Cognition 156 (2016) 60–70

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

When pumpkin is closer to onion than to squash: The structure of the second language lexicon

Katy Borodkin [a,*], Yoed N. Kenett [b,1], Miriam Faust [a,b], Nira Mashal [b,c]

[a] Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Israel
[b] The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Gonda Brain Research Center, Bar-Ilan University, Ramat-Gan 52900, Israel
[c] School of Education, Bar-Ilan University, Ramat-Gan 52900, Israel

ARTICLE INFO

Article history:
Received 15 February 2015
Revised 25 July 2016
Accepted 29 July 2016

Keywords:
Bilingualism
Mental lexicon
Semantic fluency
Small-world network
Modularity

ABSTRACT

The current research investigated the organization of the second language mental lexicon. Twenty-seven English-Hebrew bilingual speakers (who spoke Hebrew as their second language) completed a semantic fluency task in each of their languages, and 24 native Hebrew speakers completed the task in Hebrew. Responses were compared within and across groups, using computational tools. The analyses indicated that the lexical network of the second language displayed greater local connectivity and less modular community structure than the network in the native language, both in the entire sample and in a subsample of bilinguals whose Hebrew vocabulary was matched to that of the native Hebrew speakers. These findings suggest that the lexical network of the second language is not as well-organized as is the network of the first language, even in highly proficient bilinguals. The structural characteristics of the second language lexicon might be affected by factors related to language learning history, including age of acquisition and language use.

1. Introduction

To know a word is to know its spelling and pronunciation, its grammatical class and syntactic constraints, as well as its meaning (Nation, 2001). Word knowledge also refers to usage and associations with other words. A word may have semantic links (e.g., *pumpkin* is likely linked with *zucchini* or with *squash*) as well as associative links with other words with which it tends to co-occur (e.g., *pumpkin* may be linked with *pie*, *orange*, or *Halloween*). Although the structural characteristics of this lexical network have been extensively investigated in monolinguals (e.g., McRae, de Sa, & Seidenberg, 1997; Plaut, 1997), far less is known about the organization of the mental lexicon of bilingual speakers who acquired their languages consecutively. In such individuals, connections are often established first between words in their native language (L1). Second language (L2) words are initially connected only to their L1 translation equivalents; however, they become associated with other L2 words later on, as the L2 vocabulary is acquired, thus giving rise to an autonomous L2 lexical network (Frenck-Mestre & Prince, 1997; Kroll & Stewart, 1994). The current study focuses on the organization of this L2 network in relation to L1 network and explores it with advanced network tools.

Existing research on meaning representation in bilinguals has been largely dedicated to studying cross-linguistic connections through diverse experimental methods, such as cross-language semantic and translation priming, picture naming or Stroop (reviewed in Altarriba & Basnight-Brown, 2009; de Groot, 2011). This line of research is less informative with regard to the connections within the L2 lexical network and the principles governing the organization of this network. Studies addressing the topic more directly have often applied the word association task, in which participants are asked to generate one (or more) associative responses that come to their mind upon presentation of a target word (Kruse, Pankhurst, & Smith, 1987; Söderman, 1993). Responses in this task can be categorized as syntagmatic (words that belong to different lexical classes, such as *pumpkin-orange*), phonological (words that resemble the target word in form but not in meaning, such as *pumpkin-napkin*), and paradigmatic (words that belong to the same lexical class as the target word, such as *pumpkin-squash*). Although there is evidence suggesting that adult L2 speakers, like children in L1 (Ervin, 1961), produce more syntagmatic and phonological responses relative to adult L1 speakers (Meara, 1978; Namei,

* Corresponding author currently at: Department of Communication Disorders, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel.

E-mail addresses: katyborodkin@tauex.tau.ac.il (K. Borodkin), yoedk@sas.upenn.edu (Y.N. Kenett), Miriam.Faust@biu.ac.il (M. Faust), Nira.Mashal@biu.ac.il (N. Mashal).

[1] Present address: Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA.

http://dx.doi.org/10.1016/j.cognition.2016.07.014
0010-0277/© 2016 Elsevier B.V. All rights reserved.

2004; Söderman, 1993), other studies failed to observe differences between L1 and L2 speakers (Kruse et al., 1987; Nissen & Henriksen, 2006). Thus, it is still unclear if and how the lexical structure of L1 and L2 differ. Furthermore, while the word association task is often applied to assess lexical-semantic organization (Kolers, 1963; Van Hell & De Groot, 1998), some researchers have argued that in bilinguals it is also affected by other factors, such as word retrieval difficulties (Antón-Méndez & Gollan, 2010).

Another task that yields inconsistent findings among L1 and L2 speakers is the semantic fluency task. This task is often used in neuropsychological settings and in research to assess language functioning (Ardila, Ostrosky-Solís, & Bernal, 2006). Participants are asked to generate as many different words as possible that belong to a certain category (such as animals and vegetables) in a limited time. Bilinguals performing a semantic fluency task often produce fewer items than monolinguals, both when they are limited to only one of their languages or allowed to use both (Gollan, Montoya, & Werner, 2002; Portocarrero, Burright, & Donovick, 2007; Rosselli & Ardila, 2002; Rosselli et al., 2000; Sandoval, Gollan, Ferreira, & Salmon, 2010). They have been also reported to provide more correct responses in the dominant or more proficient language compared to the nondominant or less proficient language (Sandoval et al., 2010; Taler, Johns, Young, Sheppard, & Jones, 2013). This pattern of results may be explained by between-language interference, which results from the competition between words from both languages that is characteristic of bilingual language production (Gollan et al., 2002; Rosselli et al., 2000; Sandoval et al., 2010). Other researchers, however, did not find significant differences between monolinguals and bilinguals (Bialystok, Craik, & Luk, 2008; Friesen, Luo, Luk, & Bialystok, 2014; Luo, Luk, & Bialystok, 2010) or between the bilinguals' languages (Roberts & Dorze, 1997; Rosselli et al., 2000). These inconsistencies may be accounted in part by differences in vocabulary size (Bialystok et al., 2008). In that study, bilinguals with matched vocabulary scores performed at the same level as monolinguals, and both outperformed bilinguals with lower scores.

Responses on a semantic fluency task can be further analyzed in terms of clustering and switching, and this qualitative analysis has been used as a window into the structure of the bilingual lexicon (Roberts & Dorze, 1997; Rosselli & Ardila, 2002). Clustering refers to the production of sequences of words belonging to the same semantic subcategory, and switching – to the ability to shift to another subcategory (Troyer, 2000; Troyer, Moscovitch, & Winocur, 1997). For example, in the vegetables category, a participant may begin with the squashes family[2] (e.g., *pumpkin*, *squash*, *zucchini*), then switch to the flavorings (e.g., *onion*, *garlic*, *chives*), and so forth until the trial ends. Roberts and Dorze (1997) have demonstrated that clustering measures (i.e., length of clusters and percentage of words in clusters) were greater in French compared to English in bilinguals, indicating richer lexical network in French. The differences, however, were observed only for animal but not food names. The authors speculated that the differences between categories may be related to childhood acquisition patterns of the participants in their study (French-English speakers living in Ottawa, Canada). Similarly, the number of clusters was greater in Spanish than English in older bilinguals living in the US (Rosselli & Ardila, 2002; Salvatierra, Rosselli, Acevedo, & Duara, 2007), suggesting richer network for animal names in L1 (Spanish).

The clustering and switching scoring method (Troyer, 2000; Troyer et al., 1997), however, has faced some criticism. The categorization of responses relies on subjective judgment, which raises potential issues with reliability and validity (Taler et al., 2013). Others question the fundamental assumption of this method that the sequences in semantic fluency responses are indicative of internal lexical-semantic organization (Body & Muskett, 2012). Specifically, Body and Muskett point out the arbitrary nature of classification rules in the existing clustering systems (e.g., Troyer et al. (1997) categorize *camel* as belonging to either *beasts of burden* or *African animals*, but not to *Australian animals*, although camels are native to Australia). Using self-reports, they also show that random factors, such as perceptually salient shared characteristics, rather than semantic organization, determine many of the links between the words in a sequence (e.g., *panda* is followed by *penguin* because they are both black and white).

