

Interdisciplinary approaches to researching L2 lexical acquisition, processing, and use: An introduction to the Special Issue

Journal:	Second Language Research
Manuscript ID	Draft
Manuscript Type:	Special Issue
Keywords:	interdisciplinary, lexical studies, lexical processing, second language vocabulary, second language acquisition of vocabulary, aspects of lexical knowledge, interdisciplinary research methods
Abstract:	Lexical knowledge is complex, multidimensional, and difficult to pin down to a set of defined components. The development, organisation, and use of lexical knowledge in the first and additional languages are studied in a number of neighbouring disciplines beyond second language acquisition and applied linguistics, including psycholinguistics, neurolinguistics, computational linguistics, and language education. In this introduction, we highlight how the five articles in this Special Issue hone our understanding of different aspects of L2 lexical knowledge, its acquisition, and use by adopting innovative research design, methods, and approaches to data collection and analysis from these distinct but related disciplines, affording new theoretical and empirical insights.

SCHOLARONE™ Manuscripts

Interdisciplinary approaches to researching L2 lexical acquisition, processing, and use: An introduction to the Special Issue

Abstract

Lexical knowledge is complex, multidimensional, and difficult to pin down to a set of defined components. The development, organisation, and use of lexical knowledge in the first and additional languages are studied in a number of neighbouring disciplines beyond second language acquisition and applied linguistics, including psycholinguistics, neurolinguistics, computational linguistics, and language education. In this introduction, we highlight how the five articles in this Special Issue hone our understanding of different aspects of L2 lexical knowledge, its acquisition, and use by adopting innovative research design, methods, and approaches to data collection and analysis from these distinct but related disciplines, affording new theoretical and empirical insights.

1. Introduction to the Special Issue

Lexical knowledge is foundational in the first (L1) and second/additional (L2) language comprehension and production. When we read a book, watch a film, express an opinion or chat with friends from another country, words and phrases are the basic building blocks of communication. Yet, second language acquisition (SLA) research has traditionally focused on the learning of grammar and syntax, as uniquely *linguistic* knowledge components that arguably develop differently from general knowledge, to which word knowledge was assigned. Encouragingly, in recent years, research into aspects of lexical acquisition, processing, and use is starting to become a regular feature in all key SLA journals.

This Special Issue profiles studies that are representative of major trends in today's L2 lexical processing and acquisition research from an interdisciplinary perspective. Three of the studies published in this issue focus on questions related to L2 vocabulary acquisition, one study investigates bilingual lexical processing, and one study compares L1 and L2 speaker intuitions about lexical chunks. One of the L2 acquisition studies (Monaghan, Ruiz, & Rebuschat, 2020) is concerned with simultaneous acquisition of vocabulary and morphosyntax during exposure to an artificial language, exploring the effect of local and global variables on learning. The second L2 vocabulary acquisition study compares the effect of massed and spaced repetition schedules on L2 (English) vocabulary learning from reading by Japanese English as a Foreign Language (EFL) students, testing the acquisition of explicit and tacit knowledge of meaning (Nakata & Elgort, 2020). The third L2 vocabulary acquisition mega-study contributes to our knowledge about L2 (English) word acquisition order using empirical data obtained through crowdsourcing (Brysbaert, Keuleers, & Mandera, 2020). The L2 processing study by Chen, Perfetti, Fang, and Chang (2020) tests the locus of L1 activation (i.e., lexical vs. pre-lexical) in L2 word reading by Chinese-English bilinguals, using implicit and explicit tasks. Finally, L1 and L2 knowledge of Italian word combinations (collocations and idioms) is investigated by Fioravanti, Senaldi, Lenci, and Siyanova-Chanturia (2020) by comparing L1 and L2 speakers' ratings of lexical fixedness and compositionality. It is noteworthy that this contribution looks at an under-researched L2 -Italian. The majority of research in the field of SLA and, in particular, in vocabulary studies has focused on L2 English. Having a study on a non-English L2, which has to date received very little attention, is a welcome addition to our current body of knowledge about L2 lexical acquisition and use.

These five studies challenge and extend existing knowledge of how L2 words and phrases are acquired, stored, and processed. Building on earlier vocabulary research that underscores the complex, multifaceted nature of lexical knowledge (Nation, 2013; Perfetti, 2007; Read, 2000), the articles in this Special Issue heed growing calls to use sensitive measures and interdisciplinary approaches in studying different facets of L2 vocabulary knowledge and use (Godfroid, 2020). In a recent introduction to a special issue of *Language Learning*, Rebuschat, Meurers, and McEnery (2017) argued that the study of "a complex phenomenon like language acquisition can significantly benefit from insights, tools, and methods from many disciplines, yet it is still relatively rare to find studies that combine multiple approaches" (p. 7). In L2 vocabulary studies, Read warns researchers against oversimplification: "[w]hether we focus on individual lexical items or the mental lexicon as a whole, we are setting out to describe something that is inherently ill-defined, interdimensional, variable and thus resistant to neat classification" (Read, 2004, p. 224).

