Assessing automatic processing of hypernymic relations in first language speakers and advanced second language learners

A semantic priming approach

Scott Crossley
Georgia State University

This study investigates the depth of lexical knowledge in first language (L1) speakers and second language (L2) learners in reference to hierarchical word knowledge. Eighty-eight participants took part in a lexical decision task that assessed their speed and accuracy in recognizing words and nonwords. Prime and target pairs in the lexical decision task were related words (hyponynm to hypernym and hypernym to hyponym), unrelated words, or word to nonwords. The findings indicate bidirectional priming in L1 participants such that associated pairs (hyponynm to hypernym and hypernym to hyponym) were processed faster than unrelated words. For L2 participants, unidirectional priming effects were reported for the hyponynm to hypernym condition only. These findings provide evidence that hierarchical lexical networks characterize L1 lexicons but not L2 networks. Such findings provide important information about the organizational properties of L1 and L2 lexicons.

Keywords: semantic priming, hypernymy, lexical networks, second language learning, lexical competence, depth of lexical knowledge

Assessing the lexical competence of English speakers is relatively straightforward for many features of word knowledge (e.g., evaluating breadth of knowledge features such as word frequency and lexical diversity); however, some aspects of lexical competence can only be assessed indirectly (e.g., depth of knowledge features such as collocational knowledge and hierarchical relations between words). These areas of lexical knowledge are generally the result of incidental learning that occurs when speakers focus their attention on word meaning and not word form

The Mental Lexicon 8:1 (2013), **96–116.** DOI 10.1075/ml.8.1.05cro
ISSN 1871–1340 / E-ISSN 1871–1375

(Nation, 2001; Schmitt, 1998; VanPatten & Williams, 2007). For first language (L1) speakers of English, word knowledge is, by and large, acquired implicitly and fluently. For second language (L2) speakers of English, word knowledge develops at a slower rate and involves both implicit and explicit learning. Fluency is generally not expected for L2 learners, but lexical competence at a variety of levels, to include depth of knowledge, is attainable.

One approach to assessing a speakers' depth of lexical knowledge is through corpus and computation analyses of spoken or written data (e.g., Crossley, Salsbury, & McNamara, 2009, 2010a). A second approach is through direct assessment of lexical knowledge (i.e., tests of productive and passive word knowledge; Meara, 2005; Schmitt, 1998). A final approach, and potentially the most promising, is through the use of semantic priming experiments that assess implicit links between related words. Semantic priming experiments investigate the tendency for L1 and L2 users to process a word more accurately and quickly immediately after they have been exposed to a related word. For instance, users will identify, as a word, *cat* more quickly if they have recently been exposed to the word *dog* as compared to if they have been exposed to an unrelated word such as *pen* (McDonough & Trofimovich, 2009). Semantic priming experiments are used to assess how previous language exposure influences lexical processing through an analysis of how words prime lexical chains (i.e., *dog* and *cat*). Many researchers argue that semantic priming reflects the fundamental properties of lexical organization and the manner in which words are stored and retrieved (Neely, 1991; Ratcliff & McKoon, 1988). When used to assess depth of lexical knowledge, priming tasks also have an ecological validity because they depend on implicit processes that tap into implicit memory (McDonough & Trofimovich, 2009). Previously, priming experiments have been used to assess the lexical competence of L1 speakers and bilinguals in a variety of tasks related to depth of lexical knowledge (Devitto & Burgess, 2004; Frenck-Mestre & Prince, 1997; McNamara, 2005; Meyer & Schvaneveldt, 1971); however, their use in L2 environments has been rare.

The purpose of this study is to use semantic priming techniques to assess the strength of hypernymic relations (i.e., links between superordinate and subordinate terms in hierarchies such as weapon → gun → rifle) in the lexicons of L1 and L2 speakers. Such techniques allow for the assessment of the lexical competence of L1 and L2 speakers in terms of their depth of lexical knowledge. A semantic priming task was developed and tested on both an L1 and an L2 population. The priming task affords the opportunity to assess the lexical knowledge of L1 and L2 speakers in reference to priming effects for both hypernyms (superordinate terms) and hyponyms (subordinate terms), which, to my knowledge, have not been examined. Additionally, the priming task affords a comparison of priming effects between the baseline L1 population and the L2 population. Such a comparison provides an opportunity to understand potential differences in the organization of L1 and L2 lexicons.

Lexical Competence

Assessing lexical competence is important because it is an essential element of academic success (Daller, van Hout, & Treffers-Daller, 2003). Lexical knowledge forms the basis for more advanced language skills, including syntax (Ellis, N., & Ferreira-Junior, 2009; Wulff, Ellis, Romer, Bardovi-Harlig, & LeBlanc, 2009) and reading comprehension (Hu & Nation, 2000). Pragmatically, the lexicon is also crucial for communicative success with misinterpretations of lexical items identified as key elements in communication errors (Crossley, Salsbury, & McNamara, 2010b; Ellis, R., 1995; Ellis, R., Tanaka, & Yamazaki, 1994; de la Fuente, 2002). From a theoretical perspective, understanding L2 lexical acquisition in relation to its deeper, cognitive functions can lead to increased awareness of how language users process and produce language (Crossley et al., 2009, 2010a).