In the present study, a different approach was taken to explore the structural characteristics of L1 and L2 lexicon, through the use of network science tools. These tools allow for the examination of complex systems (such as the mental lexicon) as web-like structures, or networks, in which *nodes* represent individual entities and *edges* represent links between the entities. The approach has been applied in a variety of domains, including biology, social sciences, and technology (reviewed in Barabási, 2009; Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013). In bilingualism research, work within this framework has indicated that the organization of the L2 lexicon is less complex (less dense) than the organization of the L1 lexicon (Wilks & Meara, 2002). However, the authors later admitted that their assumption of many direct connections between words (several dozens) might have been over-simplistic, rendering the conclusions somewhat tentative (Meara, 2009; Wilks, Meara, & Wolter, 2005). The current study further advances this line of research by applying different computational network tools. More specifically, the small-world property and the community structure of L1 and L2 lexical networks are explored.

Networks may be defined in terms of local and global connectivity patterns. In random networks, for instance, local clustering is low (neighboring nodes are sparsely connected to each other), while global distance is short (it takes only a few steps to transverse between distant nodes). Small-world networks, on the other hand, have both high local clustering and short global distance. A network with these characteristics is called a "small-world", because every node in such network is relatively close to almost every other node. Communication transfer in this kind of networks is easy both locally and globally, and thus they are considered optimal (Watts & Strogatz, 1998). More formally, small-world networks are often defined in relation to a random network with the same number of nodes and edges. The comparison is traditionally made based on two parameters, the clustering coefficient (CC) and the average shortest path length (ASPL). The CC refers to the probability that two neighbors of a randomly chosen node will themselves be neighbors, and the ASPL represents the average shortest amount of steps that separate any two pair of random nodes. A small-world network is characterized by having a large CC despite the fact that its ASPL is relatively short and not dramatically different from a random network of comparable size.

At an intermediate level, network organization can be described in terms of community structure (in other words, modularity). A network is considered modular if it has clusters of nodes (communities) that are more densely linked to other nodes within the same community than to nodes outside the community (Newman, 2006). It has been further noted that modular systems tend to be small-world networks, whereas some small-world networks are not necessarily modular (for illustration, see Meunier, Lambiotte, & Bullmore, 2010).

Both the small-world property and modular community structure have been observed in lexical-semantic networks of monolingual adults (Borge-Holthoefer & Arenas, 2010; De Deyne & Storms, 2008; Kenett, Kenett, Ben-Jacob, & Faust, 2011;

[2] Botanically, the squashes family members are fruits, but in common language they are often treated as vegetables. The latter view was adopted for the purposes of this research.

Steyvers & Tenenbaum, 2005). In other words, these networks in native speakers may be characterized as having clusters of (mainly) semantically related words and words that link remote clusters to one another, making it possible to connect any pair of words by traversing only a few links. Interestingly, the structural properties of lexical-semantic networks are affected by the circumstances in which words are acquired. For example, children with language delay showed less local and global connectivity compared to typically developing children, perhaps reflecting the adjusted language their parents tend to adopt to fit the developmental level of their child (Beckage, Smith, & Hills, 2011). In another study, the index of global distance (ASPL) was different in children with cochlear implants compared to normal hearing children, suggesting that the initial auditory deprivation may have altered their lexical development (Kenett et al., 2013). Similarly, a recent study of bilingual first language acquisition has demonstrated that bilingual children have larger ASPL than monolinguals, presumably a consequence of a sparser lexicon (Bilson, Yoshida, Tran, Woods, & Hills, 2015).

The conditions under which L2 acquisition in sequential bilinguals takes place are often less than ideal – for instance, L2 is often acquired at an older age and used less frequently than L1 – which may lead to a suboptimal organization of the L2 lexical network. This hypothesis is consistent with previous research showing the effects of age of acquisition and language use on other aspects of lexical-semantic knowledge in L2 (i.e., lexical categorization; Malt & Sloman, 2003; Zinszer, Malt, Ameel, & Li, 2014). We thus expected the L2 network to show different patterns of connectivity compared to those of the L1 network, at the local, the intermediate, or the global level (as indexed by the CC, the Q index, and the ASPL, respectively). This hypothesis was tested in English-Hebrew bilinguals (who spoke Hebrew as their L2) and in native Hebrew speakers, who were asked to retrieve as many different exemplars as they can in the categories of fruits and vegetables and animals. The animal category was used to test the replicability of the findings observed for the fruits and vegetables category. To study the effects of L2 vocabulary knowledge, we also analyzed a subsample of L2 speakers with the highest vocabulary scores.

2. Method

2.1. Participants

The sample included 51 undergraduate students at Bar-Ilan University between the ages of 20 and 29, who received academic credit or were paid for participation. Twenty-four of the participants (13 women, mean age = 22.63, *SD* = 1.53) were English-Hebrew bilinguals, who were born in the US and lived in Israel at the time of testing. They learned English as L1 and Hebrew as L2. Bilingual language profiles were assessed using the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), which indicated that they were exposed to spoken and written Hebrew at early age (*M* = 7.32, *SD* = 6.15 and *M* = 7.38, *SD* = 3.32, respectively), before coming to Israel at the average age of 13.94 (*SD* = 6.59). They estimated that they used English somewhat more often than Hebrew (59% and 41% of time, respectively). Bilinguals rated themselves as highly proficient (on a scale ranging from 0 = *none* to 10 = *perfect*) in speaking (*M* = 7.81, *SD* = 1.74), reading (*M* = 7.55, *SD* = 1.79), and listening (*M* = 8.59, *SD* = 1.30) in Hebrew.

To control for vocabulary knowledge in Hebrew, we administered a Hebrew version of the Peabody Picture Vocabulary Test-IV (PPVT; Patael, Segal, Kaplan, & Kishon-Rabin, 2010; after Dunn & Dunn, 2007). Eighty-four words from sets 13–19 were used. Bilingual speakers provided significantly fewer correct responses (*M* = 54.66,

Table 1
Individual scores on PPVT in the PPVT-matched sub-samples.

Bilinguals		Native Hebrew speakers	
Participant #	PPVT	Participant #	PPVT
20	55	4	66
21	58	5	63
22	55	11	68
27	63	12	56
29	59	43	68
33	56	44	66
34	58	49	67
35	64	51	68
37	68	52	67
38	62	58	63
56	55	59	66
57	82	60	64

Note. PPVT = Peabody Picture Vocabulary Test.

SD = 10.66) than did native Hebrew speakers (*M* = 68.96, *SD* = 4.70), $t(49) = 6.25$, $p < 0.001$. We then selected 12 bilinguals who scored above the median PPVT score of their group (58.5) and matched their mean performance (*M* = 61.25, *SD* = 7.74) to that of 12 native Hebrew speakers (*M* = 65.17, *SD* = 3.41), $t(22) = 1.61$, $p = 0.12$. Given the small sample size, the individual scores in each of the matched groups are listed in Table 1.

The remaining 27 participants were native Hebrew speakers (18 women, mean age = 24.40, *SD* = 2.10). They were born in Israel and had not resided in an English-speaking country for more than 6 months.

2.2. Materials and procedure

2.2.1. Semantic fluency tasks

Participants were asked to generate as many different words as possible that were either fruits and vegetables or animals. After Kavé (2005), fruits and vegetables were treated as a single category to avoid the confusion between botanical definition and common use (as in *pumpkin*). Production on each category was limited to 1 min. Native Hebrew speakers performed the semantic fluency tasks only in Hebrew. Bilinguals were asked to generate words in both Hebrew and English with an hour in between the two administrations. The order of the languages was counterbalanced. Responses were recorded both manually and by audiotape, for later verification during scoring, and the number of correctly produced words was counted following common scoring procedures (e.g., Gollan et al., 2002; Kavé, 2005).