In order to match the subject matter complexity and piece together a more precise picture of L2 lexical acquisition and processing, it behoves SLA vocabulary investigations to adopt research methods from adjacent disciplines that study memory and language, including psycholinguistics, cognitive psychology, and education. Such interdisciplinary approaches facilitate the creation of a new, deeper understanding of what it means to acquire L2 lexical knowledge and how this knowledge is stored and accessed, establishing more precise theoretical accounts of acquisition, learning, processing, and use. Novel methods of data collection and analysis, combined with researchers adopting open science practices of sharing data and instruments, increase the robustness and transparency of the findings, make them more accessible for replications and suitable for meta-analyses, and bring L2 vocabulary research in line with current science research standards (Brysbaert &

Stevens, 2018; Lindstromberg, 2016; Lindstromberg & Eyckmans, 2017; Marsden, Morgan-Short, Thompson, & Abugaber, 2018; Marsden & Plonsky, 2018; Porte, 2012; Rebuschat, Meurers, & McEnery, 2017). Below we elaborate on the theoretical, empirical, and methodological contributions of the five articles included in this Special Issue and detail new insights into different aspects of L2 lexical acquisition and processing offered by these interdisciplinary and methodologically innovative studies.

2. Facets of lexical knowledge

Vocabulary knowledge comprises a number of aspects or dimension that have been described, categorised, and studied in overlapping and distinct ways in different fields. A word's identity has been described as a nexus of three groups of component representations (Perfetti & Hart, 2001, 2002): linguistic form (phonology and morphosyntax), literacy form (orthography), and meaning (lexical semantic representations and the knowledge of core meaning/s). One of the most comprehensive taxonomies of lexical knowledge, proposed by Nation (1990, 2001, 2013), comprises receptive and productive knowledge of three core aspects, each consisting of three constituent components: form (spoken, written, word parts), meaning (form-meaning mapping, conceptual and referential meaning, associations), and use (grammatical, collocational, pragmatic constraints).

The articles in the Special Issue investigate a number of aspects of lexical knowledge, from the basic L2 word form recognition (Brysbaert et al., 2020), to form-meaning mapping and lexical-semantic representations (Nakata & Elgort, 2020), to the knowledge of use, including grammatical function (Monaghan et al., 2020) and collocational constraints (Fioravanti et al., 2020); as well as cross-language access to different aspects of word

knowledge (i.e., sub-lexical orthographic and lexical semantic representations) in bilingual word processing (Chen et al., 2020).

2.1 Which words do L2 speakers know?

The Special Issue includes an article by Brysbaert et al. (2020) that uses a *Yes-No* test to estimate how many and which L2 words are recognised by speakers of English as an additional language from diverse L1 backgrounds. Vocabulary size (i.e., the number of known words in a person's lexicon) is an important indicator of L2 proficiency. Brysbaert et al. adopted a bottom-up approach to estimating L2 vocabulary size based on a large dataset of responses to individual words. Unlike studies interested primarily in estimating how *many* words a language learner knows, Brysbaert et al. shift the focus to *what* English words are known by L2 users. This approach affords inferences about the order of L2 word acquisition. Having obtained more than 17 million responses in an online *Yes-No* test from L2 speakers of English, the authors create a new word ranking that takes into account the number of L2 speakers who know a word and the time it took them to respond, in addition to the more traditional top-down criterion, i.e., corpus word frequency.

2.2 Form-meaning mapping and access to L2 lexical-semantic knowledge

Beyond the basic ability to recognize a word in the target language, quality of knowledge
and its integration into the existing L2 lexical semantic networks of the learner is critical for
fluent L2 comprehension. A contextual L2 word learning study by Nakata and Elgort (2020)
explored whether theoretical predictions of memory research can explain the effect of
spaced versus massed repetition on the development of L2 word knowledge from reading.
Going beyond controlled offline retrieval of form and meaning, Nakata and Elgort compared

the effect of different repetition schedules on the integration of new and old L2 lexical semantic knowledge. This study compared the development of multiple aspects of L2 word knowledge, testing the hypothesised dissociation between the development of explicit and tacit lexical knowledge.

- 2.3 Learning the word meanings and grammatical function from cross-situational exposure

 Monaghan et al. (2020) used an artificial language to probe how adult learners acquire
 language inductively, focusing on both vocabulary and syntax and drawing parallels
 between first and second language learning. The knowledge of grammatical functions is a
 component aspect of vocabulary knowledge (Nation, 2001, 2013); in fact, some linguistic
 theories and approaches to language acquisition have questioned the formal separation of
 the lexicon and grammar, instead, emphasising the role of statistical learning (e.g., Bod,
 2006). Monaghan et al. tested learners' ability to extract the meaning and the grammar of
 novel linguistic forms while exposed to complex utterances and visual scenes under
 different instructional and feedback conditions.
- 2.4 Collocational constraints fixedness and compositionality of word combinations

 On a par with words and highly idiosyncratic phrases, such as idioms, the mental lexicon also contains thousands of word sequences that vary along the continua of frequency of use and lexical properties, such as compositionality and fixedness. The contribution by

 Fioravanti and colleagues focused on the knowledge of word use, represented by L1 and L2 speakers' sensitivity to sequences above the word level. Employing acceptability ratings

 (Study 1) and forced choice decisions (Study 2), Fioravanti et al. (2020) investigated L1 and L2 speakers' perceptions of two phrasal properties lexical fixedness and compositionality –

focusing, in particular, on combinatorial constraints of Italian Verb+Noun word combinations.