A key element of lexical competence is depth of lexical knowledge (i.e., the degree of organization of known words; Meara, 1996, 2005; Qian, 1999; Read, 1998; Wesche & Paribakht, 1996). Unlike breadth of lexical knowledge, depth of lexical knowledge is not based on the number or variety of words a learner produces, but on constraints at the phonemic, morphemic, and syntactic level and at deeper levels related to word associations (e.g. semantic co-referentiality, hypernymy, polysemy, and collocation knowledge; Qian & Schedl, 2004). Word association features, which are of primary interest for this study, are often integrated under the term "lexical network," which serves as a convenient metaphor to describe the manner in which lexical features combine to form complex association models that act categorically to form entire lexicons (Haastrup & Henriksen, 2000; Huckin & Coady, 1999). From an acquisitional perspective, as language learners develop lexical competence, they build lexical networks that are strengthened by differentiating relations between words and within words (Haastrup & Henriksen).

Lexical Networks and Semantic Priming

A key component in semantic priming theory is the notion that lexical networks result from spreading activation. Spreading activation theory argues that activation among related semantic concepts in memory results in priming effects for associated concepts, which are represented as interconnected nodes. The more similar the concepts are, the closer they are stored to one another. The more distant the concept, the farther apart they are stored and the fewer number of links that are shared. Within the network, the activation of a single concept will trigger the activation of all its interconnected conceptual nodes (Collins & Loftus, 1975). Findings from a variety of priming studies support spreading activation theory in that strongly associated words appear to be stored together or, at minimum, to be linked in the mind of the language user. As a result, exposure or production of a word primes the activation of related words (Neely, 1991), resulting in a priming effect in which related words are processed and identified more quickly. Alternatively, priming may not be based on associations between discrete items (i.e., individual words), but through sets of semantic features shared between that facilitate processing (McRae & Boisvert, 1998; Plaut, 1995). Thus, spreading may not be the result of associative relations between words alone (e.g., cradle-baby), but rather shared semantic features (e.g., the shared semantic features found in Zebra-Horse; Perea & Rosa, 2002).

While semantic priming is a common approach to assessing the lexical knowledge of L1 speakers (see McNamara, 2005 for an overview), it has been used less infrequently in assessing the lexical knowledge of L2 learners, with most studies focusing on bilingual lexicons. Two primary research strands are found in L2 priming studies: (1) investigations to assess if similar patterns of priming exist in both the L1 and the L2; and (2) investigation into whether the priming patterns in bilinguals and L2 speakers resemble those found in monolingual speakers (McDonough & Trofimovich, 2009). In reference to the first strand, research has indicated that priming effects are stronger in speakers' L1 than their L2, especially when proficiency differences exist between the languages (Phillips, Segalowitz, O'Brien, & Yamasaki, 2004). In reference to the second strand, research indicates that L1 speakers and advanced L2 speakers access and use some aspects of semantic information in a similar manner for strongly associated words (antonyms, synonyms, and collocations, Frenck-Mestre & Prince, 1997) but not weakly associated words (Devitto & Burgess, 2004).

The majority of within language priming studies in bilingual and L2 research have focused on semantically related words such as *sugar-sweet* (see McDonough & Trofimovich, 2009 for an overview) or collocations (Frenck-Mestre & Prince, 1997). A few studies have also examined synonyms and antonyms (Dong, Gui, & MacWhinney, 2005; Frenck-Mestre & Prince, 1997), but none, to my knowledge, have examined conceptual or hierarchical lexical knowledge (e.g., priming effects between superordinate and subordinate words). Such relations between words have been hypothesized to be the single most important organizational system for lexical relations (Miller & Teibel, 1991).

Hypernymic Relations

Hypernymic relations are hierarchical associations between hypernyms (superordinate words) and hyponyms (subordinate words). A hypernym is less specific than a related word (*animal* compared to *dog*) and a hyponym is more specific than a related word (*dog* as compared to *animal*). A good example of hypernymy is the relationship between *car* and *vehicle* in which *car* is the hyponym of the hypernym *vehicle* because *car* has a narrower and more specific definition than *vehicle*. Hypernymic relations allow for hierarchical categorizations that define how hyponyms inherit properties from their related hypernyms and allow set inclusion among category members. Hypernymy is a foundational lexical relationship that is consistent with network models that allow for the economical representation of lexical properties (Chaffin & Glass, 1990; Murphy, 2004). Such properties allow for learners to generalize terms as well as generate cognitive economy because every object is part of a conceptual category and not isolated (Murphy). A categorical relation such as hypernymy that is based on an economy of representations should afford faster processing of language as compared to sense relations such as synonymy or polysemy, which are argued to be more difficult to process (Chaffin & Glass).

Knowledge of hypernymic relations between words is associated with a speaker's aptitude for managing academic and formal registers (Snow, 1990; Snow, Cancino, Gonzalez, & Shriberg, 1989; Snow, Cancino, De Temple, & Schley, 1991). For instance, the more informal speech is, the more hypernyms are used (i.e., *I drove the car there*), and the more formal speech is, the more hyponyms are used (i.e., *I drove the Cadillac to the country club*; Ordonez, Carlo, Snow, & McLaughlin, 2002). From a developmental perspective, hypernymic relations are more likely acquired as the learners advance cognitively (Anglin, 1993; Snow, 1988; Vygotsky, 1962), as they increase their levels of education (LeVine, 1980; Snow, 1990), as they acquire more specific lexical knowledge (Wolter, 2001), and in an L1 rather than an L2 (Meara, 1982; Söderman, 1993). Studies have demonstrated that at a very young age, children can distinguish between hypernyms and hyponyms in their L1 and that children first develop superordinate categories (Berlin, Breedlove, & Raven, 1974; Brown, 1958; Mervis & Crisafi, 1982; Murphy, 2004) followed by subordinate categories (Berlin et al., 1974).