2.2.2. Word correlation lexical networks

The semantic fluency data were analyzed using a recently developed small-world network modeling methodology (Kenett, Beaty, Silvia, Anaki, & Faust, 2016; Kenett et al., 2013; see also Goñi et al., 2011; Lerner, Ogrocki, & Thomas, 2009). In this network, nodes represent the generated nouns and edges represent word correlations or the tendency to generate a word *b* given that a word *a* was generated.

The lexical networks were constructed as follows. First, separate data matrices were created for responses of bilinguals in English, bilinguals in Hebrew, and native Hebrew speakers in Hebrew. These data matrices were structured such that each row contained all answers of a single participant, and each column was a unique word given by the entire sample. Each cell consisted of either 1, when a participant *i* generated word *j* or 0 when that participant did not say the word. The analyses for each contrast were carried out based on words generated by at least two participants in the two comparison groups. This way, networks that were directly compared to each other were comprised of the same nodes, which is the recommended procedure to control for confounding factors (Kenett et al., 2013; van Wijk, Stam, & Daffertshofer, 2010). Specifically, network parameters are sensitive to differences in the number of nodes; direct comparison of the parameters between networks may therefore yield spurious results (van Wijk et al., 2010). Furthermore, comparing empirical networks with different nodes may result in alternative explanations of the results, such as the effects of the specific nodes examined on the properties of the networks (Kenett et al., 2013). Next, we computed word correlations from the data matrices. The correlations between the words were calculated using Pearson's correlation. This correlation is based on the word generation profile, which specifies the participants who generated that specific word. The more similar the word generation profiles of two words is, the higher the word correlation between them (see Kenett et al., 2013). A word correlation matrix is then created, which contains the word correlations between all pairs of words generated in the sample.

The word correlation matrix can be studied as an adjacency matrix of a fully connected, weighted, undirected lexical network. An adjacency (also known as connectivity) matrix is a means of representing which nodes are adjacent to which other nodes in the network. That is, we created an $n \times n$ matrix in which n represents the number of nodes (words), and each cell represents the relation (word correlation) between all word pairs. Each word is a node in the network, and the correlation between the words is represented as a weighted edge between them. Since most of the edges have small values (weak correlations), the relevant information about the network can be obscured. Several methods have been developed to overcome this obstacle by constructing a subgraph that captures the most relevant information embedded in the original network. Here we used the Planar Maximally Filtered Graph method (Kenett et al., 2011; Tumminello, Aste, Di Matteo, & Mantegna, 2005).

Current network studies of language often treat these networks as undirected (symmetrical relations between nodes) and unweighted (all weights are treated as equal; Borge-Holthoefer & Arenas, 2010). Such an approach is taken to explore coarse grain structural properties of these networks and to control for any higher order information (such as directionality of relations between nodes or the strength of the weights). Since we were interested in the structure of the networks, the networks were binarized such that all edges were converted to a uniform weight = 1, and were then analyzed as unweighted, undirected networks.

2.2.3. Types of network contrasts

Three different types of contrasts were created, with each contrast comparing a separate pair of lexical networks:

Contrast #1 – Hebrew as L2 versus English as L1: Lexical networks were computed for bilingual participants separately for each language. First, each of the unique Hebrew words generated by the bilingual sample was translated into English. Next, we constructed the network for each language based on the words that were generated in both languages.

Contrast #2 – Hebrew as L2 versus Hebrew as L1: Lexical networks were computed separately for the bilinguals and the native Hebrew speakers, based on Hebrew words produced by both groups.

Contrast #3 – Hebrew as L2 versus Hebrew as L1 in the PPVT-matched sub-sample: Lexical networks were computed separately for the Hebrew words produced by the subgroup of bilinguals and native Hebrew speakers who were matched on Hebrew vocabulary.

2.2.4. Network analyses

Analyses were performed with the Brain Connectivity Toolbox for Matlab (Rubinov & Sporns, 2010). The clustering coefficient (CC) and the average shortest path length (ASPL) were calculated (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). The network's CC and ASPL were evaluated qualitatively against the equivalent parameters in a random network with the same number of nodes and edges (CC_{random} and $ASPL_{random}$, respectively). To confirm quantitatively the small-world nature of the networks, S index, which examines the trade-off between the network's CC and ASPL, was also computed (Humphries & Gurney, 2008). S values larger than 1 indicate the network is a small-world. Lastly, the modularity (Q) index was calculated (Newman, 2006). The Q index statistically quantifies how much a network partitions into communities. The larger the modularity index is, the more the network is comprised of communities (Newman, 2006).

To statistically analyze our findings, we used two complementing approaches. The first approach, simulation of random networks, was applied to statistically test the null hypothesis that the network parameters are equal to parameters of a random network. To this end, we generated a large sample of Erdös-Rényi random networks with a fixed edge probability (Boccaletti et al., 2006). The Erdös-Rényi random network model was chosen as it does not assume any assumptions regarding the structure of the network (Erdös & Rényi, 1960). This was done for each category and each contrast independently. For each simulated random network, the CC and ASPL parameters and the Q index were computed. This procedure was simulated with 1000 realizations. This resulted in a random reference distribution for each parameter. The empirical network parameters of each lexical network were then compared to their reference distribution to evaluate their statistical significance. This was achieved via a one-sample Z-test for each network parameter.

Second, we used the bootstrap method (Efron, 1979) to simulate and then compare random partial lexical networks for each of the contrasts. We reasoned that if the two lexical networks of a specific contrast differed from each other, then any subnetwork consisting of the same nodes in both networks should also be different. Furthermore, the bootstrap method makes it possible to generate many simulated partial lexical networks, allowing for statistical examination of the difference between any two networks. The bootstrapping procedure involved random selection of half of the nodes in each contrast. Partial lexical networks were constructed for each group separately for these random words. This method is known as without replacement bootstrap (Bertail, 1997; Politis & Romano, 1994; Shao, 2003). Finally, for each partial lexical network, the CC, ASPL, and the Q index were computed. This procedure was simulated with 1000 realizations. The difference between the bootstrapped partial networks on each network parameter was then tested using an independent samples t-test.

3. Results

3.1. Number of correct responses

For the fruits and vegetables category, the Contrast #1 comparison of English and Hebrew output within the bilingual participants was significant, $t(23) = 3.49$, $p = 0.002$, with more words produced in L1 ($M = 21.52$, $SD = 4.99$) than in L2 ($M = 17.7$, $SD = 4.95$). The Contrast #2 comparison of Hebrew output in L1 and L2 speakers was also significant, $t(49) = 4.25$, $p < 0.001$, with more words produced by the native Hebrew speakers ($M = 23.07$, $SD = 5.13$) than by the bilinguals. The Contrast #3 comparison of Hebrew output in vocabulary-matched bilinguals ($M = 19.00$, *SD* = 4.57) and in native Hebrew speakers (*M* = 21.59, *SD* = 4.34) was not significant, $t(22) = 1.42$, $p = 0.17$.

These results were replicated in the animal category. Bilinguals provided significantly fewer animal names in Hebrew (*M* = 16.04, *SD* = 6.89) than they did in English (*M* = 24.13, *SD* = 5.30), $t(23) = 6.23$, $p = 0.001$. They also generated significantly fewer Hebrew animal names than did native Hebrew speakers (*M* = 24.41, *SD* = 7.31), $t(49) = 4.25$, $p < 0.001$. There was, however, no significant difference between PPVT-matched bilinguals (*M* = 19.08, *SD* = 6.86) and native Hebrew speakers (*M* = 23.25, *SD* = 5.07), $t(22) = 1.69$, $p = 0.10$.