2.5 The locus of cross-language activation

Word reading involves visual, orthographic word processing and activation of the lexical and semantic representations. Reading a word in a second language may also involve activation of the corresponding word in the first language, a phenomenon known as 'non-selective lexical access' (e.g., Dijkstra & Van Heuven, 2002; Wen & Van Heuven, 2018). Although evidence of non-selective access has been mainly obtained from pairs of orthographically close languages (such as, Dutch and English), some studies were able to show L2-L1 interaction for distant languages, with different writing systems, such as Hebrew and English (e.g., Degani, Prior, & Hajajra, 2018), Japanese and English (e.g., Miwa, Dijkstra, & Bolger, 2014) and Chinese and English (e.g., Thierry & Wu, 2007; Zhang, Van Heuven, & Conklin, 2011). L1 co-activation in L2 processing may occur at the sub-lexical level (i.e., at the letter or syllable level in alphabetic languages, and at the level of phonetic and semantic radical in logographic languages), or at a later lexical and semantic processing stages (i.e., at the whole word level in alphabetic languages and at the character level in logographic languages). Although there is evidence of both lexical and sub-lexical within-language activation (e.g., in Chinese, between alphabetic Pinyin and morphosyllabic Chinese characters), and across alphabetic languages (e.g., Bijeljac-Babic, Biardeau, & Grainger, 1997; Brysbaert, Van Dyck, & Van de Poel, 1999; Van Heuven, Dijkstra, & Grainger, 1998), whether both sub-lexical and lexical activation occur in the processing of English words by Chinese-English bilinguals is not yet clear. The findings reported by Chen et al. (2020) supply an important piece of the cross-language lexical processing puzzle.

3. Innovative interdisciplinary research methods

Rich insights into L2 lexical processing, knowledge and its development can be gained by enacting a research agenda that builds on advances in theoretical, computational, and experimental methods from SLA and the neighbouring disciplines that study language and memory, particularly cognitive psychology, psycholinguistics, and computational linguistics. To develop such an agenda, SLA researchers need to engage in a cross-disciplinary dialogue and adopt and adapt interdisciplinary research methods, such as behavioural response time paradigms, including priming, eye-tracking, event-related brain potentials, distributional semantics techniques, computational modelling, and others. The contributions that make up this special issue are a testament to how such innovative research methods can inform and help move the field forward. Below we highlight some of these insights into the acquisition, processing, and use of different facets of vocabulary knowledge afforded by the use of such methods.

3.1 Innovative approaches to the data collection

The mega-study by Brysbaert et al. (2020) offers a new approach to estimating which words L2 learners know and how L2 vocabulary knowledge develops. Their crowdsourcing webbased approach to the data collection resulted in millions of data points from L2 learners representing diverse L1 backgrounds, who took the online test of English word knowledge. This method of data collection (Brysbaert, Stevens, Mandera, & Keuleers, 2016; Keuleers, Stevens, Mandera, & Brysbaert, 2015) establishes a new landmark in L2 vocabulary size research because it adopts a bottom-up approach to estimating vocabulary knowledge,

starting with empirical evidence collected from thousands of L2 participants and theorizing on the basis of this evidence.

Traditionally, it has been assumed that L2 vocabulary acquisition is, by and large, driven by the frequency of words' occurrence in the target language, calculated using large corpora – collections of spoken and written language. Ranked frequency wordlists (usually, 1,000 lexical items of a similar corpus frequency) are created and vocabulary size tests are constructed by drawing a stratified sample from these lists in such a way that each item in the test represents a specified proportion of the words in the corresponding wordlist (e.g., 10 words representing 1% of a 1,000-item wordlist). Higher-frequency words are expected to be learned earlier and known better than lower frequency words (which is what the test data normally shows). However, this approach is not without its challenges (as detailed in Brysbaert et al., 2020).

The distribution of word frequencies is extremely asymmetrical; the first two to three thousand most frequent words are extremely frequent but the frequency drops quickly after that, with a very long tail of many thousands of low frequency words. The relative frequency of these lower frequency words varies depending on the corpus chosen for the development of the wordlists (e.g., BNC and COCA prioritise written texts and SUBTLEX is based on the language of film subtitles). Thus, although the rank order of the first couple of thousand words can be estimated more or less reliably, it is more variable for low frequency words. This makes it more challenging to accurately measure learners' vocabulary size, especially for more advanced learners, or make predictions about the order in which L2 words are acquired. Another limitation of this top-down approach is that it assumes L2 vocabulary development could be almost entirely explained by the frequency of

use in the target language. Yet, the input learners get in the foreign language classroom and from textbooks is not always aligned with natural language use.

A large dataset of word knowledge responses obtained by Brysbaert et al. shows that a bottom-up approach can reveal discrepancies between expected or theorised word knowledge and real knowledge of L2 learners. Thus, the crowdsourcing methodology offers a novel approach to verifying hypotheses about the relationship between corpus word frequencies and the actual development of L2 vocabulary knowledge. Combining L2 ranks obtained on the basis of word prevalence and word recognition times, Brysbaert et al. (2020) created new wordlists that represent the order of L2 word acquisition in English.

