BRIEF REPORT

Language experience shapes relational knowledge of compound words

Daniel Schmidtke[1] · Christina L. Gagné[1] · Victor Kuperman[2] · Thomas L. Spalding[1]

Abstract

Prior studies of noun–noun compound word processing have provided insight into the human capacity for conceptual combination (Gagné and Shoben *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 71 1997; Spalding, Gagné, Mullaly & Ji *Linguistische Berichte Sonderheft, 17*, 283–315 2010). These studies conclude that relational interpretations of compound words are proposed and appraised by the language system during online word recognition. However, little is known about how the capacity for creating new meanings from existing conceptual units develops within an individual mind. Though current theories imply that individual relational knowledge about the combinability of concepts develops as language experience accumulates, this hypothesis has not been previously tested experimentally. Here, we addressed this hypothesis in a task that assesses individual relational knowledge of English compound words. We report that greater experience with printed language shapes relational knowledge of compound words in two ways. Firstly, individuals with more experience with printed language were able to select a greater number of possible relational meanings for individual compound words. Secondly, individuals with greater experience with printed language were also more precise about which relational meaning was the most semantically plausible out of all possible meanings. Our results confirm that language experience affects an individual's ability to use relational knowledge in order to combine conceptual units. Our findings offer further support for the Lexical Quality Hypothesis (Perfetti, 2007), which states that lexical representations of words become simultaneously more flexible and precise as a result of repeated exposure to their orthographic forms in language usage.

Keywords Conceptual combination · Individual differences · Morphology · Compound word knowledge

Introduction

Producing new complex meanings from existing conceptual units is a hallmark of human creativity. Compound words, in particular, offer a unique opportunity to understand the mental apparatus that are involved in this capacity. Because compound word meanings are linguistically expressed by combining two or more existing words (e.g., the compound *sunburn* is formed by joining together the modifier

constituent *sun* and the head constituent *burn*), the study of the cognitive resources underpinning their comprehension can provide insight into the nature of conceptual composition.

Prior psycholinguistic research (for a review see Gagné & Spalding, 2009a, 2014) has pointed to the important function of *conceptual relations* for the online composition of compound word meaning. This evidence indicates that comprehending a compound involves proposing a semantic link between the conceptual units that are denoted by its constituents. For example, *sunburn* can be interpreted relationally as 'a burn that is caused by the sun'. However, though much behavioral research has provided evidence for the role of relational structures during compound word recognition (e.g., Coolen, van Jaarsveld & Schreuder, 1991; Estes, 2003; Gagné, Spalding, Figueredo & Mullaly, 2009b; Gerrig & Murphy, 1992), little is understood about how knowledge of the relational structure of compound words is *learned* by individuals. Such an understanding is crucial for psychological theories that require an

✉ Daniel Schmidtke
schmiddf@mcmaster.ca

[1] University of Alberta, Edmonton, Canada

[2] McMaster University, Hamilton, Canada

accurate description of human conceptual organization and development, including accounts of compound word learning and processing. We address this issue in the current study by examining how experience with language shapes an individual's relational knowledge of compound words.

The foundational assumption of relational theories of conceptual combination (Gagné & Shoben, 1997; Spalding, Gagné, Mullaly & Ji, 2010) is that nominal combinations, such as compound words, are understood via access to an unspecified relation that exists between their constituent concepts (for linguistic theories of relational structure in compounding, see Downing, 1977; Gleitman & Gleitman, 1970; Levi, 1978; Warren, 1978). The competition-among relations-in-nominals (CARIN; Gagné and Shoben, 1997) and the relational-interpretation-competitive-evaluation (RICE; Spalding et al., 2010) theories are psychological models which describe how relational knowledge is implemented in the mind of the language user. A central tenet of both accounts is that during the processing of novel complex meanings, people use knowledge about how certain relations are likely to be semantically associated with certain concepts. For example, *chocolate* is most likely to be used with a MADE OF relation, e.g., *chocolate cake* and *chocolate egg*. However, other relations are also possible, e.g., *chocolate shop*—'a shop FOR chocolate' or *chocolate factory*—'a factory that MAKES chocolate'. Importantly, a claim of the CARIN and RICE models is that sensitivity to which relations are used for a particular concept in a specific position (e.g., head vs. modifier) partly reflects the current state of that individual's knowledge about concepts and the way that they combine.