3.2. Word correlation lexical networks

Network parameters were calculated for each of the three contrasts (see Table 2). For the fruits and vegetables category, all six lexical networks were small-world in nature, as evident when comparing the CC and the ASPL of each lexical network to the corresponding parameters in a random network ($CC > CC_{random}$, $ASPL \geqslant ASPL_{random}$). For example, the CC of the Hebrew as L2 lexical network in Contrast #1 was larger than the CC of the random network ($CC_{L2} = 0.66 > CC_{random} = 0.10$), and the ASLP of this empirical network was also larger than but comparable to its random network counterpart ($ASPL_{L2} = 2.94 \geqslant ASPL_{random} = 2.47$). The small-world-ness index further suggests that the lexical networks had a small-world property (all S values above 1). Similar results were observed for the animal category output, indicating that the small-world-ness property was characteristic of this category as well (see bottom panel of Table 2).

Turning to the comparison between languages, further inspection of Table 2 indicates that the CC was larger and the Q was smaller for Hebrew as L2 than for L1, both in the fruits and vegetables and the animal category (e.g., $CC_{L2} = 0.66 > CC_{L1} = 0.61$ in Contrast #1 of the fruits and vegetables category). This was true across the three contrasts; in other words, regardless of whether the L1 was English or Hebrew. The ASPL of the L2 network was smaller than the ASPL of the L1 network, but only in Contrast #1 of both categories. The results across the two semantic categories thus indicate that the L2 network is more densely connected locally and less modular compared to the L1 network.

These properties of the L2 lexical network are further illustrated in Figs. 1 and 2 (generated by the Cytoscape software; Shannon et al., 2003). The graphs demonstrate that the L2 lexical network, as a whole, is less spread-out and less compartmentalized compared to the L1 network, both in the entire bilingual sample (Fig. 1) and in the sub-sample of the most proficient bilinguals (Fig. 2).

3.3. Simulation of random networks

The values of the three network parameters (CC, ASPL, and Q) included in each contrast were compared against their reference simulated random network distribution to assess their statistical significance. This analysis revealed that the network parameters of each lexical network, for both categories, were significantly different from their simulated random parameters (all *p*'s < 0.001; Table 3).

3.4. Bootstrapped partial networks

To test the significance of the differences between the lexical networks, statistical analyses were conducted on the bootstrapped partial networks. In the fruits and vegetables category, independent samples *t*-tests on the CC and the Q index revealed a significant difference between the two distributions in each contrast. This indicates that the CC was larger and the Q was smaller for Hebrew as L2 than they were for English and for Hebrew as L1 (see Table 4). The ASPL showed less stable group differences across the comparisons. While it was smaller for Hebrew as L2 network than for L1 networks in Contrast #1 and #2, the difference was no longer significant in the PPVT-matched sub-sample (Contrast #3).

The results for the animal category were similar: across contrasts, the CC was larger and the Q was smaller for the Hebrew as L2 network than the L1 networks (Table 4). Notably, the effect size was large in most of these comparisons (Cohen's *d* greater than 0.8). The ASPL was significantly smaller for the Hebrew as L2 network compared to the English network (Contrast #1); however, this difference was no longer significant in Contrast #2 and #3.

3.5. Testing order effects

Bilinguals show reduced verbal fluency in a dominant language after performing the task in a non-dominant language, but production in a non-dominant language is not affected by prior production in the other language (Van Assche, Duyck, & Gollan, 2013).

Table 2
Parameters of the word correlation lexical networks for two semantic categories.

	Contrast #1		Contrast #2		Contrast #3	
	Hebrew as L2	English as L1	Hebrew as L2	Hebrew as L1	Hebrew as L2	Hebrew as L1
Fruits and vegetables						
N	60	60	68	68	50	50
CC	0.66	0.61	0.67	0.62	0.63	0.60
ASPL	2.94	3.01	3.63	3.49	3.19	3.07
S	6.20	5.79	5.62	5.38	4.11	3.58
Q	0.53	0.56	0.55	0.58	0.54	0.56
CC_{random}	0.10	0.08	0.08	0.09	0.10	0.13
$ASPL_{random}$	2.47	2.40	2.52	2.59	2.37	2.32
Animals						
N	78	78	80	80	63	63
CC	0.69	0.66	0.70	0.65	0.70	0.63
ASPL	3.75	3.82	3.43	3.32	3.51	3.43
S	6.73	6.26	7.62	7.30	5.89	5.38
Q	0.59	0.63	0.58	0.60	0.56	0.60
CC_{random}	0.09	0.08	0.09	0.07	0.09	0.09
$ASPL_{random}$	2.58	2.63	2.61	2.60	2.49	2.47

Note. L1 = native language; L2 = second language; N = number of nodes in the network; CC = clustering coefficient; ASPL = average shortest path length; S = small-world-ness index; Q – modularity index; CC_{random} = clustering coefficient of random network; $ASPL_{random}$ = average shortest path length of random network.