3.2 Using multi-experiment task design from bilingual processing research

Chen et al. (2020) used a number of behavioural response time tasks (explicit and implicit) to determine the locus of involvement of a logographic L1 (Chinese) in the processing of an alphabetic L2 (English). In the initial task, participants first read an English word silently and then made a colour judgement (red or blue) on the target Chinese character related (or unrelated) to the English word. This relationship was manipulated in a way that allowed the researchers to determine whether sub-lexical (sub-character) Chinese orthography was automatically activated when Chinese speakers silently read English words, or whether the activation occurred at the lexical (meaning-related character) level. As the first task only showed lexical but not sub-lexical level of L1 activation in L2 processing, the researchers attempted to direct participants' explicit attention to the meaning by instructing them to make semantic judgements on the English word, instead of reading it silently, while keeping the colour decision task on the Chinese character the same. This manipulation did not change the outcomes, as the L1 activation was again observed at the lexical but not sub-

lexical level. In both experiments, however, lexical access to the Chinese character occurred through implicit processing, because the task was to judge the colour. Because L2 processing may not be robust enough to be reliably detected in an implicit processing task, Chen et al. (2020) conducted an additional experiment, replacing the colour judgement with the naming task that involves explicit lexical processing. Still, only lexical character-level activation of L1 was observed in this task, but sub-lexical (radical) level activation did not occur. This methodical, step-by-step, careful manipulation of the task characteristics to test the locus of L1 activation in L2 processing is a characteristic of experimental psycholinguistic approaches to hypothesis testing. The researchers used this multi-experiment methodology to answer an important theoretical question about the nature of cross-language influences in L2 word processing.

3.3 Using multiple measures to understand word learning as a process and an outcome

In the study that investigated the effect of the repetition schedule (spaced versus massed)
in contextual L2 word learning, Nakata and Elgort (2020) went beyond measuring the
outcome (explicit and tacit) word knowledge resulting from these two treatments, also
tracking the development of the knowledge of word meanings during reading. This careful
experimental design enabled the researchers to piece together a detailed picture of
contextual word learning, as a process, revealing that a single learning episode in the
massed condition (three contextual encounters prior to receiving feedback) led to a more
accurate word meaning inference than the first two contextual encounters, followed by
feedback, in the spaced condition. It was only the final (third) learning episode in the spaced
learning condition that resulted in a more accurate inference than in the massed condition.

Importantly, using mixed effects regressions, Nakata and Elgort were also able to test whether the accuracy of meaning inferences during the learning phase differentially affected explicit and tacit knowledge of meaning, gained under the massed and spaced learning conditions. They found that the meaning inference accuracy was a reliable predictor of the outcome explicit word knowledge, but only when the words were learned in the spaced, but not massed, condition. No effect of inference accuracy was observed on the tacit knowledge (operationalised as a semantic priming effect). By administering both traditional offline tests of explicit knowledge (form-meaning and meaning-form mapping) and psycholinguistic online tests of tacit knowledge (semantic priming) and using mixed effects statistical models, Nakata and Elgort were able to simultaneously ascertain the effect of the repetition schedule and contextual inference accuracy on the establishment of explicit and tacit word knowledge in contextual word learning. Their interdisciplinary approach to the study design, testing and measurement procedures, and statistical modelling of the data allowed the researchers to match the complexity associated with investigating multiple dimensions of vocabulary knowledge and its development under different learning conditions.

3.4 Statistical modelling of the learning data to account for global and local variables

The article by Monaghan, et al. (2020) is an example of the tangible research advantages

afforded by sophisticated and flexible approaches to the data analysis. Using mixed-effects

regressions, Monaghan et al. were able to zoom in on the interplay between global factors

(i.e., explicit versus implicit instruction and the provision of feedback) and local factors (i.e.,

accuracy of responses on the preceding learning trial) in the learning of an artificial

language. The logistic mixed-effect data analysis (Jaeger, 2008) allowed the researchers to

model accuracy (correct vs. incorrect response) for each participant on each item during the training task, accounting for the variation among individual participants and items in the same model. Their approach to data modelling shows how the learning of an artificial language happens over time, from one situational encounter to the next. By testing both linear and quadratic effect in the modelling of the trial-by-trial learning data, the researchers created a dynamic view of learning; they showed how global factors affect the language learning trajectory and which word types and syntactic positions are likely to affect and be affected by prior learning outcomes. The study further highlighted affordances associated with a laboratory-based paradigm (with learning measures trial-to-trail, establishing a precise learning trajectory), compared with the more traditional classroom-based learning settings wherein such detailed and accurate measurements may not be possible.

3.5 Combining computational corpus-based approaches and human ratings in evaluating compositionality of word combinations

Fioravanti, et al. (2020) adopted a multi-study approach in their investigation of lexical fixedness and compositionality of Verb+Noun word combinations – free combinations, collocations, and idioms – in L1 and L2 Italian. Modifying target sequences and contexts, and employing two off-line tasks – *acceptability ratings* and *lexical choice* – the authors explored L1 and L2 speakers' intuitions about the extent to which free combinations, collocations, and idioms may be deemed acceptable. The authors also used computational indices of compositionality to investigate lexical flexibility across the three types of combinations. Although distributional semantic indices of compositionality have previously been shown to predict phrase similarity in human judgments of acceptability and syntactic flexibility for

idioms, this is the first study to have triangulated L1 and L2 data with such computational indices, providing new insights into how L1 speakers and L2 learners perceive free combinations, collocations, and idioms – word combinations that vary markedly along the continua of lexical fixedness and compositionality. The paper shows that considering human judgements and computational indices together is a powerful method for investigating lexical properties of different types of word combinations.

