles des réseaux sociaux pour un échantillon linguistiquement diversifié de locuteurs du français et de l'anglais dans la ville de Montréal (n = 95). Plus précisément, nous avons utilisé l'analyse factorielle exploratoire pour caractériser les structures factorielles résultant de l'entropie basée sur le questionnaire et de l'entropie basée sur les réseaux sociaux. En outre, nous avons examiné la généralisation et la stabilité de la relation entre les deux entropies au sein de trois groupes bilingues présentant des compositions de réseaux sociaux différentes : simultané, à dominance anglaise et à dominance française. Nos résultats indiquent que les entropies basées sur les questionnaires et les entropies basées sur les réseaux sociaux se chargent sur les mêmes facteurs et que la relation entre elles n'est pas affectée par les différences de groupe dans la composition des réseaux sociaux ou par le contexte. Cela suggère que l'entropie du langage basée sur les questionnaires s'aligne bien avec l'entropie basée sur les réseaux sociaux et que cette relation est stable dans différentes réalités sociolinguistiques, validant l'entropie du langage comme un outil utile pour quantifier la diversité linguistique.

Mots-clés : multilinguisme, entropie du langage, diversité linguistique, réseau social personnel, validité externe

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. https://doi.org/10.1017/S1366728916000225

Anderson, J. A. E., Hawrylewicz, K., & Grundy, J. G. (2020). Does bilingualism protect against dementia? A meta-analysis. *Psychonomic Bulletin & Review*, *27*(5), 952–965. https://doi.org/10.3758/s13423-020-01736-5

Arnaboldi, V., Guazzini, A., & Passarella, A. (2013). Egocentric online social networks: Analysis of key features and prediction of tie strength in Facebook. *Computer Communications*, *36*(10–11), 1130–1144. https://doi.org/10.1016/j.comcom.2013.03.003

Atkinson, D., Byrnes, H., Doran, M., Duff, P., Ellis, N. C., Hall, J. K., Lantolf, J., Larsen-Freeman, D., Negueruela, E., Norton, B., Ortega, L., Schumann, J., Swain, M., & Tarone, E. (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, *100*(Suppl. 1), 19–47. https://doi.org/10.1111/modl.12301

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

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1022–1047. https://doi.org/10.1037/xlm0000770

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual language profile: An easy-touse instrument to assess bilingualism*. Center for Open Educational Resources and Language Learning, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Bruine de Bruin, W., Parker, A. M., & Strough, J. (2020). Age differences in reported social networks and well-being. *Psychology and Aging*, *35*(2), 159–168. https://doi.org/10.1037/pag0000415

Burt, R. S. (2015). Reinforced structural holes. *Social Networks*, *43*, 149–161. https://doi.org/10.1016/j.socnet.2015.04.008

Cuartero, M., Rossi, E., Navarro, E., & Cabo, D. P. (2023). Mind the net! Unpacking the contributions of social network science for heritage and Bilingualism research. *Research Methods in Applied Linguistics*, *2*(1), Article 100041. https://doi.org/10.1016/j.rmal.2023.100041

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

de Winter, J. C., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research*, *44*(2), 147–181. https://doi.org/10.1080/00273170902794206

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

Dhand, A., McCafferty, L., Grashow, R., Corbin, I. M., Cohan, S., Whittington, A. J., Connor, A., Baggish, A., Weisskopf, M., Zafonte, R., Pascual-Leone, A., & Barabási, A. L. (2021). Social network structure and composition in former NFL football players. *Scientific Reports*, *11*, Article 1630. https://doi.org/10.1038/s41598-020-80091-w

Dray, S., & Josse, J. (2015). Principal component analysis with missing values: A comparative survey of methods. *Plant Ecology*, *216*(5), 657–667. https://doi.org/10.1007/s11258-014-0406-z

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition*, *12*(3), 273–289. https://doi.org/10.1017/S1366728909990113

Duval, A., Beatty-Martínez, A. L., Pasvanis, S., Crestol, A., Snytte, J., Rajah, M. N., Titone, D. A., Rajah, M. N., & Titone, D. A. (2024). Language diversity across home and work contexts differentially impacts age- and menopause-related declines in cognitive control in healthy females. *Journal of Experimental Psychology: General*, *153*(6), 1537–1550. https://doi.org/10.1037/xge0001564