Second language research has demonstrated that L2 learners do not have the same access to hypernymic relations as L1 speakers. For instance, Levenston and Blum (1977) found that L2 learners use more words of general than of specific meanings, causing the L2 learners to make inappropriate overgeneralizations and to not adhere to expected uses of registers and collocations. Ijaz (1986) found that L2 learners associated prepositions most often with contexts that were prototypical, but they failed to attribute similar relationships to non-prototypical categories. More recent studies have demonstrated that L2 learners produce more words that are less specific as time is spent studying English (Crossley et al., 2009) indicating that L2 learners may first learn general words and then move toward the production of hypernyms.

Methods

The purpose of this study is to examine associative priming effects for hypernyms and hyponyms in a within-language priming context in order to investigate whether hierarchical relations between words influence how words are stored and processed in the mental lexicons of advanced L1 speakers and L2 speakers. In this study, L1 and advanced L2 speakers of English were given a lexical decision task to complete in which priming effects for both hypernyms and hyponyms were tested. A semantic priming approach was selected because it is argued to test automatic processes and to reflect fundamental properties of speakers' mental lexicons such as how the lexicon is accessed and how words are retrieved. Semantic priming approaches also afford the examination of whether prior exposure to a related word facilitates processing (McDonough & Trofimovich, 2009). It is predicted that hypernyms and hyponyms will show significant facilitative effects for lexical decisions involving related words. This hypothesis is premised on the presumed centrality of hypernymic relations in developed lexicons (Chaffin & Glass, 1990; Miller & Teibel, 1991) and on the notion that less economical sense relations (i.e., synonymy; Chaffin & Glass, 1990) have demonstrated priming effects for both L1 and L2 speakers (Frenck-Mestre & Prince, 1997).

Participants

Eighty-four undergraduate students from Georgia State University volunteered for the experiment. Participants were assigned to two language conditions according to their self-identification as either second language (L2) speakers of English or first language (L1) speakers of English. Forty-one of the participants were L2 speakers of English and 43 of the participants were native speakers of English. All students received class credit in a freshman Psychology course for participating in the experiment.

Demographic data was self-reported by the participants. The average age for the L1 speakers was 21 years old. The average grade point average (GPA) for the L1 participants was 3.24. Thirty-seven percent of the L1 speakers in the study were male, while the other 63% were female. The average age of the L2 speakers was 25 years old and they also reported an average GPA of 3.24. As with our L1 participants, 37% of the L2 speakers in the study were male, while the other 63% were female. The L2 speakers came from 22 different L1 backgrounds of which the three most common were Spanish (ten L1 speakers), Chinese (five L1 speakers), and Urdu (three L1 speakers). On average, the L2 participants had studied English for 12 years and began studying English at the age of 12.

All L2 speakers admitted to Georgia State university are required to score above a 430 on the verbal section of the Scholastic Aptitude Test (SAT) or a 17 on the English section of the American College Test. If the L2 speakers reside in a country other than the United States where English is not the dominant language, they must either score a 550 on the TOEFL exam or pass through six levels of the University's Intensive English Program (generally a year-long program). Knowing these constraints, all L2 speaking participants were classified as advanced level English speakers. For those L2 participants that reported a TOEFL score, the average score for the TOEFL Internet-Based Test was 106.666 ($n = 12$). The average score for the paper-based test was 595.666 ($n = 3$).

Materials and Study Design

For this experiment, 60 prototypical words from English (i.e., *doctor, dog, job, movie*) were selected as primes along with 60 related hypernyms and 60 related hyponyms as target words. We operationalized prototypicality as a function of word length and word frequency (under the presumption that more prototypical words would be shorter and more frequent, Vergès, 1992). The hypernyms and hyponyms for each prototypical word were selected from the WordNet database (Fellbaum, 1998; Miller, Beckwith, Fellbaum, Gross, & Miller, 1990), which provides lists of hypernyms and hyponyms for over 100,000 English words. The principle semantic category for the majority of the 180 primes and targets was noun; however, some of the words had multiple semantic categories (i.e., the words could function as nouns and verbs such as *ground, name,* and *drink*). None of the selected words contained inflectional or derivational endings. Two experts reviewed the word pairs to ensure that phonological similarities among the pairs were minimal. Sixty nonwords were also randomly selected from ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). These nonwords were used as targets in the lexical decision task in which participants were asked to judge if a string of letters was an English word or a nonword. All nonwords contained orthographically existing onsets and bodies and thus followed the phonotactic structure of English. None of the nonwords contained inflectional or derivational endings. Table 1 contains a list of the words and nonwords used in this experiment.