4. Theorising L2 lexical acquisition and processing

The research questions considered in the Special Issue on L2 lexical acquisition and processing are not very different from those posed in canonical SLA research. The topics include the order of L2 acquisition, incidental versus intentional learning, and the relationship between explicit and implicit knowledge, learning and instruction, and the organisation and development of L2 knowledge, including questions about the L2–L1 interface. The contributions to the Special Issue consider these questions in reference to the acquisition and processing of L2 *lexical* knowledge.

Brysbaert et al. (2020) test the assumption that the order of L2 vocabulary acquisition can be explained primarily in reference to corpus word frequency. Monaghan, et al. (2020) examine how explicit and implicit instruction and feedback affect cross-situational acquisition of vocabulary and morphosyntax of an artificial language. Nakata and Elgort (2020) test the effect of the repetition schedule on the development of explicit and tacit L2 word knowledge during and from reading by Japanese learners of English. Fioravanti et al. (2020) investigate language users' intuitions about fixedness and compositionality of Italian word combinations, testing the phraseological approach to formulaicity. Research by Chen et al. (2020) identifies the locus of L1 (Chinese) activation during L2 (English) word

processing, contributing to theories and models of bilingual processing for distant L1-L2 pairs of languages. Below we expand on theoretical contributions of the article in this issue.

4.1 Beyond word frequency

Brysbaert et al.'s (2020) study showed that word frequency based on existing English corpus data accounts for 46% of the variance in the likelihood that a word will be familiar to an L2 speaker. Clearly, word frequency in the language, in general, is an important predictor of whether a word is likely to be known, but there are a number of other factors that need to be considered. Firstly, the finding reported by Brysbaert and colleagues confirm the assumption that authentic texts and speech may not be the primary source of the target language input in foreign language learning contexts. In such contexts, words related to the object of study – the language itself (such as verb, subject, vocabulary) – may be used more frequently than in the target language more generally. On the other hand, words associated with early childhood (such as doll, dad, smallpox, parakeet) that most native speakers know are less likely to be known by language learners whose age of L2 acquisition coincides with the start of formal schooling or takes place even later in life. Secondly, Brysbaert et al. argue that not all words are created equal, i.e., some words are easier to learn that others. For example, words that are similar in spelling, pronunciation, and meaning in the L1 and L2 are easier to learn and are more likely to be known than other words. Finally, similar to He and Godfroid (2019) and Hashimoto and Egbert (2019), Brysbaert et al.'s conjecture that learner motivations and perceptions of usefulness play a part in what L2 words are known. Having used a bottom-up approach to the question of what words L2 learner of English know, empirical evidence from Brysbaert et al. suggests that the order and likelihood of word acquisition in the L2 differs somewhat from that in the L1, as a result of factors beyond

corpus word frequency, i.e., L2 word difficulty, frequency of encounters in the learning context and learner motivations. Taking these additional factors into consideration, an important practical contribution of research by Brysbaert et al. is the developed new L2 ranked 1,000 word-family wordlists, adjusted for L2 word prevalence and recognition times.

4.2 Cross-situational acquisition of vocabulary and morphosyntax

To understand how explicit and implicit instruction and feedback affect cross-situational acquisition of vocabulary and morphosyntax, Monaghan, et al. (2020) trained participants to acquire a complex artificial language from utterances and scenes that contained objects, properties of object, actions, grammatical role markers, and grammar (word order). Monaghan et al. tested hypotheses generated from the holistic view of SLA that predicts simultaneous and interconnected cross-situational learning of the target grammar and syntactic structures as well as vocabulary, offering new evidence on how learning and instructional approaches differentially affect components of linguistic knowledge. This holistic research paradigm removes the boundaries between research into acquisition of vocabulary and acquisition of morphosyntax, capable of revealing similarities and differences in their learning. Thus, the provision of feedback during learning had a positive effect on the learning trajectory and supported the acquisition of vocabulary, but did not affect the acquisition of word order. The authors argue that although some learning interventions (such as the provision of feedback) may generally improve learning, the effect may not be the same for different aspects of the language. This finding offers tentative support for the proposed dissociation between the acquisition of grammar and vocabulary, as predicted by the models of learning that distinguish cognitive processing systems serving vocabulary and grammar acquisition (Paradis, 2009; Ullman, 2004), although further

evidence is needed to test this dissociation, as far as the nondeclarative knowledge of vocabulary is concerned (e.g., Nakata & Elgort, 2020).

While addressing quintessential SLA research questions, this interdisciplinary contribution uses comparative design methods from instructed language acquisition research and taps into a number of theoretical and computational models of learning. For example, Monaghan et al. use insights from child L1 acquisition research from spoken input to motivate their cross-situation study of an artificial language and draw on theories of cross-situational learning and cross-situational statistics (e.g., propose-but-verify, Trueswell, Medina, Hafri, & Gleitman, 2013; and associative learning theory, McMurray, Horst, & Samuelson, et al., 2012) in their discussion of the study results. In line with these theories, the study showed that feedback can support inductively derived cross-situational information. However, as the authors argue, these theories are insufficient to explain the observed effects of local context on learning and will need to be revised to include attentional mechanisms needed to explain learning from more realistic learning contexts. Thus, Monaghan et al.'s contribution to the Special Issue highlights advantages of adopting an interdisciplinary stance in conducting L2 acquisition research.