Edwards, J. (2012). *Multilingualism: Understanding linguistic diversity*. Bloomsbury Publishing. https://doi.org/10.5040/9781472541697

Feng, R. Y., Tiv, M., Kutlu, E., Gullifer, J. W., Palma, P., O'Regan, E., Vingron, N., Doucerain, M. M., & Titone, D. (2023). A systems approach to multilingual language attitudes: A case study of Montréal, Québec, Canada. *The International Journal of Bilingualism*, *28*(3), 454–478. https://doi.org/10.1177/13670069221133305

Fu, F., Nowak, M. A., Christakis, N. A., & Fowler, J. H. (2012). The evolution of homophily. *Scientific Reports*, *2*(1), Article 845. https://doi.org/10.1038/srep00845

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

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*, *117*, 123–134. https://doi.org/10.1016/j.neuropsychologia.2018.04.037

Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., Klein, D., Baum, S., Phillips, N., & Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, *42*(2), 245–278. https://doi.org/10.1017/S0142716420000521

Gullifer, J. W., & Titone, D. (2018) *Compute language entropy with {languageEntropy}*. https://github.com/jasongullifer/languageEntropy

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294. https://doi.org/10.1017/S1366728919000026

Hamers, J. F., & Hummel, K. M. (1994). The francophones of Quebec: Language policies and language use. *International Journal of the Sociology of Language*, *1994*(105–106), 127–152. https://doi.org/10.1515/ijsl.1994.105-106.127

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. *Cognition*, *150*, 10–19. https://doi.org/10.1016/j.cognition.2016.01.016

Hegde, D., & Tumlinson, J. (2014). Does social proximity enhance business partnerships? Theory and evidence from ethnicity's role in US venture capital. *Management Science*, *60*(9), 2355–2380. https://doi.org/10.1287/mnsc.2013.1878

Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, *79*(1), 269–283. https://doi.org/10.1086/687533

Ilin, A., & Raiko, T. (2010). Practical approaches to principal component analysis in the presence of missing values. *Journal of Machine Learning Research*, *11*, 1957–2000. https://www.jmlr.org/papers/volume11/ilin10a/ilin10a.pdf

Iniesta, A., Beatty-Martínez, A., Gullifer, J. W., & Titone, D. (2024, July 31). *Leveraging social network data to ground multilingual background measures: The case of general and socially based language entropy*. https://doi.org/10.17605/OSF.IO/9ARWC

Jung, S. (2013). Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Behavioural Processes*, *97*, 90–95. https://doi.org/10.1016/j.beproc.2012.11.016

Kałamała, P., Senderecka, M., & Wodniecka, Z. (2022). On the multidimensionality of bilingualism and the unique role of language use. *Bilingualism: Language and Cognition*, *25*(3), 471–483. https://doi.org/10.1017/S1366728921001073

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition*, *204*, Article 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kim, S. Y., Wang, Y., Weaver, S. R., Shen, Y., Wu-Seibold, N., & Liu, C. H. (2014). Measurement equivalence of the language-brokering scale for Chinese American adolescents and their parents. *Journal of Family Psychology*, *28*(2), 180–192. https://doi.org/10.1037/a0036030

Kircher, R. (2014). Thirty years after Bill 101: A contemporary perspective on attitudes towards English and French in Montreal. *Canadian Journal of Applied Linguistics*, *17*(1), 20–50. https://files.eric.ed.gov/fulltext/EJ1029181.pdf

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. https://doi.org/10.1016/j.bandc.2017.06.003

Leimgruber, J. R., & Fernández-Mallat, V. (2021). Language attitudes and identity building in the linguistic landscape of Montreal. *Open Linguistics*, *7*(1), 406–422. https://doi.org/10.1515/opli-2021-0021

Li, J., Steffensen, S. V., & Huang, G. (2020). Rethinking ecolinguistics from a distributed language perspective. *Language Sciences*, *80*, Article 101277. https://doi.org/10.1016/j.langsci.2020.101277

Li, P., Zhang, F. A. N., 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. https://doi.org/10.1017/S1366728913000606

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, *3*(3), 635–694. https://doi.org/10.2466/pr0.1957.3.3.635