Table 1. Words and Nonwords Used in Experiment

Prototype Prime Words	Hyponym to Hypernym Condition	Hypernym to Hyponym Condition	Nonwords
TEAM	UNIT	CREW	CRARP
CAR	VEHICLE	AMBULANCE	SKERK
COUNTRY	TERRITORY	CHINA	SWEEL
BOSS	LEADER	FOREMAN	NIPE
PARTY	AFFAIR	DANCE	KNARSE
GUN	WEAPON	RIFLE	CLEINT
FRIEND	PERSON	BUDDY	VISPE
HOSPITAL	STRUCTURE	ASYLUM	TWURN
MONEY	CURRENCY	DOLLAR	RUP
OFFICE	PLACE	ROOM	DEUD
TOWN	LOCATION	PARIS	WHUFF
DOCTOR	SCHOLAR	SURGEON	GLERTH
DREAM	FANTASY	NIGHTMARE	DWEF
YELLOW	COLOR	GOLD	SMOAP
BAR	AREA	LOUNGE	SLOUR
DECISION	CHOICE	RESOLUTION	PRINE
DAY	TIME	TOMORROW	CLALK
SEAT	SPACE	CHAIR	GRUCH
NEWS	INFO	REPORT	PAFT
WORLD	OBJECT	NATURE	DRERGE
WOMAN	FEMALE	WIFE	CLERN
FATHER	PARENT	DAD	SMIN
PEACE	ORDER	TRUCE	FRING
LOVE	EMOTION	WORSHIP	WROOB
PICTURE	ART	PHOTO	CHEUTH
HOUSE	HOME	CABIN	GWEIGH
GIFT	POSSESSION	PRIZE	ZIM
SCHOOL	INSTITUTION	ACADEMY	FRAWP
MOVIE	SHOW	MUSICAL	SPRORD
WIND	WEATHER	BREEZE	PHLOLL
KID	JUVENILE	ORPHAN	PSAWK
BAG	CONTAINER	PURSE	WRUMF
PLAN	IDEA	PLOT	KNUD
STREET	ROAD	ALLEY	SNIMPSE
DOG	ANIMAL	PUPPY	FURK
HAIR	COVERING	BEARD	GWERB
SONG	MUSIC	CHANT	SHOAF
GAME	CONTEST	FOOTBALL	FOWN
CLASS	EVENT	LECTURE	PHRALT
GROUND	MATERIAL	DIRT	FAIVE

Table 1. (*continued*)

Prototype Prime Words	Hyponym to Hypernym Condition	Hypernym to Hyponym Condition	Nonwords
DRESS	CLOTHING	GOWN	CRIM
LETTER	TEXT	INVITATION	GROAD
NAME	TITLE	JOHN	FRUILL
FIGHT	CONFLICT	ENCOUNTER	HORT
FOOD	SUBSTANCE	BREAD	TWARGE
BED	FURNITURE	BUNK	GHULK
DOOR	BARRIER	SCREEN	GHINCE
MAN	MALE	BACHELOR	FULGE
SOUL	SPIRIT	GHOST	THRAFE
JOB	ACTIVITY	POSITION	BROSH
DINNER	MEAL	BANQUET	PREET
MACHINE	DEVICE	COMPUTER	TWUIT
APARTMENT	HOUSING	STUDIO	SPAUCE
BUSINESS	GROUP	FIRM	ZOM
STORY	FICTION	MYSTERY	TIBE
ACCIDENT	TROUBLE	CRASH	SHUIFF
DRINK	SERVING	COFFEE	PRISK
HOTEL	BUILDING	MOTEL	KEF
OFFICER	WORKER	CAPTAIN	CREANT
TRAIN	TRANSPORT	SUBWAY	FRUST

The words were arranged into three equal groups with each group consisting of 20 prime-hypernym pairs (the hyponym to hypernym condition; e.g., *team-unit*), 20 prime-hyponym pairs (the hypernym-hyponym condition; e.g., *team-crew*), 20 hypernym-hyponym unrelated pairs (the unrelated condition; e.g., *unit-Paris*), and 60 words selected from the remaining primes (n = 20), hypernyms (n = 20), and hyponyms (n = 20) that were paired with the 60 nonwords (the nonword condition; e.g., *unit-skerk*). As a result, each of the three groups contained the same divisions of conditions and the same words, but each condition in each group contained different words (except the nonword condition which always contained the same 60 nonwords). Thus, participants only saw each word once in the experiment. This arrangement also permitted an even division between target words and nonwords such that participants could not guess, based on frequency, if a target was more likely to be a word or a nonword.

All words were controlled for word frequency and word length. Word frequency counts came from the SUBTLEXus corpus developed by Brysbaert and New (2009). The SUBTLEXus frequency counts come from a corpus of 51 million words from subtitles of American television shows and films. As a result, the

Table 2. Descriptive Statistics for Word Selection

Index	Prototype primes	Hypernyms	Hyponyms
Word frequency	2.845 (0.415)	2.927 (0.475)	3.512 (0.206)
Word length	5.917 (1.619)	6.383 (1.814)	4.983 (1.501)

SUBTLEX corpus focuses on everyday interactions and more likely reflects the distributional properties of the natural language to which L1 and L2 speakers are exposed. The primary frequency count in SUBTLEX is contextual diversity (CD), which provides frequency counts from 8,388 different contexts and provides a means to calculate frequency without overestimating word form frequencies that result from multiple occurrences in a limited number of sources or even a single source (Adelman, Brown, & Quesada, 2006; Brysbaert & New, 2009). All words selected for this study were within the 9,000 most frequent words in the English language. Analysis of variance (ANOVA) analyses demonstrated that no significant differences in word frequency or word length existed between the selected hypernyms and hyponyms. Significant differences did exist between the word frequency scores for the prototype primes and the hyperyms ($p < .001$) and the prototype primes and hyponyms ($p < .001$). Significant differences were also reported for word length between the prototype primes ($p < .001$) and the hyperyms and the prototype primes and hyponyms ($p < .001$). Descriptive statistics for this analysis are reported in Table 2.