A crucial entailment of this claim is that relational knowledge is determined by a person's individual experience with relational information that is present in the linguistic environment. In other words, both theories imply that the capacity for producing complex meanings from smaller units of meaning is contingent upon the current state of a person's knowledge about which concepts are suitable for particular meanings (see also Štekauer, 2005). However, while many studies have assumed that relational knowledge relies on experience with language, none have explicitly examined whether and how relational knowledge might develop as a function of experience with language. This gap in our understanding leads to a number of questions about how humans develop mastery in conceptual combination. For example, do individuals become more productive in their ability to compute complex meanings as language experience accumulates? If so, how might a greater repertoire of combinatorial possibilities influence how the conceptual system is able to compute complex meanings? In what follows, we motivate the current study by connecting two streams of research, which together guide our predictions about the developmental trajectory of relational knowledge. The first concerns the specific cognitive underpinnings of relational knowledge and the evidence which points to its experience-dependent flexibility. The second strand of research specifies a general principle of how lexical representations are expected to develop with increasing exposure to language.

Sources of flexibility in relational knowledge

Experimental evidence suggests that distributional information about how relational interpretations combine with compound word modifiers (e.g., *cheese* in *cheesecake*) in language usage has an influence on the processing of compound words. In a sense-nonsense task and a lexical decision task, Gagné and Shoben (1997) demonstrated that the ease of comprehending a noun-noun combination is determined by the statistical likelihood of the association between a modifier constituent and a conceptual relation in language usage. For example, a combination such as *mountain cabin* was comprehended faster than *mountain magazine* because the LOCATED conceptual relation in *mountain cabin* is also used in many other compound phrases, which have *mountain* as a modifier, e.g., *mountain climber*, *mountain gorilla*, and *mountain retreat*. However, the ABOUT relation is not frequently used in compounds which have *mountain* as a modifier, and thus slowed down comprehension when processing *mountain magazine*. This *relational dominance effect* (see also Pham and Baayen, 2013) indicates that modifiers are flexible in that they combine with a great range of conceptual relations, and that the difficulty of selecting a conceptual relation during compound word recognition is a function of the likelihood of that relation for a particular concept as found in language use.

Recently, Marelli, Gagné and Spalding (2017) recaptured Gagné and Shoben's (1997) original relational dominance effect using a computational measure derived from Distributional Semantic Models (DSMs), an approach which is guided by the principle that a word meaning can be determined from the contextual usage of the word itself (Firth, 1957; Miller & Charles, 1991). Based only on co-occurrence statistics of words in text-based corpora and without any explicit hand-coded information concerning compound relational structure, Marelli et al.'s (2017) computational simulation replicated the behavioral pattern of Gagné and Shoben's (1997) sense-nonsense judgment task. Taken together, both Gagné and Shoben's original findings and Marelli et al.'s computational replication demonstrate that the processing of novel compound word phrases is influenced by distributional patterns concerning how a conceptual relation is to be used with a modifier constituent. Novel compound word phrases are processed faster when

they contain modifiers that are associated with a distinct set of conceptual relations in written language usage.

The flexible nature of conceptual knowledge for compound words is also reflected in the collective judgment of human participants. This has been demonstrated in the results from studies which use the *possible relations task* (Spalding & Gagné, 2014a; Schmidtke, Kuperman, Gagné & Spalding, 2016; Schmidtke, Gagné, Kuperman, Spalding & Tucker, 2017), a crowdsourcing method which collects human judgments about the appropriateness of conceptual relations for existing English compound words. These studies show that some compound words can be used with a more diverse set of conceptual relations than others, and that this matters for visual and auditory compound word processing. In the possible relations task, participants select the most appropriate conceptual relation (out of a choice of 16 different conceptual relations) for an individual compound word. The resulting frequency distribution of selections per compound show that for some compounds, participants converge on a single dominant meaning, while for other compounds they do not. For example, the relational interpretation 'man MADE OF snow' is almost always selected for the compound *snowman*, while for the compound *folksong*, the frequency of participants' selections are equally apportioned across 'song ABOUT folk', 'song FOR folk' and 'song BY folk'.