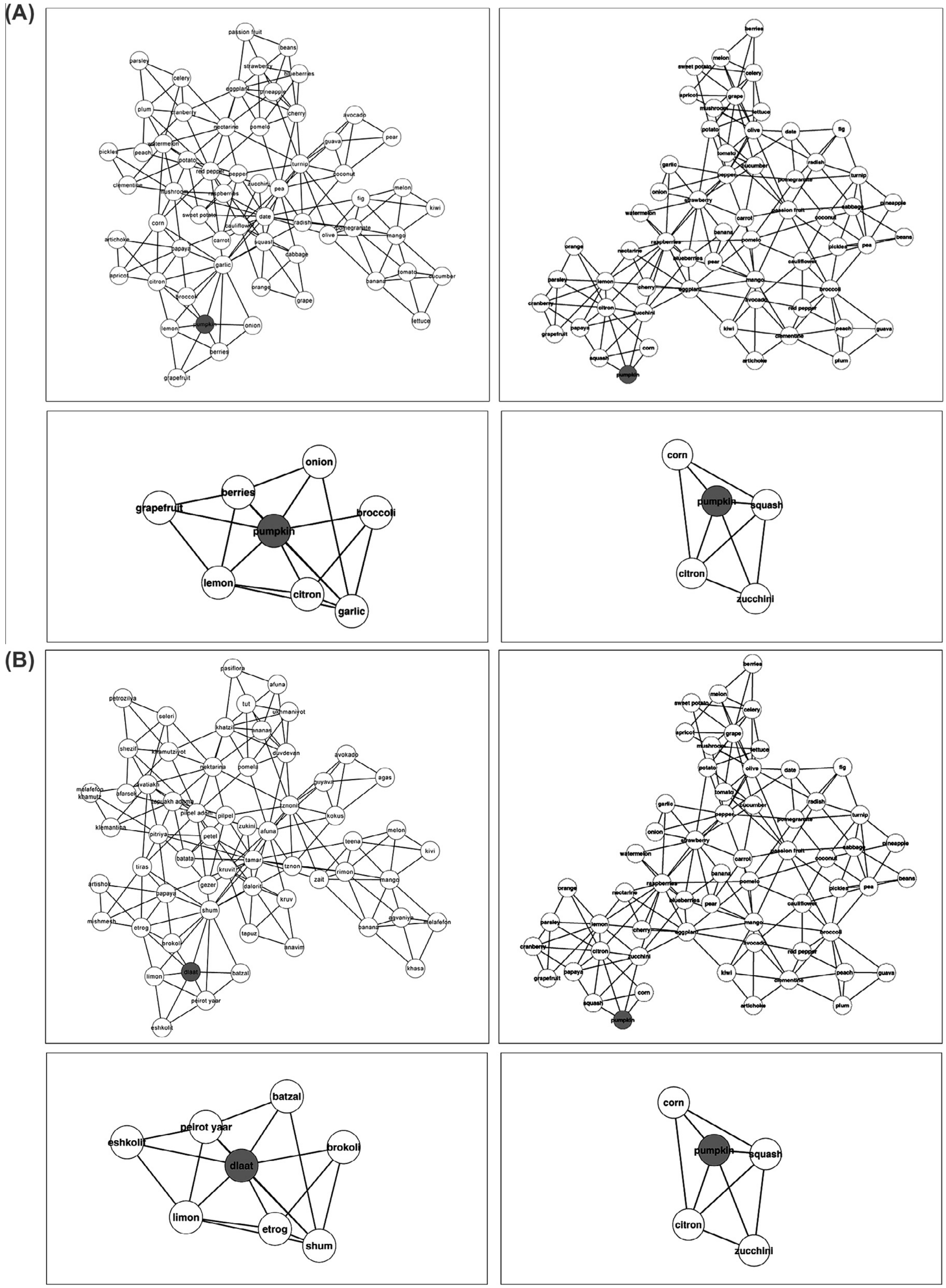

Fig. 1. 2D visualization of the lexical networks of fruits and vegetables category, as analyzed in Contrast #1 (A: words in Hebrew were translated to English; B: responses in Hebrew were transliterated using Latin alphabet). Hebrew as L2 data are represented on the left and English as L1 – on the right. Upper panels show the full network and the lower panels – the node *pumpkin* and its direct neighbors. It should be mentioned that, although many words are grouped the way one would expect based on a taxonomic or a thematic classification (e.g., *pumpkin* is linked to *squash*, *corn*, and *zucchini* in the English as L1 network), semantic-lexical organization may rely on additional factors, including personal, experience-based relations (Body & Muskett, 2012; De Deyne, Verheyen, & Storms, 2016). Consequently, some connections may seem random (such as *citron* [a lemon-like fruit with a thick rind] as a direct neighbor of *pumpkin* in the aforementioned English as L1 network) or identical nodes may have unequal numbers of neighbors across samples (as the lower panels illustrate for *pumpkin*).

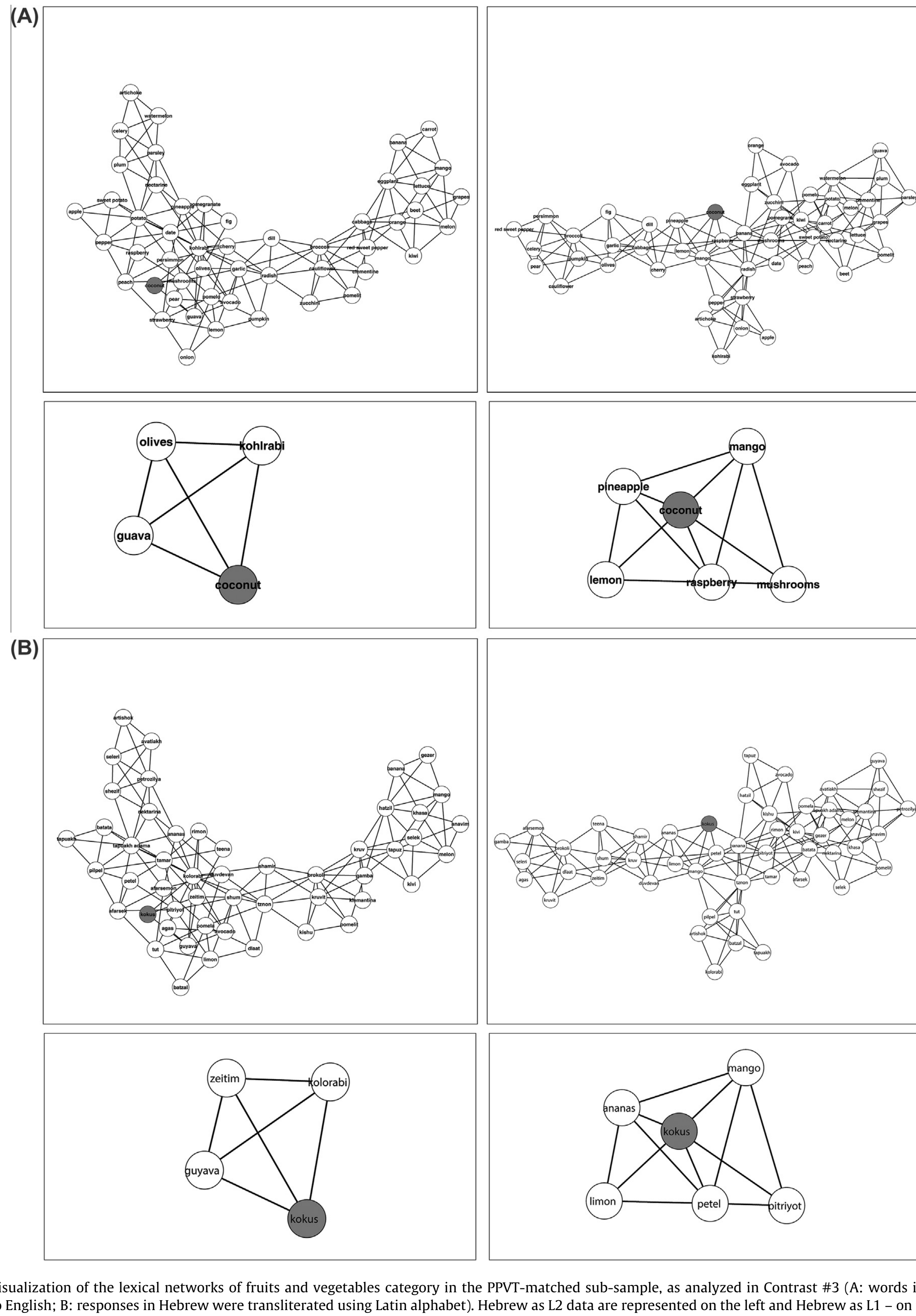

Fig. 2. 2D visualization of the lexical networks of fruits and vegetables category in the PPVT-matched sub-sample, as analyzed in Contrast #3 (A: words in Hebrew were translated to English; B: responses in Hebrew were transliterated using Latin alphabet). Hebrew as L2 data are represented on the left and Hebrew as L1 – on the right. Upper panels show the full network and the lower panels – the node *coconut* and its direct neighbors.

In the current research, languages were tested in a counterbalanced order across bilinguals, and thus the analyses of the English output (L1, the dominant language) may have been affected by the bilinguals who were tested first in Hebrew (L2, the non-dominant language). To examine this possibility, we compared the fluency performance in English of the two bilingual subsamples, those tested in L2 first and those tested in L1 first. In the fruits and vegetables category, consistent with previous findings, bilinguals

Table 3
Results of simulated random network analysis.

	Contrast #1		Contrast #2		Contrast #3	
	Hebrew as L2	English as L1	Hebrew as L2	Hebrew as L1	Hebrew as L2	Hebrew as L1
Fruits and vegetables						
CC	27.38	24.95	34.45	31.51	20.02	18.85
ASPL	5.61	6.41	13.29	11.61	9.00	7.73
Q	8.35	9.61	8.71	9.94	8.31	9.05
Animals						
CC	39.03	37.13	39.60	36.44	32.16	28.47
ASPL	14.13	14.99	9.97	8.65	11.75	10.83
Q	10.78	12.53	10.10	10.96	9.15	10.77

Note. L1 = native language; L2 = second language; CC = clustering coefficient; ASPL = average shortest path length; Q – modularity index.
Values represent Z-scores.

tested in Hebrew first ($M = 15.27$, $SD = 3.29$) produced significantly fewer responses in English compared to bilinguals tested in English first ($M = 21.46$, $SD = 3.78$), $t(22) = 4.24$, $p < 0.001$. The same trend was observed in the animals category ($M = 18$, $SD = 3.63$ for the Hebrew first order and $M = 20.77$, $SD = 4.80$ for the English first order), but here the difference was not significant, $t(22) = 1.57$, $p = 0.13$. Unfortunately, we were unable to complete the analyses comparing the English as L1 network measures for the English first and Hebrew first subsamples due to low statistical power of these contrasts (the total number of nodes was 35 and 40 for the fruits and vegetables and animal category, respectively).

4. Discussion

Using network science tools, we were able to demonstrate that the CC measure was larger and the Q index was smaller in the L2 network compared to these measures in the L1 network, indicating that the L2 lexical network was more densely connected at the local level and its community structure was less modular. These differences in network parameters were observed across two semantic categories (fruits and vegetables and animals) and across different contrasts, both when L1 and L2 were compared using a within-subjects design (in a bilingual sample) and a between-subjects design (Hebrew was compared as L1 in one sample and as L2 in another). The same pattern of results also emerged in a sub-sample of highly proficient bilinguals, whose L2 vocabulary was comparable to that of native speakers.