Apparatus and Procedure

Display of stimuli along with response time recordings were collected on a desktop computer using E-Prime software. Stimuli were displayed in uppercase letters in the center of the screen. The trials began with a center fixation point (+) followed by the presentation of the prime word for 250 milliseconds. The prime word was then replaced immediately by the target word, which was either a hypernym or hyponym of the prime word, an unrelated word, or a nonword. The target word remained displayed until the participants provided a response (i.e., was the target a word or a nonword). Responses were collected using a serial response box. Subjects received 10 practice trials followed by a single block of 120 trials, presented in random order. Demographic information was collected prior to the experiment using the E-Prime software. The entire experiment took about 20 minutes on average.

Data Analysis

A variety of statistical analyses were conducted to assess whether L1 speakers and advanced L2 learners showed priming effects for hypernyms and hyponyms in both response times (i.e., speed of judgments) and accuracy (i.e., number of correct judgments) when compared to the unrelated condition. Mixed design and within-subjects Analysis of Variance (ANOVA) tests were conducted to evaluate priming effects for speed. Logistic regression were used to evaluate the accuracy for L1 and L2 participants.

Results

Response Time Data

Descriptive analysis. A series of descriptive statistical analyses were conducted on the response time data to assess for outliers or errors. Descriptive statistics for the word raw response times for the L1 participants are presented in Table 2. Descriptive analyses demonstrated that the data for the hyponym-hypernym condition response times were normally distributed. However, the response time data for the hypernym-hyponym condition and the unrelated condition were not normally distributed. There was a single outlier in the hypernym-hyponym condition and two outliers in the unrelated condition. To correct the contaminated distribution and allow accurate interpretations through parametric statistics (Turkey, 1960), the data was first transformed using logarithmic formulas (logarithms to the base of 10; Field, 2005; Larson-Hall & Herrington, 2010). The logarithmic transformation corrected for one outlier in the unrelated condition. Not wanting to risk the independence of the data by removing the remaining outliers (Larson-Hall & Herrington), the scores for these outliers were transformed by assigning them a score that equaled the mean plus two standard deviations (Field).

Descriptive statistics for the word raw response times for the L2 participants are presented in Table 3. Descriptive analyses demonstrated that the response times for the unrelated condition were normally distributed. However, the response time data for the hyponym-hypernym condition, and the hypernym-hyponym condition were not normally distributed. There was a single outlier in the hyponym-hypernym condition and two outliers in the hypernym-hyponym condition. As with the L1 data, the data was first corrected for contaminated distribution through logarithmic formulas (logarithms to the base of 10; Field, 2005; Larson-Hall & Herrington-Hall, 2010). The logarithmic transformation corrected for all outliers in all the conditions.

Table 3. L1 Participant Response Times ($n = 43$)

	Mean	Standard Deviation	Minimum	Maximum
Hyponym-hypernym condition	745	132	533	1024
Hypernym-hyponym condition	762	156	482	1250
Unrelated condition	833	205	570	1496

Table 4. L2 Participant Response Times ($n = 41$)

	Mean	Standard Deviation	Minimum	Maximum
Hyponym-hypernym condition	994	308	516	1881
Hypernym-hyponym condition	1083	426	570	2588
Unrelated condition	1109	397	548	2044

Response Time Data

A mixed design ANOVA was conducted to assess overall effects of condition (hypernym-hyponym, hyponym-hypernym, and unrelated conditions) regardless of nativeness (i.e., L1 or L2 participants), differences in overall response times with respect to nativeness, and interactions between condition and nativeness. This ANOVA was followed by within subjects ANOVAs for L1 and L2 participants to examine differences in responses time between the conditions based on nativeness.

Mixed design ANOVA. There was a significant main effect for condition, $F(2, 164) = 12.477$, $p < .001$, $hp^2 = .132$. Tukey contrasts revealed that responses times for all participants were significantly faster for the hyponym-hypernym condition than for the unrelated condition, $F(1, 82) = 28.335$, $p < .001$, $hp^2 = .257$, and faster for the hypernym-hyponym condition than for the unrelated condition, $F(1, 82) = 6.918$, $p < .010$, $hp^2 = .078$. There was a significant effect for nativeness indicating that response times for L1 participants were faster than response times for L2 participants, $F(1, 82) = 24.632$, $p < .001$, $hp^2 = .231$. However, Levene's tests for equality of error for all conditions between L1 and L2 participants were significant, indicating response times between L1 and L2 participants were not equal in variance. There was not a significant interaction effect between the conditions and the nativeness of the participant, $F(2, 164) = 1.254$, $p > .050$, $hp^2 = .015$, indicating the response times between L1 and L2 participants were not significantly faster among the conditions.

Within-subjects ANOVA L1 participants. The within-subjects ANOVA of the response time data for the L1 participants showed a significant overall effect for condition, $F(2, 84) = 12.161$, $p < .001$ $hp^2 = .225$. Tukey contrasts revealed that responses times were significantly faster for the hyponym-hypernym condition than for the unrelated condition, $F(1, 42) = 20.980$, $p < .001$, $hp^2 = .333$, and faster for the hypernym-hyponym condition than for the unrelated condition, $F(1, 42) = 11.651$, $p < .001$, $hp^2 = .217$.