The RICE theory of conceptual combination (Spalding et al., 2010) captures the processing consequences of the flexibility of relational knowledge of established and novel compound words outlined in the studies reported above. This theory predicts that online compound word processing is guided by an automatic process of meaning construction and competition. At the first stage of processing, a number of relational meanings of established compounds are concurrently proposed, based on the ease of availability with the compound's modifier constituent. Next, the range of possible relational interpretations are evaluated. During the evaluation stage, the most dominant competitor out of the set of proposed meanings, i.e., the most likely relational candidate, is selected as the compound word meaning. Evidence for the competitive process of conceptual combination was recently reported by Schmidtke et al. (2016, 2018). They reported that the speed of auditory and visual word recognition is inhibited for compounds with greater entropy among the distribution of possible conceptual relations, i.e., when processing a compound such as *folksong* (see above example).

Importantly, while the possible relations task highlights the internal flexibility of relational information of compound words themselves, it may also tacitly point to the existence of systematic variability in relational knowledge that exists across *individuals*. Although it has not been shown yet, it is possible that the same experimental data can be used to evaluate the driving force behind individual variability in relational knowledge. A shift of focus from compound-specific relational flexibility to variability across individuals exposes a crucial gap in knowledge about the potential sources of such individual variability, and raises the important question about what this can tell us about how humans learn and develop knowledge about how concepts can be combined. This study aims to use the possible relations task to fill this gap in knowledge.

Altogether, the experimental and computational studies reviewed above provide reasons to believe that relational knowledge may develop as a function of language experience. The emerging picture from these studies is that people learn and exploit statistical regularities about the relational structure of combined concepts via their exposure to language. This view leads to the general prediction that individual relational knowledge about compound words is subject to constant flux as experience with distributional patterns of conceptual relations accumulates. As exposure to written language increases, we predict that the same modifiers and heads are encountered within different compound words across a more diverse range of contexts. As a consequence, it is conceivable that this increase in experience with distributional patterns of relational knowledge extends the set of potential conceptual relations that are able to be combined with modifiers and heads during the process of interpreting specific compound words. Based on this hypothesis, the current study therefore targets individual exposure to printed materials as a potential determinant of greater flexibility in relational knowledge of compound words.

Exposure to language and lexical quality

Our goal of investigating the impact of exposure to print on relational knowledge is also motivated by a wealth of evidence which shows that individual variation in language experience leads to individual differences in performance in behavioral tasks and psychometric assessments, which rely on general lexical knowledge. For example, individuals who have accumulated more exposure to printed materials exhibit shorter lexical decision latencies to isolated words (Chateau & Jared, 2000; Sears, Campbell, & Lupker, 2006), shorter gaze durations on words during naturalistic reading (Gordon, Lowder, & Hoedemaker, 2016; Moore & Gordon, 2015; Taylor & Perfetti, 2016), increased spelling accuracy (Burt & Fury, 2000; Stanovich & West, 1989), greater vocabulary knowledge (Beech, 2002; Lewellen, Goldinger, Pisoni, & Greene, 1993; Martin-Chang & Gould, 2008; Stanovich, West & Harrison, 1995; West & Stanovich, 1991), greater verbal fluency (Stanovich & Cunningham, 1992),

and greater reading comprehension (Martin-Chang & Gould, 2008). Furthermore, based on the results of a cross-sectional meta-analysis of print exposure from infancy to early adulthood, Mol and Bus (2011) claim that exposure to printed materials is the central driving force behind the acquisition of many facets of reading skill, including vocabulary knowledge and comprehension.