Our findings suggest that L2 network is less well-organized than its L1 equivalent, in that words in this network are less likely to group into identifiable subcategories (here, of fruits and vegetables/animals), as illustrated in the figures. In Fig. 2, for instance, the target *coconut* in the Hebrew as L1 network (upper left panel) is surrounded by other representatives of the tropical fruits family (e.g., *mango, pineapple, banana*). The cluster below contains several items that can be described as root vegetables (e.g., *artichoke*, *onion*, *kohlrabi*, *radish*). To the left of the tropical fruits community, there is a cluster of vegetables often used in a garden vegetables soup (e.g., *cabbage*, *cauliflower, dill*, *garlic*, *broccoli*, *celery*). The Hebrew as L2 network (upper right panel), on the other hand, shows less separation among the communities (less modularity) and denser neighborhoods (larger CC). For example, the target *coconut* is embedded within a large word cluster, where tropical fruits (e.g., *pineapple*, *guava*) are located alongside the more common summer fruits (e.g., *pear*, *peach*) as well as root vegetables (e.g., *kohlrabi*, *onion*, *potato*). The suboptimal organization of the L2 lexical network is also evident when zooming into the single word level (e.g., Fig. 1, lower panels). For instance, the word *pumpkin* in the English as L1 network shares its (squashes) family with half of its direct neighbors. In L2, on the other hand, the same target word has more diverse neighbors, and none of them is a member of the squashes family (e.g., *garlic* and *onion*, which are often perceived as flavorings).

Our findings support and extend previous research using the semantic fluency task in bilinguals. In the past, the clustering and switching patterns in responses were analyzed to explore the differences in bilingual lexical networks (Roberts & Dorze, 1997; Rosselli & Ardila, 2002; Salvatierra et al., 2007). Besides the element of subjective judgment inherent to this scoring method, there is a disagreement as to what the scores represent: the amount of associative links, the strength of associative links, or the search strategies (Roberts & Dorze, 1994, 1997; Roberts & Le Dorze, 1998). The network science approach adopted here offers more specific tools, which enabled us to show that the L1 and L2 lexical networks differ at the local and the intermediate level of network connectivity. L2 words are more densely connected to their neighbors and less tend to group into neighborhoods compared to L1 words. Further, the previously reported differences in lexical organization were observed in one semantic category, animals (Roberts & Dorze, 1997; Rosselli & Ardila, 2002; Salvatierra et al., 2007). In simultaneous bilinguals, these differences were attributed to a rather specific context in which the animal names were learned by the bilinguals in the sample (Roberts & Dorze, 1997). We demonstrate that in sequential bilinguals, the structural differences between the L1 and the L2 lexical networks are present in more than one semantic category, and thus, are likely more general.

There may have been several factors, including L2 proficiency, that contributed to the differences between the L1 and L2 network

Table 4
Results of bootstrapped partial networks analysis.

	Contrast #1				Contrast #2				Contrast #3			
	Hebrew as L2	English as L1	*t* value	Cohen's *d*	Hebrew as L2	Hebrew as L1	*t* value	Cohen's *d*	Hebrew as L2	Hebrew as L1	*t* value	Cohen's *d*
Fruits and vegetables												
CC	0.673 (0.01)	0.666 (0.01)	7.12***	0.32	0.674 (0.02)	0.670 (0.02)	3.92***	0.18	0.682 (0.02)	0.671 (0.02)	2.42*	0.11
ASPL	2.89 (0.21)	2.91 (0.24)	−2.02*	0.09	2.89 (0.21)	2.92 (0.23)	−2.6*	0.12	2.55 (0.16)	2.55 (0.16)	0.35	0.02
Q	0.48 (0.05)	0.50 (0.03)	−7.94***	0.49	0.52 (0.04)	0.56 (0.01)	−26.45***	1.38	0.48 (0.04)	0.50 (0.01)	−21.61***	0.69
Animals												
CC	0.683 (0.02)	0.665 (0.02)	20.3***	0.91	0.688 (0.02)	0.669 (0.03)	20.27***	0.93	0.682 (0.01)	0.671 (0.02)	12.43***	0.56
ASPL	2.87 (0.19)	2.90 (0.22)	−3.3***	0.15	2.80 (0.18)	2.81 (0.21)	−1.48	0.07	2.97 (0.25)	2.95 (0.27)	1.33	0.06
Q	0.54 (0.03)	0.56 (0.02)	−13.14***	0.78	0.52 (0.04)	0.55 (0.02)	−20.42***	0.95	0.53 (0.04)	0.56 (0.01)	−22.04***	1.03

Note. L1 = native language; L2 = second language; CC = clustering coefficient; ASPL = average shortest path length; Q – modularity index. Values represent distribution mean (standard deviations are in parentheses).
* $p < 0.05$.
*** $p < 0.001$.

parameters. Bilinguals in our sample produced fewer correct responses in Hebrew (L2) compared to English (L1) and compared to native speakers of Hebrew. These differences were no longer significant in the PPVT-matched sub-samples, in agreement with previous research on the effects of L2 vocabulary size on verbal fluency performance (Bialystok et al., 2008; Friesen et al., 2014; Luo et al., 2010). Despite the comparable semantic fluency performance, the L2 lexical network in these highly proficient bilinguals was still characterized by greater local connectivity and reduced modularity compared to the network of native speakers, much like the L2 network of the entire bilingual sample. This result might, in fact, indicate that proficiency alone does not have a profound effect on L2 network structure. This conclusion, however, should be taken with caution, given the small sample size of the matched sub-samples and the large variability in the PPVT scores of the PPVT-matched bilinguals. As Table 1 shows, many of the bilinguals scored below 60, while some performed particularly well; the scores of native Hebrew speakers, on the other hand, all ranged between 63 and 68 (except for one participant with a lower score). The question of whether or not a more homogenous bilingual group, with greater vocabulary size, would have yielded a L2 lexical network with structural characteristics more similar to those of a L1 network is an empirical one and clearly deserves further attention in future studies.

The structural properties of a lexical-semantic network can be also modified by an atypical language learning environment (Beckage et al., 2011; Bilson et al., 2015; Kenett et al., 2013). L1 and L2 language learning history in our bilingual sample differed in important ways. The participants were exposed to Hebrew at a later age compared to English (L1) and received extensive exposure (through immersion) to Hebrew only during puberty. They also reported using Hebrew less than English, despite living in Israel. Thus, later age of language acquisition and insufficient language use and exposure might have had resulted in a suboptimal lexical network in L2 that deviates in its structural characteristics from that of the L1 network. The effects of such factors have been explored on some aspects of lexical knowledge, such as lexical categorization (Malt & Sloman, 2003; Zinszer et al., 2014; see also Jiang & Forster, 2001; Li, 2009; Silverberg & Samuel, 2004), but not in relation to connectivity between words. Thus the following account of the effects we observed in the current research is only speculative. It has been proposed that in learners who became exposed to L2 after early childhood, the lexical-semantic network of L1 is first borrowed to represent the links between L2 words (e.g., Pavlenko, 2009). With exposure to and use of L2, especially in immersion context, this network undergoes reorganization, which continues for many years (Malt & Sloman, 2003), to accommodate L2 language- and culture-specific features. This rewiring of L2 lexicon may involve various processes, such as establishing fine-grained distinctions for contrasts not encoded in L1. For instance, a speaker of Hebrew as L2 may initially group together the words *sharav*, *khamsin*, and *khom* (all of which roughly correspond to *heat* in English), and only after a sufficient exposure to Hebrew and experience living in Israel differentiate between the three by distancing the first two (which refer to heat waves and dust storms that happen a few times in a year, mostly in the spring and the early summer) from the last. Higher local connectivity in the L2 lexical network may thus be in part a product of this process of differentiation between closely related words. Rewiring can also occur as a result of re-adjustment of the semantic features of extant concepts to match the new environment, leading to less modularity in L2 community structure. For example, for an American English speaker who acquires Hebrew while living in Israel, *fig* may become more connected to the more common summer fruits, such as *pear* and *plum*, because it is native to Israel and easily available, but it may also preserve its L1 connections to the more exotic, tropic fruits, such as *pineapple* and *papaya*. Clearly, further research is required to directly examine the effects of factors related to language learning history (e.g., age of acquisition effects and frequency of language use) and their mechanisms of influence on the structural properties of bilingual lexical networks.