4.3 Contextual L2 words learning from reading under massed and spaced conditions

The study by Nakata and Elgort (2020) concerns one of the most robust phenomena in experimental psychology, the effect of spacing on learning. However, traditionally, this effect has been observed under deliberate paired-associated learning conditions, while the present contribution investigates whether the superiority of spacing holds for contextual learning, when participants are encountering learning targets in meaningful contexts and are not explicitly instructed to memorise them. Because experimental L2 word learning

research into spacing under contextual learning conditions is limited, the authors used learning and memory research from cognitive psychology (e.g., Karpicke & Roediger, 2008; Kornell & Bjork, 2008) to generate a (non-directional) hypothesis in relation to the spacing effect in contextual learning. On the one hand, spaced encounters result in multiple learning episodes and multiple opportunities for knowledge retrieval, creating *desirable difficulty* (Schmidt & Bjork, 1992) predicted to improve learning. On the other hand, massing may facilitate inductive learning from context, with learners discovering core features of meaning by juxtaposing multiple instances of use; whereas spacing may reduce learners' ability to notice similarities between individual instances of use.

Nakata and Elgort's research into the spacing effect in contextual word learning also considers a theoretically important dissociation between explicit/declarative and tacit/nondeclarative knowledge, commonly considered in SLA studies in relation to the acquisition of grammar (Suzuki & DeKeyser, 2017) and in learning and memory studies in cognitive psychology (Reber, 2013; Schacter, 1987). The finding that the choice of a repetition schedule affects explicit but not tacit lexical knowledge (operationalised as semantic priming in a lexical decision task) is aligned with the predictions of the explicit/tacit knowledge dissociation. Tacit knowledge is assumed to be acquired implicitly through repeated exposure in contexts that provide sufficient meaning constraint and is less likely to be affected by the success of deliberate contextual inferences or explicit feedback than explicit knowledge. Since the number of exposures and contexts were held constant in both spaced and massed learning schedules, the learning condition was less likely to affect implicit learning and, consequently, tacit knowledge. Nakata and Elgort's study thus provides a level of support for the relevance of the explicit/implicit memory dissociation in the acquisition of lexical (and not only grammatical) L2 knowledge.

4.4 Language users' intuitions about fixedness and compositionality of word combinations Adopting the phraseological framework, Fioravanti et al. (2020) investigated L1 and L2 Italian speakers' sensitivity to and intuitions about fixedness and compositionality of free combinations, collocations, and idioms. The phraseological tradition classifies word combinations in terms of lexical fixedness and compositionality, where fixedness refers to the substitutability of a word within a phrase, and compositionality refers to how much the meanings of individual components contribute to the meaning of the phase. The phraseological approach places word combinations along a continuum, with the least restricted free combinations and the most restricted idioms being at the far ends, and collocations somewhere in the middle of the continuum (Cowie, 1994; Howarth, 1998; Hausmann, 1989; Mel'čuk, 1998; Nesselhauf, 2003). The results of the present study offer novel theoretical insights into how lexical fixedness of three combination types is perceived by more and less experienced language users. While L1 speakers operate on the idiom principle, according to which "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices" (Sinclair, 1991, p. 110), L2 speakers operate on the open choice principle, which is "a way of seeing language text as the results of a very large number of complex choices" (Sinclair, 1991, p. 110). It seems L1, but not L2, users are able to select conventional ways among a range of perfectly grammatical and meaningful, but "non-native-like or highly marked usages" (Pawley & Syder, 1993, p. 191). Further, L1, but not L2, speakers' intuitions closely mirror the lexical flexibility-based distinction between free combinations, collocations, and idioms proposed by phraseological models (Cowie, 1981; Howarth, 1998) and, in particular, by Ježek's (2005) and Masini's (2009) classifications specific to the Italian language.

In addition, Fioravanti et al (2020) provide further support to the claim that collocations are challenging for L2 learners. Both low and high proficiency learners judged the modified forms as more acceptable than the original ones. This finding aligns closely with earlier studies that report on L2 learners' tendency to substitute the original verb with a synonym (Nesselhauf, 2003; Gyllstad, 2005; Laufer & Waldman, 2011).

A question that arises in second language – the locus of L1 activation during L2 word processing

A question that arises in second language acquisition is whether L1 is activated during L2

word processing and, if yes, what aspects of L1 word knowledge are activated. Although

activation of L1 during L2 word processing, both at the lexical and sub-lexical level, has now

been firmly established for pairs of alphabetic languages (Dijkstra & Van Heuven, 2002), the

contribution by Chen et al. (2020) tests whether cross-language activation also occurs in

bilingual processing of pairs of distant languages with different writing systems, such as

Chinese and English. Since automatic cross-language activation has been observed in some

studies with Chinese speakers of English (e.g., Zhang et al., 2011) but not others (e.g., Wen

& Van Heuven, 2018), Chen et al.'s study offers additional evidence to show that L1

meaning equivalents of L2 words are activated, even when the task does not require explicit

L1 lexical processing (i.e., colour judgements). Chen et al. interpreted this L1 involvement in

L2 processing as the legacy of prolonged L1 (Chinese) involvement in the acquisition of

English as a foreign language.