The link between experience with printed language and the development of lexical knowledge is schematized in the framework of the Lexical Quality Hypothesis (Perfetti, 1985, 2007; Perfetti & Adlof, 2012; Perfetti & Hart, 2002). According to Perfetti (2007), the reading outcomes outlined above can be attributed to the repeated exposure of printed forms of words, which gradually entrenches lexical knowledge in memory. With continual reinforcement via exposure to printed text, the hypothesis predicts that lexical representations eventually contain strong memory traces between orthographic, phonological, and semantic representations. Put simply, the Lexical Quality Hypothesis claims that more encounters with words results in stronger and more "crisp" lexical representations.

For compound words specifically, the costs and benefits of exposure to print on lexical quality have been demonstrated in two eye-tracking studies in which compound words were embedded within naturalistic contexts and read silently. First, Falkauskas and Kuperman (2015) reported that individuals with greater exposure to printed materials were more sensitive to the probabilistic bias of the presented spelling format (either spaced or concatenated) of compound words. For example, in natural language usage, the spaced format is more likely for certain compounds (e.g., *horse race* is more frequent than *horserace*), and the concatenated format is more probable for others (e.g., *ringtone* occurs more often than *ring tone*). Falkauskas and Kuperman (2015) reported that the effect of spelling bias was contingent on reading experience and compound frequency. That is, the facilitative effect of compounds presented in their more frequently occurring spelling format was greater for more experienced readers reading less frequent compound words. Thus, Falkauskas and Kuperman (2015) concluded that readers exploit certain statistical properties of the orthographic form of compound words (i.e., spelling bias), but only when these forms are sufficiently entrenched in memory.

In a second compound word reading study, Schmidtke, Van Dyke, and Kuperman (2017) identified a disadvantage of impoverished exposure to print on a semantic component of compound word processing. They reported that total reading times were longer for individuals with less exposure to print when reading semantically transparent compound words, i.e., reading speed was impeded when reading compounds with greater semantic similarity between the head constituent and the whole word, such as *storm* in *rainstorm*. However, the opposite behavioral pattern of facilitation was observed among individuals with more exposure to print reading the same semantically transparent compounds. This pattern of results demonstrated that increased exposure to printed words is associated with greater ease in discriminating between the closely related concepts that are denoted by transparent compounds and their heads (e.g., *rainstorm*). Based on these results, in the present study we consider the potential effect of experience with language on the sensitivity to semantic transparency of compound words.

As we reviewed, evidence from the literature on conceptual relations strongly suggests that knowledge about conceptual relations is acquired through greater experience with language. Furthermore, the evidence from online word reading studies outlined above shows that experience with language is important for fine-tuning semantic knowledge about compound words, and also in strengthening lexical representations of infrequent compound words. Consequently, individual variability in exposure to printed language may also be expected to lead to inter-individual differences in relational knowledge about compound words. In what follows, we address this issue in a study which adapts the design of the possible relations experiment originally implemented by Spalding and Gagné (2014a).

The current study

The purpose and design of the research reported here is motivated by the current gap in our understanding about the developmental trajectory of relational knowledge about compound words. Our predictions rely on drawing connections between the theorized experience-dependent nature of relational knowledge and two underlying developmental adjustments to lexical representations that are proposed within the Lexical Quality Hypothesis. Specifically, the Lexical Quality Hypothesis proposes two critical dimensions of lexical quality: *flexibility* and *precision* (2007, p. 359). Firstly, a "precise" lexical representation is argued to consist of specific and stable representations of form and meaning connections (Perfetti, 1992; 2007). Further specification of lexical precision is provided by Andrews and Hersch (2010, p. 312), who state that lexical precision "facilitates activation of the lexical representation corresponding to the sensory input and minimizes activation of competing alternatives". Secondly, a more "flexible" high-quality semantic representation, is assumed to be "more generalized", "less context-bound", and can also encompass a "fuller range of meaning dimensions" (Perfetti, 2007, p. 360). In the present study, we quantify these theoretical constructs of "flexibility" and "precision" for the purpose of testing the development of relational knowledge.