Within-subjects ANOVA L2 participants. The within-subjects ANOVA of the response time data for the L2 participants showed a significant overall effect for condition, $F(2, 80) = 4.418$, $p < .050$ $hp^2 = .099$. Tukey contrasts revealed that responses times were significantly faster for the hyponym-hypernym condition than for the unrelated condition, $F(1, 40) = 10.244$, $p < .010$, $hp^2 = .204$. However, response times were not significantly faster for the hypernym-hyponym condition when compared to the unrelated condition, $F(1, 40) = 0.618$, $p > .050$, $hp^2 = .015$.

Accuracy Data

Because a ceiling effect existed for the accuracy data (i.e., participants could not receive a score above 100 percent accuracy and thus there was a predefined level above which variance could not be measured), no distribution corrections were made. Thus, all data points were analyzed as raw scores (see Table 5). The ceiling effect also means that the data is not normally distributed and ANOVAs could not be conducted. Instead, logistic regressions were conducted to assess differences between conditions and between L1 and L2 participants.

All participants. Accuracy scores significantly predicted the hyponym-hypernym condition from the hypernym-hyponym condition, $\chi^2 = 7.556$, $df = 1$, N = 84, $p < .010$ and the hyponym-hypernym from the unrelated condition, $\chi^2 = 7.885$,

Table 5. Descriptive Statistics for Accuracy Rates

Condition	Participant	Mean	Standard Deviation
Hyponym-hypernym condition	NS	0.990	0.026
	NNS	0.988	0.024
	Combined	0.989	0.025
Hypernym-hyponym condition	NS	0.990	0.023
	NNS	0.961	0.041
	Combined	0.976	0.036
Unrelated condition	NS	0.984	0.028
	NNS	0.968	0.035
	Combined	0.976	0.032

df = 1, N = 84, $p < .010$. Accuracy scores did not significantly predict the hypernym-hyponym condition from the unrelated condition. This finding indicates that participants were more accurate in the hyponym-hypernym condition than in the other conditions.

L1 participants. Accuracy scores did not significantly predict any of the conditions indicating that participants were as accurate in all conditions.

L2 participants. Accuracy scores significantly predicted the hyponym-hypernym condition from the hypernym-hyponym condition, $\chi^2 = 12.494$, *df* = 1, N = 41, $p < .001$ and the hyponym-hypernym from the unrelated condition, $\chi^2 = 8.631$, *df* = 1, N = 41, $p < .010$. Accuracy scores did not significantly predict the hypernym-hyponym condition from the unrelated condition. This finding indicates that L2 participants were more accurate in the hyponym-hypernym condition than in the other conditions.

Discussion

This study has demonstrated that priming effects exist for both the hypernym-hyponym and the hyponym-hypernym conditions for L1 participants indicating that fluent lexical networks contain hierarchical clusters of related items that promote spreading activation between superordinate and subordinate terms and between subordinate and superordinate terms. However, for the L2 participants, the results indicate unidirectional priming effects suggesting that L2 lexicons are characterized by network connections that activate between subordinate and superordinate terms, but not between superordinate and subordinate terms. Such findings can provide important information about the organizational properties of L1 and L2 lexicons and the differences between fluent and developing lexicons.

For L1 speakers, the results provide additional evidence for the spreading activation model and the strength of hierarchical relations in the organization of the mental lexicon. In fluent lexicons, hyponyms activate hypernyms (i.e., the prime word *car* stimulates the target word *vehicle)* and hypernyms activate hyponyms (i.e., the prime word *car* stimulates the target word *ambulance*). This finding provides evidence that fluent lexicons can be characterized by organizational principles that allow for set inclusion among category members. These organizational features likely allow users to generalize terms because the terms are part of a conceptual category and are not isolated from one another. Such organizational principles provide for economic lexical representations and the subsequent priming effects reported by the response times in this study.

For advanced L2 speakers, the results suggest that lexical networks do not develop to the same degree as found in L1 participants and call into question the strength of L2 lexical networks and the use of spreading activation models to explain lexical associations between hierarchically related words in L2 mental lexicons. The results indicate that priming effects for L2 speakers in this study are not only unidirectional (only between a hyponym as a prime and a hypernym as a target), but, more importantly, the targets that shared fewer conceptual cues with the prime word, and were thus less similar, demonstrated activation. A spreading activation model would predict that concepts that were more similar would be stored more closely together and thus have more rapid activation. However, the findings of this study indicate that only hyponyms prime hypernyms, which contain fewer shared items with the prime. Conversely, hypernyms do not demonstrate priming effects for hyponyms even though hyponyms should share a greater number of cues with the hypernym prime and would thus be more similar (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).[1] An example of this relationship is contained in the prime word *car*, which acts as a hyponym of *vehicle* and a hypernym of *ambulance* (see Figure 1). *Ambulance*, as a hyponym of car contains more shared items with *car* than *vehicle* does (e.g., four wheels, doors, a steering wheel, a windshield, seats, an internal combustion engine). *Vehicle* is less similar than a car and may or may not contain reliable cues (e.g., not all vehicles have wheels, doors, steering wheels, windshields, seats, and internal combustion engines). However, even though *vehicle* is less similar to *ambulance*, the results from the L2 participants support the notion that the word *car* will prime *vehicle*, whereas it will not prime *ambulance*.