The present study has several limitations. As mentioned before, our findings may be restricted to bilinguals who received extensive exposure to L2 only after puberty and who continue to use L1 more often than L2. Second, previous studies have demonstrated that semantic relations among words as opposed to associative relations are particularly difficult for L2 learners to process (Kotz, 2001; Kotz & Elston-Güttler, 2004), thus it is possible that our findings are limited to the representation of semantic information. On a related note, our findings do not necessarily mean that the differences between languages in lexical networks were driven exclusively by differences in the representation of semantic relations. Other factors, for example, phonological relatedness may have also affected L2 semantic fluency performance, as has been previously argued for other tasks (Meara, 1978; Namei, 2004; Söderman, 1993). Although inspection of the transliterated neighboring words in the Hebrew as L2 networks (see Figs. 1B and 2B) provides little evidence supporting this suggestion, we cannot entirely rule it out. The role of such links might be more directly studied in a task that allows a greater range of responses (e.g., word associations) or if a phonological criterion for what establishes an association between words (e.g., Vitevitch, 2008) is applied.

Third, we found evidence of language order effects in our behavioral data, suggesting an inhibitory effect of L2 on L1 production (Van Assche et al., 2013). It is thus possible that the observed structural differences between the languages were a by-product of inhibition of the dominant language processing. Although we were unable to directly test language order effects on network measures due to low statistical power of these comparisons, we believe this account is less likely given that the same L1-L2 differences were observed when only responses in the non-dominant language, Hebrew, were analyzed (Contrast #2 and #3), where no inhibition is expected. Finally, although the L1 and L2 lexical networks showed consistent differences in the CC and the Q parameters across comparisons, there was no stable trend in the ASPL parameter. In Contrast #1 the L2 network had a shorter ASPL, while in Contrast #2 and #3 the ASPL index was greater in the L2 than the L1 network. This inconsistency was also evident in the bootstrap analysis, which resulted in weak or no significant differences between the two bootstrapped distributions. While network analyses in Contrasts #2 and #3 were performed on the Hebrew words directly provided by participants, the English words comprising the L2 lexical network in Contrast #1 were first translated from Hebrew. The words in the two semantic categories were highly concrete; nevertheless, post hoc translations by three fluent Hebrew-English speakers suggested a few of them had more than one translation equivalent (4 out of 68 fruit and vegetable names and 6 out of 78 animal names). The multiplicity in translation equivalents may have affected the selection of common nodes for L1 and L2 networks and thus possibly the stability of the ASPL index across the contrasts. A more theoretically interesting explanation is that factors related to L2 language learning history may affect more the local and the intermediate than the global connectivity of L2 lexical network. Clearly, this finding deserves further investigation.

In conclusion, computational tools allowed us to study the organization of the L2 mental lexicon. Our results show that even in highly proficient speakers, whose L2 vocabulary was matched to the vocabulary of native speakers, the lexical network had a different, less optimal structure compared to the structure of the lexical network in L1. Recently, researchers have begun to examine the effects of structural relationships in the lexicon on language processing. It has been demonstrated, for example, that the CC computed from a network of phonologically related words (Vitevitch, 2008) influences monolingual word recognition (Chan & Vitevitch, 2009), word production (Chan & Vitevitch, 2010), and word learning (Goldstein & Vitevitch, 2014). Thus, given the many differences observed between L1 and L2 (in the language production domain, for example, slower and less accurate naming in L2 is often reported; reviewed in Hanulová, Davidson, & Indefrey, 2010), our findings may have significant implications for the research of bilingual language processing.

Acknowledgments

This work was partially supported by the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation (grant no. 51/11). We thank Dr. Gitit Kavé for her insightful comments on a previous draft of this manuscript and Omer Maliniak for his assistance in collecting and coding the data for this research.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.07.014.

References

Altarriba, J., & Basnight-Brown, D. (2009). An overview of semantic processing in bilinguals: Methods and findings. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 79–98). Bristol, UK: Multilingual Matters.

Antón-Méndez, I., & Gollan, T. H. (2010). Not just semantics: Strong frequency and weak cognate effects on semantic association in bilinguals. *Memory & Cognition, 38*(6), 723–739. http://dx.doi.org/10.3758/mc.38.6.723.

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology, 41*(5), 324–332. http://dx.doi.org/10.1080/00207590500345542.

Barabási, A.-L. (2009). Scale-free networks: A decade and beyond. *Science, 325*(5939), 412–413.

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences, 17*(7), 348–360. http://dx.doi.org/10.1016/j.tics.2013.04.010.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE, 6*(5), e19348. http://dx.doi.org/10.1371/journal.pone.0019348.

Bertail, P. (1997). Second-order properties of an extrapolated bootstrap without replacement under weak assumptions. *Bernoulli, 3*(2), 149–179.

Bialystok, E., Craik, F., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics, 21*(6), 522–538. http://dx.doi.org/10.1016/j.jneuroling.2007.07.001.

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition, 140*, 122–134. http://dx.doi.org/10.1016/j.cognition.2015.03.013.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(175–308).

Body, R., & Muskett, T. (2012). Pandas and penguins, monkeys and caterpillars: Problems of cluster analysis in semantic verbal fluency. *Qualitative Research in Psychology, 10*(1), 28–41. http://dx.doi.org/10.1080/14780887.2011.586104.

Borge-Holthoefer, J., & Arenas, A. (2010). Semantic networks: Structure and dynamics. *Entropy, 12*(5), 1264–1302.

Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering-coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1934–1949. http://dx.doi.org/10.1037/a0016902.

Chan, K. Y., & Vitevitch, M. S. (2010). Network structure influences speech production. *Cognitive Science, 34*(4), 685–697. http://dx.doi.org/10.1111/j.1551-6709.2010.01100.x.

De Deyne, S., & Storms, G. (2008). Word association: Network and Semantic properties. *Behavior Research Methods, 40*(1), 213–231.

De Deyne, S., Verheyen, S., & Storms, G. (2016). Structure and organization of the mental lexicon: A network approach derived from syntactic dependency relations and word associations. In A. Mehler, A. Lücking, S. Banisch, P. Blanchard, & B. Job (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 47–79). Berlin Heidelberg: Springer.

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York: Psychology Press.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test fourth edition (PPVT-IV)*. Bloomington, MN: NCS Pearson.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 1–26.

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Science, 5*, 17–61.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *The American Journal of Psychology, 74*, 361–372. http://dx.doi.org/10.2307/1419742.

Frenck-Mestre, C., & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language, 37*(4), 481–501. http://dx.doi.org/10.1006/jmla.1997.2526.

Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2014). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience, 1–13*. http://dx.doi.org/10.1080/23273798.2014.918630.

Goldstein, R., & Vitevitch, M. S. (2014). The influence of clustering coefficient on word-learning: How groups of similar sounding words facilitate acquisition. *Frontiers in Psychology, 5*. http://dx.doi.org/10.3389/fpsyg.2014.01307.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*(4), 562–576. http://dx.doi.org/10.1037/0894-4105.16.4.562.

Goñi, J., Arrondo, G., Sepulcre, J., Martincorena, I., Vélez de Mendizábal, N., Corominas-Murtra, B., ... Villoslada, P. (2011). The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cognitive Processing, 12*(2), 183–196. http://dx.doi.org/10.1007/s10339-010-0372-x.