Lorenzo-Seva, U., & Van Ginkel, J. R. (2016). Multiple imputation of missing values in exploratory factor analysis of multidimensional scales: Estimating latent trait scores. *Anales de Psicología*, *32*(2), 596–608. https://doi.org/10.6018/analesps.32.2.215161

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*(2), 95–107. https://doi.org/10.1037/h0056029

Mann, A., & de Bruin, A. (2022). Bilingual language use is context dependent: Using the Language and Social Background Questionnaire to assess language experiences and test–rest reliability. *International Journal of Bilingual Education and Bilingualism*, *25*(8), 2886–2901. https://doi.org/10.1080/13670050.2021.1988049

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. https://doi.org/10.1044/1092-4388(2007/067)

Marsden, P. V. (1988). Homogeneity in confiding relations. *Social Networks*, *10*(1), 57–76. https://doi.org/10.1016/0378-8733(88)90010-X

McCarty, C., Lubbers, M. J., Molina, J. L., & Vacca, R. (2019). *Conducting personal network research: A practical guide*. Guilford Press.

McGill University. (2021). *Student demographic survey*. https://www.mcgill.ca/senate/files/senate/09_d20-59_student_demographic_survey.pdf

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in America: Changes in core discussion networks over two decades. *American Sociological Review*, *71*(3), 353–375. https://doi.org/10.1177/000312240607100301

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Nassiri, V., Lovik, A., Molenberghs, G., & Verbeke, G. (2018). On using multiple imputation for exploratory factor analysis of incomplete data.

Behavior Research Methods, *50*(2), 501–517. https://doi.org/10.3758/s13428-017-1013-4

Navarro, E., & Rossi, E. (2023). Using latent variable analysis to capture individual differences in bilingual language experience. *Bilingualism: Language and Cognition*. Advance online publication. https://doi.org/10.1017/S1366728923000846

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From bilingualism to bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition*, *21*(4), 867–879. https://doi.org/10.1017/S1366728918000020

Paradowski, M. B., Cierpich–Kozieł, A., Chen, C., & Ochab, J. K. (2022). How output outweighs input and interlocutors matter for study-abroad SLA: Computational social network analysis of learner interactions. *Modern Language Journal*, *106*(4), 694–725. https://doi.org/10.1111/modl.12811

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2017). Spontaneous neural encoding of social network position. *Nature Human Behaviour*, *1*, Article 0072. https://doi.org/10.1038/s41562-017-0072

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2018). Similar neural responses predict friendship. *Nature Communications*, *9*(1), Article 332. https://doi.org/10.1038/s41467-017-02722-7

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. https://doi.org/10.1037/a0035583

Pot, A., Keijzer, M., & de Bot, K. (2018). Intensity of multilingual language use predicts cognitive performance in some multilingual older adults. *Brain Sciences*, *8*(5), Article 92. https://doi.org/10.3390/brainsci8050092

Revelle, W. (2018). *psych: Procedures for personality and psychological research* (R package Version 1.8.6). Northwestern University.

RStudio Team. (2020). *RStudio: Integrated development for R*. https://www.rstudio.com/

Scott, J. (2017). *Social network analysis*. SAGE Publications. https://doi.org/10.4135/9781529716597

Siegelman, N., Elgort, I., Brysbaert, M., Agrawal, N., Amenta, S., Arsenijević Mijalković, J., Chang, C. S., Chernova, D., Chetail, F., Clarke, A. J. B., Content, A., Crepaldi, D., Davaabold, N., Delgersuren, S., Deutsch, A., Dibrova, V., Drieghe, D., Filipović Đurđević, D., Finch, B., … Kuperman, V. (2024). Rethinking first language–second language similarities and differences in English proficiency: Insights from the English reading online (ENRO) project. *Language Learning*, *74*(1), 249–294. https://doi.org/10.1111/lang.12586

Statistics Canada. (2016) *Census of population*. https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3901

Statistics Canada. (2021) *Census of population*. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E

Steenkamp, J.-B. E. M., & Maydeu-Olivares, A. (2023). Unrestricted factor analysis: A powerful alternative to confirmatory factor analysis. *Journal of the Academy of Marketing Science*, *51*(1), 86–113. https://doi.org/10.1007/s11747-022-00888-1