Flexibility of relational knowledge The CARIN and RICE theories imply that knowledge about conceptual relations for compound words becomes more diverse as exposure to language increases. That is, with more reading experience, more unique modifiers and heads are encountered in a greater number of different compound words and across a greater range of lexical environments. This increase in diversity of contexts in turn broadens the range of possible relations that can be plausibly combined with specific modifiers or heads. Thus, we hypothesize that increased exposure to printed materials enables an individual to conceive of a more diverse set of potential relational meanings for a specific compound word. We test this hypothesis experimentally using a modified version of the possible relations task. In this task, participants were asked to choose as many potentially meaningful relational interpretations as possible for a given compound word. We predict that individuals who have accumulated greater exposure to print during their lifetime will be able to select more possible relational meanings for a given compound word.

Precision of relational knowledge Although flexibility of relational knowledge may increase with print exposure, we also seek to characterize whether the precision of such knowledge changes with language experience. Entropy of conceptual relations (Schmidtke et al., 2016, 2017), a measure of competition between relational meanings, is an ideal operationalization of precision of relational knowledge. For instance, consider that while a specific person might be able to select three different possible interpretations for the compound *folksong*, e.g., 'song ABOUT folk', 'song FOR folk' and 'song BY folk', at the same time that person's most dominant interpretation relative to others might be 'song FOR folk'. The extent to which this meaning is the strongest competitor relative to others is an indication of that person's ability to precisely associate a single dominant relational meaning out of a set of competitors. We predict that increased exposure to language plays a role in refining the distinct ability to isolate a single compound meaning out of a set of competitors.

Entropy of conceptual relations is a valid metric of precision because it estimates convergence among members of the same population in the ability to single out one relational interpretation for a compound word at the expense of other potential relational meanings. Low entropy would mean that participants are collectively more certain and precise about the relation interpretation for a compound, yet high entropy would indicate that participants are collectively less specific about the most plausible conceptual relation. With data collected using the original possible relations task, we separately compute entropy of the distribution of first-choice relational interpretations for (1) the group of

participants with greatest printed exposure, and for (2) the group of participants with least exposure to print. Based on our hypothesis, we predict that as compared with the lower print exposure group, the higher print exposure group will be collectively more precise when selecting the most plausible relational interpretation for individual compounds, i.e., compounds will be associated with lower entropy values.

Based on the results of prior studies reviewed earlier, it is also possible that we observe differential effects of exposure to print on either of the above outcome variables to be modulated by compound frequency (e.g., Falkauskas & Kuperman, 2015) and semantic similarity between the head and the compound word (Schmidtke et al., 2017). To be explicit about these expectations, first it is possible that individuals with greater print exposure will demonstrate greater susceptibility to effects of compound frequency. The logic here follows from that of Falkauskas and Kuperman (2015), which is that only experienced readers have amassed a sufficient amount of lexical exposure to have encountered low-frequency compounds. For the current hypothesis, we argue that this increased exposure to low-frequency compounds provides sufficient learning opportunities to refine flexibility and precision in relational knowledge. We therefore might expect to find the greatest changes in both flexibility and precision for more experienced readers.

Secondly, it is also expected that the semantic transparency of the meanings of the compound and its head will influence the precision of relational knowledge about a compound. More transparent compounds, such as *carwash*, are likely to be associated with more easily computable relational meanings, e.g., 'wash FOR car' or 'wash USED BY car'. On the other hand, for a more opaque compound, such as *hogwash*, it is more difficult to propose a relational meaning that conforms with the established meaning of the word. This is because the concepts denoted by opaque compounds are not predictable from the meanings of their constituents. For this reason, should an effect of semantic similarity arise, we predict that more transparent compounds will prompt greater precision in relational knowledge as compared to opaque ones. However, based on the results of the eye-movement results of Schmidtke et al. (2017), we expect to find this pattern only in individuals with greater experience with printed language. That is, less-experienced readers may not benefit from the conjunctive activation that could be made available from similar concepts that are denoted by transparent compound words and their embedded head constituents. For this reason, we expect that the effect of semantic transparency on precision of relational knowledge is stronger among the most experienced readers.