Hanulová, J., Davidson, D. J., & Indefrey, P. (2010). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes, 26*(7), 902–934. http://dx.doi.org/10.1080/01690965.2010.509946.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE, 3*(4), 1–10. http://dx.doi.org/10.1371/journal.pone.0002051.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*(1), 32–51. http://dx.doi.org/10.1006/jmla.2000.2737.

Kavé, G. (2005). Phonemic fluency, semantic fluency, and difference scores: Normative data for adult Hebrew speakers. *Journal of Clinical and Experimental Neuropsychology, 27*(6), 690–699. http://dx.doi.org/10.1080/13803390490918499.

Kenett, Y. N., Beaty, R. E., Silvia, P. J., Anaki, D., & Faust, M. (2016). Structure and flexibility: Investigating the relation between the structure of the mental lexicon, fluid intelligence and creative achievement. *Psychology of Aesthetics, Creativity and the Arts*. http://dx.doi.org/10.1037/aca0000056.

Kenett, Y. N., Kenett, D. Y., Ben-Jacob, E., & Faust, M. (2011). Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. *PLoS ONE, 6*(8), e23912. http://dx.doi.org/10.1371/journal.pone.0023912.

Kenett, Y. N., Wechsler-Kashi, D., Kenett, D. Y., Schwartz, R. G., Ben Jacob, E., & Faust, M. (2013). Semantic organization in children with cochlear implants: Computational analysis of verbal fluency. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00543.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior, 2*(4), 291–300. http://dx.doi.org/10.1016/S0022-5371(63)80097-3.

Kotz, S. A. (2001). Neurolinguistic evidence for bilingual language representation: A comparison of reaction times and event-related brain potentials. *Bilingualism: Language and Cognition, 4*(02), 143–154. http://dx.doi.org/10.1017/S1366728901000244.

Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics, 17*(2–3), 215–235. http://dx.doi.org/10.1016/S0911-6044(03)00058-7.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174. http://dx.doi.org/10.1006/jmla.1994.1008.

Kruse, H., Pankhurst, J., & Smith, M. S. (1987). A multiple word association probe in second language acquisition research. *Studies in Second Language Acquisition, 9*(02), 141–154.

Lerner, A. J., Ogrocki, P. K., & Thomas, P. J. (2009). Network graph analysis of category fluency testing. *Cognitive and Behavioral Neurology, 22*(1), 45–52. http://dx.doi.org/10.1097/WNN.0b013e318192ccaf.

Li, P. (2009). Lexical organization and competition in first and second languages: Computational and neural mechanisms. *Cognitive Science, 33*(4), 629–664.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29–41. http://dx.doi.org/10.1016/j.cognition.2009.08.014.

Malt, B. C., & Sloman, S. A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition, 6*(01), 47–67. http://dx.doi.org/10.1017/S1366728903001020.

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

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*(2), 99–130. http://dx.doi.org/10.1037/0096-3445.126.2.99.
Meara, P. (1978). Learners' word associations in French. *Interlanguage Studies Bulletin-Utrecht, 3*(2), 192–211.
Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam: John Benjamins Publishing Co.
Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience, 4*, 200. http://dx.doi.org/10.3389/fnins.2010.00200.
Namei, S. (2004). Bilingual lexical development: A Persian-Swedish word association study. *International Journal of Applied Linguistics, 14*(3), 363–388.
Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences, 103*(23), 8577–8582. http://dx.doi.org/10.1073/pnas.0601602103.
Nissen, H. B., & Henriksen, B. (2006). Word class influence on word association test results. *International Journal of Applied Linguistics, 16*(3), 389–408. http://dx.doi.org/10.1111/j.1473-4192.2006.00124.x.
Patael, S., Segal, O., Kaplan, D., & Kishon-Rabin, L. (2010). *Peabody picture vocabulary test [Hebrew version of Dunn & Dunn's Peabody Picture Vocabulary Test]* (4th ed.). Tel Aviv, Israel: Tel Aviv University.
Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 125–160). Clevedon: Multilingual Matters.
Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*(5–6), 765–805. http://dx.doi.org/10.1080/016909697386682.
Politis, D. N., & Romano, J. P. (1994). A general theory for large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics, 22*, 2031–2050.
Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*(3), 415–422. http://dx.doi.org/10.1016/j.acn.2007.01.015.
Roberts, P. M., & Dorze, G. L. (1994). Semantic verbal fluency in aphasia: A quantitative and qualitative study in test-retest conditions. *Aphasiology, 8*(6), 569–582. http://dx.doi.org/10.1080/02687039408248682.
Roberts, P. M., & Dorze, G. L. (1997). Semantic organization, strategy use, and productivity in bilingual semantic verbal fluency. *Brain and Language, 59*(3), 412–449. http://dx.doi.org/10.1006/brln.1997.1753.
Roberts, P. M., & Le Dorze, G. (1998). Bilingual aphasia: Semantic organization, strategy use, and productivity in semantic verbal fluency. *Brain and Language, 65*(2), 287–312. http://dx.doi.org/10.1006/brln.1998.1992.
Rosselli, M., & Ardila, A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience, 112*(6), 759–776. http://dx.doi.org/10.1080/00207450290025752.
Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology, 7*(1), 17–24. http://dx.doi.org/10.1207/s15324826an0701_3.
Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage, 52*(3), 1059–1069. http://dx.doi.org/10.1016/j.neuroimage.2009.10.003.
Salvatierra, J., Rosselli, M., Acevedo, A., & Duara, R. (2007). Verbal fluency in bilingual Spanish/English Alzheimer's disease patients. *American Journal of Alzheimer's Disease and Other Dementias, 22*(3), 190–201. http://dx.doi.org/10.1177/1533317507301792.
Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*(02), 231–252. http://dx.doi.org/10.1017/S1366728909990514.
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: A software for integrated models of biomolecular interaction networks. *Genome Research, 13*(11), 2498–2504.
Shao, J. (2003). Impact of bootstrap on sample surveys. *Statistical Science, 18*(2), 191–198.
Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*(3), 381–398. http://dx.doi.org/10.1016/j.jml.2004.05.003.
Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41–78. http://dx.doi.org/10.1207/s15516709cog2901_3.
Söderman, T. (1993). Word associations of foreign language learners and native speakers – Different response types and their relevance to lexical development. In B. Hammarberg (Ed.), *Problems, process, and product in language learning*. Stockholm: University of Stockholm.
Taler, V., Johns, B. T., Young, K., Sheppard, C., & Jones, M. N. (2013). A computational analysis of semantic structure in bilingual verbal fluency performance. *Journal of Memory and Language*. http://dx.doi.org/10.1016/j.jml.2013.08.004.
Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 370–378. http://dx.doi.org/10.1076/1380-3395(200006)22:3;1-v;ft370.
Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138.
Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America, 102*(30), 10421–10426. http://dx.doi.org/10.1073/pnas.0500298102.
Van Assche, E., Duyck, W., & Gollan, T. H. (2013). Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(6), 1781–1792. http://dx.doi.org/10.1037/a0032859.
Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(3), 193–211. http://dx.doi.org/10.1017/s1366728998000352.
van Wijk, B. C. M., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE, 5*(10), e13701. http://dx.doi.org/10.1371/journal.pone.0013701.
Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research, 51*(2), 408–422. http://dx.doi.org/10.1044/1092-4388(2008/030).
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature, 393*(6684), 440–442.
Wilks, C., & Meara, P. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research, 18*(4), 303–324. http://dx.doi.org/10.1191/0267658302sr203oa.
Wilks, C., Meara, P., & Wolter, B. (2005). A further note on simulating word association behaviour in a second language. *Second Language Research, 21*(4), 359–372. http://dx.doi.org/10.1191/0267658305sr251oa.
Zinszer, B. D., Malt, B., Ameel, E., & Li, P. (2014). Native-likeness in second language lexical categorization reflects individual language history and linguistic community norms. *Frontiers in Psychology, 5*. http://dx.doi.org/10.3389/fpsyg.2014.01203.