That a spreading activation model does not explain L2 priming effects for hypernymic relations may be a result of dimensionality. For each hyponym, there is only one hypernym; however, for each hypernym there may hundreds of hyponyms (i.e., greater dimensionality). As a result, there may be a dispersion effect for hyponyms in that activation is spread across a greater number of items, thus increasing processing time and reducing the priming effect. No such dispersion effect would exist for hypernyms and, as a result, a stronger priming effect would occur. The results reported in this study may therefore be the result of the number of related conceptual links and not the distance between these concepts.

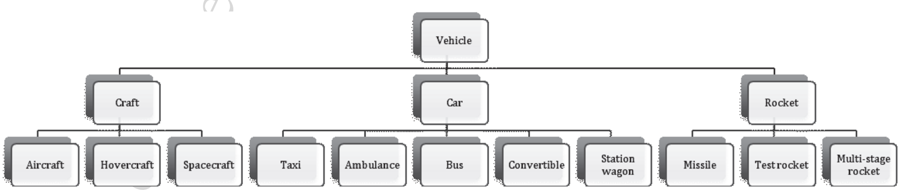

Figure 1. Immediate hierarchical relationships for the word car.

Alternatively, the reported findings may be the result of markedness in that a hyponym is semantically more complex than its hypernym because it contains more semantic features. As a result the hyponym also has less potential as a referent for its hypernym (Andersen, 2001; Jacobson, 1983; Kearns, 2006). Since marked forms are argued to be more difficult to learn, it is possible that L2 speakers have lower fluency and accuracy with hyponyms as compared to hypernyms, affecting the speed of word recognition. Such a position is supported by the accuracy data reported in this study, which indicated L2 participants were more accurate in the hyponym-hypernym condition than in the other conditions. Such a position is also supported by previous studies that showed that L2 learners produce words that are less specific as a function of time spent learning English (Crossley et al., 2009) and that L2 learners use more general words than specific words (Levenston & Blum, 1977).

Overall, this study provides some evidence to support the notion that hierarchical lexical networks are not fully developed in advanced L2 speakers to the same degree as found in L1 speakers. This is in contrast to past L2 studies investigating other aspects of lexical knowledge (i.e., collocations, antonyms, and synonyms; Dong et al., 2005; Frenck-Mestre & Prince, 1997), which have demonstrated priming effects that are similar to those found for L1 speakers.

Conclusion

This study has provided additional evidence for the existence of bi-directional hierarchical lexical networks in L1 lexicons. However, for advanced L2 speakers, such evidence was lacking suggesting that advanced L2 speakers may not access and store hierarchical lexical information in a similar manner to L1 speakers.

In general, the experimental designs used in this study provide confidence in the reported findings. However, the L2 RTs and the standard deviations for these RTs were greater than expected. Such variability in the L2 RTs, as compared to the L1 RTs means that statistical power between the two analyses was not equal and that a greater L2 sample size may be needed to provide appropriate power in order to generalize the findings across the population. Thus, the L2 results should be viewed with caution. Also, only subject analyses and not item analyses were not conducted for the data. This should also be considered a limitation. Demographically, this study used a generalized definition of L2 learner and did not control for age of learning a second language or for the amount of time spent in an environment wherein English was the dominant language. Future studies may want to control for age of exposure and learning environment, both of which may play a role in the development of lexical networks. Additionally, this study used a within language design in which the prime and the target words were in English. It is possible that priming effects may exist between languages (i.e., a word in the L1 may prime a word in the L2) and not within languages.

Acknowledgments

The author would like to thank Christie Collins, Laura Varner, and Ashely Titak for their help in collecting participant data. The author is also indebted to Danielle McNamara, YouJin Kim and Kim McDonough for their assistance with the experimental design. Lastly, the author thanks the anonymous reviewers and the editors at the Mental Lexicon for their assistance in developing this paper.

Note

1. Alternative theories support the notion that categories contain only information that is common across all members (e.g., Collins & Quillian, 1969). In the examples provided in Figure 1, craft, car, and rocket would thus share *has frame* and missile, test rocket, and multi-stage rocket would share *has jet engine,* but not *has frame*.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, *17*, 814–823.

Andersen, H. (2001). Markedness and the theory of change. In H. Andersen (Ed.), *Actualization: Linguistic change in progress* (pp. 21–58). Amsterdam: John Benjamins.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, *58*(10), 1–166.

Berlin, B., Breedlove, D., & Raven, P. (1974). *Principles of Tzeltal plant classification*. New York: Academic.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*, 14–21.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

Chaffin, R., & Glass, A. (1990). A comparison of hyponym and synonym decisions. *Journal of Psycholinguistic Research*, *19*(4), 265–280.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, *8*, 240–248.