Verification of either of these potential outcomes on flexibility and precision of relational knowledge was sought by testing for interaction effects between (1) compound

frequency and print exposure, and (2) head-compound semantic similarity and print exposure.

Methods

Stimuli

A total of 600 concatenated English compound words were randomly sampled from a larger corpus of compounds which was compiled by the first author. We ensured that compounds were selected to represent an equal spread of high and low frequencies.

Procedure

Possible relations task An online version of the possible relations task was administered on the Amazon Mechanical Turk data collection platform. The experiment was a modified version of the possible relations task (for an example of implementations of the original possible relations task, refer to Spalding & Gagné, 2014a; Schmidtke et al., 2016, 2017). The task was designed to elicit from participants possible ways of combining two words, as if they had never seen them used together before. For each trial of the task, a compound word was presented in a spaced format, such that both constituents were presented in the order that they appear in the compound itself (e.g., *snow ball*). Presenting the compound with a space ensures that the stimuli format matches the task's instructions of combining two separate words for the first time.

At the beginning of the experiment, participants were made aware that for each trial they will be presented with two words (both constituents of the compound word stimulus). Participants were instructed to pretend that they are learning English and are familiar with each of the constituents, but have never seen them used together. During each trial, participants were asked to complete two possible relations sub-tasks. Their first sub-task was to indicate the most likely literal meaning for the pair of constituents (this task is the same as the original possible relations task). Next, in a separate column during the presentation of the same compound, the second sub-task required participants to indicate as many other potential relations that might also make sense for the pair of constituents. Both sub-tasks were presented on the same page. Participants selected a checkbox to make each choice. An experimental session contained 100 critical compound word trials and four initial compound word practice trials.

For each sub-task, participants were asked to make their choices out of the set of 16 possible relational interpretations which were presented beneath the spaced compound (e.g., 'ball CAUSES snow', 'ball CAUSED BY snow', 'ball

HAS snow', 'ball MAKES snow', 'ball BY snow', 'ball FROM snow', 'ball MADE OF snow', 'ball IS snow', 'ball USED BY snow', 'ball USES snow', 'ball LOCATED snow', 'snow LOCATED ball', 'ball FOR snow', 'ball ABOUT snow', 'ball DURING snow', and 'ball BY snow'). The set of possible relations were adapted from Levi's (1978) original range of conceptual relations. Prior to the experiment, participants were provided with an exemplar of each relation.

Assessments of exposure to print After participants completed the modified possible relations task, they were asked to complete an updated version of Stanovich and West's (1989) Author Recognition Test and Magazine Recognition Test (ART and MRT; Acheson et al., 2008). The ART elicits information about the reading habits of individuals via an author identification task. This test estimates exposure to printed literature. Participants were presented with an inter-mixed list of 130 real and false author names printed on a computer screen and were required to select a checkbox next to a name they recognized as a real author. Of the entire list of names, 65 names are real authors (e.g., George Orwell) and the remaining 65 names are foils (e.g., Beatrice Dobkin). Participants are rewarded with one point for every author they guess correctly and penalized by one point for every author they guess incorrectly. The highest point score that could be achieved was 65. The MRT is an adapted form of the ART, which substitutes author names with current magazine titles. Combined with the possible relations task, the entire experiment took about 40–50 min to complete.