Steffensen, S. V., & Fill, A. (2014). Ecolinguistics: The state of the art and future horizons. *Language Sciences*, *41*, 6–25. https://doi.org/10.1016/j.langsci.2013.08.003

Subramaniapillai, S., Rajah, M. N., Pasvanis, S., & Titone, D. (2019). Bilingual experience and executive control over the adult lifespan: The role of biological sex. *Bilingualism: Language and Cognition*, *22*(4), 733–751. https://doi.org/10.1017/S1366728918000317

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*(2), 401–415. https://doi.org/10.1017/S1366728917000682

Titone, D., Gullifer, J., Subramaniapillai, S., Rajah, N., Baum, S., Sullivan, M. D., & Bialystok, E. (2017). History-inspired reflections on the bilingual advantages hypothesis. In E. Bialystok & M. D. Sullivan (Eds.), *Growing old with two languages effects of bilingualism on cognitive aging* (Vol. *53*, 265–295). John Benjamins Publishing. https://doi.org/10.1075/sibil.53.13tit

Titone, D. A., & Tiv, M. (2023). Rethinking multilingual experience through a systems framework of Bilingualism. *Bilingualism: Language and Cognition*, *26*(1), 1–16. https://doi.org/10.1017/S1366728921001127

Tiv, M., Gullifer, J., Feng, R., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics*, *56*, Article 100913. https://doi.org/10.1016/j.jneuroling.2020.100913

Tiv, M., Kutlu, E., Gullifer, J. W., Feng, R. Y., Doucerain, M. M., & Titone, D. A. (2022). Bridging interpersonal and ecological dynamics of cognition through a systems framework of bilingualism. *Journal of Experimental Psychology: General*, *151*(9), 2128–2143. https://doi.org/10.1037/xge0001174

Tiv, M., Kutlu, E., O'Regan, E., & Titone, D. (2022). Bridging people and perspectives: General and language-specific social network structure predict mentalizing across diverse sociolinguistic contexts. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, *76*(4), 235–250. https://doi.org/10.1037/cep0000273

Tiv, M., O'Regan, E., & Titone, D. (2021). In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing. *Bilingualism: Language and Cognition*, *24*(5), 918–931. https://doi.org/10.1017/S1366728921000225

Tiv, M., O'Regan, E., & Titone, D. (2023). The role of mentalizing capacity and ecological language diversity on irony comprehension in bilingual adults. *Memory & Cognition*, *51*(2), 253–272. https://doi.org/10.3758/s13421-022-01349-4

Vacca, R., Cañarte, D., & Vitale, T. (2021). Beyond ethnic solidarity: The diversity and specialisation of social ties in a stigmatised migrant minority. *Journal of Ethnic and Migration Studies*, *48*(13), 3113–3141. https://doi.org/10.1080/1369183X.2021.1903305

van den Berg, F., Brouwer, J., Tienkamp, T. B., Verhagen, J., & Keijzer, M. (2022). Language entropy relates to behavioral and pupil indices of executive control in young adult bilinguals. *Frontiers in Psychology*, *13*, Article 864763. https://doi.org/10.3389/fpsyg.2022.864763

Venturelli, V., Pedrazzoli, A., & Gallo, G. (2020). Birds of a feather flock together: The inclusive effect of similarity patterns in equity crowdfunding. *Sustainability*, *12*(9), Article 3539. https://doi.org/10.3390/su12093539

Vitevitch, M. S. (Ed.). (2019). *Network science in cognitive psychology*. Taylor and Francis Group. https://doi.org/10.4324/9780367853259

Wagner, D., Bekas, K., & Bialystok, E. (2023). Does language entropy shape cognitive performance? A tale of two cities. *Bilingualism: Language and Cognition*, *26*(5), 998–1008. https://doi.org/10.1017/S1366728923000202

Weaver, B., & Maxwell, H. (2014). Exploratory factor analysis and reliability analysis with missing data: A simple method for SPSS users. *The Quantitative Methods for Psychology*, *10*(2), 143–152. https://doi.org/10.20982/tqmp.10.2.p143

Received February 15, 2024
Revision received June 18, 2024
Accepted July 17, 2024 ■