Importantly, Chen et al. went beyond the commonly investigated cross-language interaction at the lexical level. They tested whether L1 sub-lexical orthography is activated during L2 word processing by manipulating the relationship between the translation equivalent of the English (L2) prime and the target Chinese character. This relationship

varied from a complete overlap (same Chinese character); to semantic and orthographic overlap (i.e., the two characters were similar in meaning and shared a semantic radical); to orthographic but not semantic overlap (a shared semantic radical only), to no overlap. Because the authors found no effect of the orthographic only overlap in any of their three experiments, they proposed that cross-linguistic activation of L1 Chinese during L2 English processing is contingent on the meaning overlap and is, thus, restricted to lexical (characterlevel) processing. More crucially, Chen et al. (2020) compared the activation of characters in reading English with within-language activation during reading Pinyin (an alphabetic writing system in Chinese) using the same implicit and explicit tasks. Different from findings in reading English, reading Pinyin activated the corresponding character and the sub-lexical orthography (Chen, Perfetti, Fang, Chang & Fraundorf, 2019). The differences in the activation at the sub-character level between reading English and Pinyin indicate that the sub-lexical activation is constrained by the language relatedness. Unlike models of bilingual word processing based on evidence from alphabetic L1-L2 pairs that model cross-language activation at the level of sub-lexical (orthographic) representations (e.g., BIA+, Dijkstra & Van Heuven, 2002), Chen et al. conjecture that, for pairs of unrelated languages with unrelated writing system, models of bilingual word processing may need to be revised to reflect the absence of interaction at the sub-lexical level of processing.

5. Conclusion

In this Special Issue, we set out to highlight how interdisciplinary research paradigms and advances in methods of data analysis can be used to study the complex multifaceted nature of L2 lexical knowledge.

The contributions show the benefits of using interdisciplinary approaches in motivating, designing, and implementing research into different aspects of L2 lexical knowledge and processing. Crowdsourcing knowledge data from thousands of participants can create a comprehensive inventory of the words known by L2 speakers of English from diverse L1 backgrounds. General theories of learning and memory and theories of cross-situational learning from child L1 acquisition research can inform hypotheses in L2 vocabulary acquisition research. Experimental priming paradigms from psycholinguistics provide insights into aspects of L2 lexical knowledge that cannot be measured in offline word knowledge tests and offer ways of ascertaining the locus L1 involvement in L2 lexical processing. The use of advanced statistical methods of data analysis affords a more precise understanding of the mechanisms underpinning L2 lexical acquisition in a novel language, and how the development of lexical knowledge and the development of grammatical knowledge interact over time.

A growing interest in using interdisciplinary approaches in second language research has been evident in the topics of recent special issues of this journal, including the Special Issue on *Neurolinguistics and the Language Classroom*, published in 2018, edited by Benati and Rastelli, and the Special Issue on *Eye tracking*, published in 2020, edited by Godfroid, Winke, & Conklin. Our Special Issue further lobbies for a continuous, deliberate and comprehensive interdisciplinary agenda in L2 lexical studies.

Acknowledgments

We would like to thank the contributors, who are SLA researchers and scholars from the neighbouring fields, for responding to our invitation and generously sharing their innovative interdisciplinary studies in this Special Issue. We are also very grateful to the ten expert international reviewers, who took part in the double-blind peer review process, and to the *Second Language Research* journal Editors, Silvina Montrul and Roumyana Slabakova, for their guidance and editorial support that contributed to the quality of this Special Issue.



References

- Bijeljac-Babic, R., Biardeau, A. & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition*, *25*, 447-457.
- Bod, R. (2006). Exemplar-based syntax: How to get productivity from exemplars. The Linguistic Review, 23, 291–320. doi:10.1515/ TLR.2006.012.
- Brysbaert, M., Keuleers, E., & Mandera, P. (2020). Which words do English nonnative speakers know? New supernational levels based on yes/no decision. *Second Language Research*, XX-XX. https://doi.org/10.1177/0267658320934526
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models:

 A tutorial. *Journal of Cognition*, 1, 9. http://doi.org/10.5334/joc.10.
- Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 441–458.
- Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals:

 Evidence from masked phonological priming. *Journal of Experimental Psychology:*Human Perception and Performance, 25, 137–148.
- Chen, L., Perfetti, C.A., Fang, X., & Chang, L.-Y. (2020) Activation of L1 orthography in L2 word reading: Constraints from language and writing system. *Second Language***Research, XX-XX. https://doi.org/10.1177/0267658320927761
- Chen, L., Perfetti, C. A., Fang, X., Chang, L. Y., & Fraundorf, S. (2019). Reading Pinyin activates sublexical character orthography for skilled Chinese readers. *Language, Cognition and Neuroscience*, 1-11.
- Cowie, A.P. (1981) The treatment of collocations and idioms in learners' dictionaries.

 Applied Linguistics, 2, 223–235.

- Cowie, A.P. (1994). Phraseology. In Asher, R. E. (Ed.) *The encyclopedia of language and linguistics*. Oxford: Oxford University Press, pp. 3168–3171.
- Degani, T., Prior, A., & Hajajra, W. (2018). Cross-language semantic influences in different script bilinguals. *Bilingualism: Language and Cognition*, *21*, 782–804.
- Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.
- Fioravanti, I., Senaldi, M. S. G., Lenci, A., & Siyanova-Chanturia, A. (2020). Lexical fixedness and compositionality in L1 speakers' and L2 learners' intuitions about word combinations: Evidence from Italian. *Second Language Research, XX-XX*. https://doi.org/10.1177/0267658320941560
- Godfroid, A. (2020). Sensitive Measures of Vocabulary Knowledge and Processing:

 Expanding Nation's Framework. In Webb, S. (Ed.) *The Routledge handbook of vocabulary studies*, New York: Routledge, pp. 433-453.
- Gyllstad, H. (2005). Words that go together well: Developing test formats for measuring learner knowledge of English collocations. In Heinat, F. and Klingval, E. (Eds.) *The Department of English in Lund: Working Papers in Linguistics*, 5, 1–31.
- He, X., & Godfroid, A. (2019). Choosing words to teach: A novel method for vocabulary selection and its practical application. *TESOL Quarterly, 53*, 348-371. doi:10.1002/tesq.483.
- Hashimoto, B. J. & Egbert, J. (2019). More than frequency? Exploring predictors of word difficulty for second language learners. *Language Learning*, *69*, 839–872.
- Hausmann, F. J. (1989). Le dictionnaire de collocations [Dictionary of collocations]. In:

 Hausmann, F. J., Wiegand, H. E., and Zgusta, L. (Eds). *Wörterbücher, dictionaries,*

- dictionnaires: Ein internationals Handbuch zur Lexikographie. Berlin: de Gruyter, pp. 1010–1019.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*, 24–44.
- Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.
- Ježek, E, (2005). Lessico: Classi di parole, strutture, combinazioni [Lexicon: Classes of words, structures and combinations]. Bologna: Il Mulino.
- Karpicke, J.D., & Roediger, H.L. (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968.
- Keuleers, M., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment.

 **Quarterly Journal of Experimental Psychology, 68, 1665–1692.
- Kornell, N., & Bjork, R.A. (2008). Learning concepts and categories: Is spacing the 'enemy of induction'? *Psychological Science*, *19*, 585–592.
- Laufer, B., & Waldman, T. (2011). Verb–noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672.
- Lindstromberg, S. (2016). Inferential statistics in language teaching research: A review and ways forward. *Language Teaching Research*, 20, 741–768.
- Lindstromberg, S., & Eyckmans, J. (2017). The particular need for replication in the quantitative study of SLA: A case study of the mnemonic effect of assonance in collocations. Journal of the European Second Language Association, 1, 126–136. https://doi.org/10.22599/jesla.26.

- Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*, 321–391.
- Marsden, E., Plonsky, L. (2018). Data, open science, and methodological reform in second language acquisition research. In Gudmestad, A. and Edmonds, A. (Eds.) *Critical reflections on data in second language acquisition*. Philadelphia, PA: John Benjamins, pp. 219–228.
- Masini, F. (2009) Combinazioni di parole e parole sintagmatiche [Combinations of words and syntagmatic words]. In Lombardi Vallauri, E. and Mereu, L. (Eds.) *Spazi linguistic: Studi in onore di Raffaele Simone* [Linguistics spaces: Studies in honor of Raffaele Simone].

 Roma: Bulzoni, pp. 191–209.
- McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*, 831–877.
- Mel'čuk, I. (1998). Collocations and lexical functions. In Cowie A. P. (Ed.) *Phraseology:*Theory, analysis and applications. Oxford: Oxford University Press, pp. 23–53.
- Miwa, K., Dijkstra, T., Bolger, P., et al. (2014). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism:*Language and Cognition, 17, 445–463.
- Monaghan, P., Ruiz, S., & Rebuschat, P. (2020). The role of feedback and instruction on the cross-situational learning of vocabulary and morphosyntax: Mixed effects models reveal local and global effects on acquisition. *Second Language Research, XX-XX*. https://doi.org/10.1177/0267658320927741

- Nakata, T., & Elgort, I. (2020). Effects of spacing on contextual vocabulary learning: Spacing facilitates the acquisition of explicit, but not tacit, vocabulary knowledge. *Second Language Research*, XX-XX. https://doi.org/10.1177/0267658320927764
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nation, I. S. P. (2013). Teaching & learning vocabulary. Boston: Heinle Cengage Learning.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*, 223–42.
- Paradis, M. (2009) *Declarative and procedural determinants of second languages*.

 Amsterdam: John Benjamins.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Richards, J. C. and Schmidt, R. W. (Eds.) *Language and communication*. London: Longman, pp. 191–225.
- Perfetti, C. A. (2007) Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.
- Perfetti, C. A, & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.),

 On the consequences of meaning selection (pp. 67–86). Washington, DC: American

 Psychological Association.
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven. C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins.
- Porte, G. (2012). *Replication research in applied linguistics*. New York: Cambridge University Press.

- Read, J. (2000). *Assessing vocabulary*. Cambridge, UK and New York: Cambridge University Press.
- Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, *24*, 146–161.
- Rebuschat, P., Meurers, D., & McEnery, T. (2017). Language learning research at the intersection of experimental, computational, and corpus-based approaches. *Language Learning*, 67, 6-13.
- Reber, P. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*, 2026–2042.
- Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 501-518.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217.
- Sinclair, J. (1991). Corpus, concordance, collocation. Oxford: Oxford University Press.
- Suzuki, Y., & DeKeyser, R. (2017). Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research*, *21*, 166–88.
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, 104, 12530–35.
- Trueswell, J.C., Medina, T.N., Hafri, A., & Gleitman, L.R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*, 126–156.

- Ullman, M.T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–70.
- Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*, 458–83.
- Wen, Y., & Van Heuven, W. J. (2018). Limitations of translation activation in masked priming: Behavioural evidence from Chinese–English bilinguals and computational modeling. *Journal of Memory and Language*, 101, 84–96.
- Zhang, T., Van Heuven, W. J., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese–English bilinguals. *Psychological Science, 22,* 1237–1242.