Participants

A total of 209 participants took part in the experiment. The master-list of 600 compounds was assigned to six separate 100-word sublists and was shown to an initial cohort of 168 participants (28 participants per list). All lists represented an equal spread of high- and low-frequency compound words. All responders were residing in the USA at the time of participation, and were paid USD $8.00 for a validly completed 100-word experimental list. Amazon Mechanical Turk does not permit (1) the enforcement of an anti-requisite criterion which would limit participants to the completion of a single experimental list, or (2) the pre-exclusion of bilingual and non-native English speakers. After the elimination of data collected from non-native English speakers or bilinguals, or from participants who had completed a second or third experimental session (we retained data only from a participant's first completed experimental list), a second wave of experimental lists was deployed. For the second wave of data collection, we re-sorted the master-list of 600 compounds into six new 100-word sub-lists based on the number of eliminated trials from the first wave of data collection. These new lists were then

shown to a further 41 participants where the distribution of participants per list was contingent on the number of trials required to reach at least 26 observations per compound word (see Results and discussion below for description of the distribution of selections per compound word). We applied the same participant removal procedures to the second wave of data. Furthermore, we also removed entire data for participants who selected the same conceptual relation more than 50% of the time. This cautionary procedure was performed in order to eliminate participants who potentially made repetitive arbitrary selections. In total, for both collection waves, we removed data from 23 participants. The final dataset contained responses collected from 186 MTurk workers (100 female; 85 male; one undisclosed) with an age range of 18–70 (median = 33, SD = 11).

Dependent variables

We tested the effect of exposure to print on two distinct item-specific measures of individual relational knowledge of compound words.

Individual relational diversity Our measure of flexibility of individual knowledge of conceptual relations, henceforth referred to as individual relational diversity, was estimated as the total count of unique relational interpretations made by each participant for each compound. For example, for the compound *barmaid*, a hypothetical participant may have selected the 'maid FOR bar' as the most likely relational interpretation, and 'bar HAS maid' and 'maid LOCATED IN bar' as other possible relational interpretations. This particular trial therefore elicited a count of three individual relational interpretations. The intuition behind this variable is that a higher count of different possible relations for a specific individual for a specific compound word, indicates that that individual is has a greater diversity of relational knowledge about that specific compound word.

Entropy of conceptual relations Entropy of conceptual relations served as the operationalization of experience-dependent precision about the possible relational interpretation for a compound word. First, we created the two quasi-experimental conditions of high and low print exposure by splitting the set of responses to the 'most likely relational interpretations' at the median value of the ART score (13). This step produced two separate matrices containing the same compound word items; one containing the possible relations task results for readers with high exposure to printed materials and the other containing the same results but for readers with low exposure to printed materials. In each matrix, each cell indicates how many

participants chose a particular relation as the best interpretation for a particular compound. We then computed Shannon entropy over the set of relations per compound word for each compound word separately (see Schmidtke et al., 2016, 2018). Entropy of conceptual relations is defined as $H = -\Sigma p_i \log_2 p_i$, where p_i is the probability of a relation within the respective distribution of chosen possible relations for a given compound. To eliminate noise, and to reduce the potential influence of accidental responses, the probability distribution was estimated only for relations that were selected more than once. The resulting dataset contains two estimates of entropy of conceptual relations for a single compound word: one for individuals with high exposure to printed materials, and one for readers with low exposure to printed materials. Entropy is low when a relational meaning is dominant and entropy is high when there are many relational interpretations in the distribution with similar selection probabilities.

We correlated both dependent variables as a test of the strength of their relationship. As indicated by a Spearman correlation, there is a weak and non-significant association between individual relational diversity and entropy of conceptual relations ($r_s = 0.005$, $p = 0.53$). This correlation indicates that both dependent variables are measures of two theoretically independent constructs. Distributional characteristics of dependent variables are provided in Table 1, along with the distributions of the dichotomized ART group conditions.

Independent variables

Frequency and family size As control variables we included compound frequency and the separate stand-alone frequencies of their modifier and head constituents. Raw word frequency counts were obtained from the 51 million-token SUBTLEX-US corpus (Brysbaert & New, 2009), based on subtitles from US film and media, for all compound words and their respective constituents. Frequency-based characteristics all pertain to compounds in their concatenated format. We also considered the positional family size of the left (modifier) and right (head) constituents (defined as the number and summed frequency of compounds that share a constituent with the fixed position of either the left or right constituent of the target compound). Family size estimates were calculated with the morphological parsing of the 18 million-token English component of the CELEX lexical database (Baayen, Piepenbrock & Van Rijn, 1995). Both family size measures did not reach significance nor improve model fit across all analyses and are therefore not discussed further.