Crossley, S. A., Salsbury, T., & McNamara, D. S. (2009). Measuring second language lexical growth using hypernymic relationships. *Language Learning*, *59*(2), 307–334.
Crossley, S. A., Salsbury, T., & McNamara, D. S. (2010a). The development of polysemy and frequency use in English second language speakers. *Language Learning*, *60*(3), 573–605.
Crossley, S. A., Salsbury, T., & McNamara, D. S. (2010b). The role of lexical cohesive devices in triggering negotiations for meaning. *Issues in Applied Linguistics*, *18*(1), 55–80.
Daller, H., van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, *24*(2), 197–222.
Devitto, Z., & Burgess, C. (2004). Theoretical and methodological implications of language experience and vocabulary skill: Priming of strongly and weakly associated words. *Brain and Cognition*, *55*, 295–299.
Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, *8*, 221–238.
Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution and function. *Modern Language Journal*, *93*, 370–385.
Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, *16*, 409–435.
Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and L2 vocabulary acquisition. *Language Learning*, *44*, 449–491.
Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
Field, A. (2005). *Discovering statistics using SPSS*. London: Sage Publications.
Frenck-Mestre, C., & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language*, *37*, 487–501.
de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition*, *24*, 81–112.
Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, *10*(2), 221–240.
Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430.
Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *21*(2), 181–193.
Ijaz, I. H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, *36*(4), 401–451.
Jacobson, S. (1983). Three types of terminologies. In R. R. Hartmann (Ed.), *Proceedings of the European Association for Lexicography* (pp. 355–366). Tubingen: Max Niemeyer Verlag.
Kearns, K. (2006). Lexical semantics. In B. Aarts & A. McMahon (Eds.), *The handbook of English Linguistics* (pp. 557–580). Malden, MA: Blackwell Publishing.
Larson-Hall, J., & Herrington, R. (2010). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, *31*(3), 368–390.
Levenston, E., & Blum, S. (1977). Aspects of lexical simplification in the speech and writing of advanced adult learners. In P. S. Corder & E. Roulet (Eds.), *The notions of simplification, interlanguages and pidgins and their relation to second language pedagogy* (pp. 51–72). Geneve: Librairie Droz.
LeVine, R. A. (1980). Influences of women's schooling on maternal behavior in the third world. *Comparative Education Review*, *24*, 78–105.

McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. New York: Routledge.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 558–572.

Meara, P. (1982). Word association in a foreign language: A report on the Birkbeck vocabulary project. *Nottingham Linguistic Circular*, *11*, 20–37.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge: Cambridge University Press.

Meara, P. (2005). Designing vocabulary tests for English, Spanish and other languages. In C. Butler, S. Christopher, M. Á. Gómez González, & S. M. Doval-Suárez (Eds.), *The dynamics of language use* (pp. 271–285). Amsterdam: John Benjamins.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate level categories. *Child Development*, *53*, 258–266.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). *Five papers on WordNet*. Princeton University: Cognitive Science Laboratory.

Miller, G. A., & Teibel, D. A. (1991). A proposal for lexical disambiguation. In *Human Language Technology Conference: Proceedings of the Workshop on Speech and Natural Language* (pp. 395–399). Pacific Grove, California: Association for Computational Linguistics.

Murphy, G. L. (2004). *The big book of concepts*. Cambridge, MA: MIT Press.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillside, New Jersey: Erlbaum.

Ordonez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer. *Journal of Educational Psychology*, *94*(4), 719–728.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*, 180–194. DOI 10.1007/s00426-002-0086-5.

Phillips, N., Segalowitz, N., O'Brien, I., & Yamasaki, N. (2004). Semantic priming in a first and second language: Evidence from reaction time variability and event-related time potentials. *Journal of Neurolinguistics*, *17*, 237–262.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37–42). Hillsdale, NJ: Erlbaum.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*, 282–308.

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, *21*(1), 28–52.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology*, *55A*, 1339–1362.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408.

Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Lawrence Erlbaum.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 573–605.

Schmitt, N. (1998). Tracking the incremental acquisition of a second language vocabulary: A longitudinal study. *Language Learning*, *48*(2), 281–317.

Snow, C. E. (1988). The problem with bilingual education research critiques: A response to Rossell. *Equity and Excellence*, *23*(4), 30–31.

Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language*, *17*, 697–710.

Snow, C. E., Cancino, H., Gonzalez, P., & Shriberg, E. (1989). Giving formal definitions: An oral language correlate of school literacy. In D. Bloome (Ed.), *Classrooms and literacy* (pp. 233–249). Norwood, NJ: Ablex.

Snow, C. E., Cancino, H., De Temple, J., & Schley, S. (1991). Giving formal definitions: Linguistic or metalinguistic skill? In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 90–112). Cambridge, England: Cambridge University Press.

Söderman, T. (1993). Word associations of foreign language learners and native speakers: The phenomenon of a shift in response type and its relevance for lexical development. In H. Ringbom (Ed.), *Near-native proficiency in English* (pp. 91–182). Abo, Finland: Abo Akademi.

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In I. Olkin, S. G. Ghwyne, W. Hoeffding, W. G. Madow, & H. B. Mann (Eds.), *Contributions to probability and statistics: Essays in honour of Harold Hotelling* (pp. 448–485). Stanford: Stanford University Press.

Van-Patten, B., & Williams, J. (2007). *Theories in second language acquisition: An introduction*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Vergès, P. (1992). L'évocation de l'argent: une méthode pour la définition du noyau central de la representation [The evocation of money: A method for defining the central core of a representation]. Bulletin de Psychologie [*Bulletin of Psychology*], *45*, 203–209.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, *53*, 13–40.

Wolter, B. (2001). Comparing the L1 and L2 mental lexicons: A depth of individual word knowledge model. *Studies in Second Language Acquisition*, *23*, 41–70.

Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & LeBlanc, C. (2009). The acquisition of tense-aspect: Converging evidence from corpora, cognition, and learner constructions. *Modern Language Journal*, *93*, 354–369.

Corresponding Address

Scott A. Crossley
Department of Applied Linguistic/ESL
Georgia State University
34 Peachtree St. Suite 1200
One Park Tower Building
Atlanta, GA 30303

Tel.: (404) 413-5179
E-mail: sacrossley@gmail.com