Semantic similarity In the present study, we wanted to ensure that the number of conceptual relations a person picks

Table 1 Descriptive statistics for the dependent and independent variables

Variable	Original			Transformed		
Dependent variables	Range	Mean	SD	Range	Mean	SD
Individual relational diversity	1:16	3	2			
Entropy of conceptual relations						
Low print exposure	0.45:2.8	1.50	0.49			
High print exposure	0.5:3	1.42	0.51			
Independent variables						
Compound frequency	1:3272	71	222	-1.7:3.1	0.01	1
Modifier frequency	3:133117	7658	14822	-3.8:2.4	0	1
Head frequency	1:152523	13129	24368	-4.1:2	0	0.99
Modifier-whole semantic similarity (LSA)	-0.01:0.64	0.27	0.13	-2.2:2.9	0.01	1
Head-whole semantic similarity (LSA)	-0.05:0.64	0.26	0.13	-2.5:3	0	1
Print exposure (continuous)	1:53	17	12	-2.5:2.3	-0.01	1
Print exposure (categorical)						
Low print exposure	1:13	7.2	3.6	-2.5:-0.12	-0.81	0.52
High print exposure	14:53	26	9.9	-0.03:2.3	0.83	0.62

Reported are the range, mean, and standard deviations of the original and transformed variables after selection and trimming procedures

(individual relational diversity), or entropy of conceptual relations, are not just a reflection of the strength of the semantic similarity between each of the constituents separately, and the whole word. To this end, we controlled for potential effects of the semantic association between constituents and compound words by estimating the semantic similarity between (i) the compound word meaning and its modifying constituent (Modifier-whole similarity; e.g., *dog* and *doghouse*), and (ii) the compound word meaning and its head constituent (Head-whole similarity; e.g., *house* and *doghouse*).

We operationalized semantic similarity using Latent Semantic Analysis (Landauer & Dumais, 1997). The term-to-term LSA scores for Modifier-whole and Head-whole similarity for the target compounds were collected with a default setting of 300 factors and a window of six words (Mandera, Keuleers, & Brysbaert, 2017). These LSA scores were calculated over word occurrences in the 51 million-token SUBTLEX-UK corpus of film subtitles (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). The degree of semantic similarity between the words that those vectors represent is estimated by the cosine of the angle between the vectors, ranging from 0 to 1. We inverted the raw scores (by multiplying by negative one) that were collected from the web interface for comparability with other LSA scores. After the inversion, values closer to 1 imply a greater semantic similarity between the pair of words under comparison.

Spacing bias To control for effects of the presentation format of the stimuli, we included spacing bias of compounds as a predictor in our statistical models. Separate frequency of occurrence counts of compounds in spaced (e.g., *snow man*) and concatenated (e.g., *snowman*) were obtained from a Wikipedia-based sample of compounds (Kuperman & Bertram, 2013a). The estimate of compound bias towards spacing was computed as the number of spaced occurrences of a compound divided by the total number of that compound's occurrences in spaced and concatenated formats.

To anticipate the result of this variable, spacing bias did not have an effect on either the flexibility (individual diversity) or precision of relational knowledge (experience-based entropy). In addition, spacing bias did not interact with frequency or reading experience in two-way interactions, nor in a three-way interaction of spacing bias, reading experience or word frequency. Furthermore, spacing bias did not improve model fit, and so this variable was excluded from final statistical models. These findings suggest that individual performance in the possible relations task individuals is not influenced by statistical information from language usage about the preference for a particular spacing format of a compound. This finding is expected since the process of relational composition, as outlined by the RICE theory (Spalding et al., 2010), is predicted to operate over semantic information, and is indifferent to formal properties of the surface form of the compound (such as the distribution of a compound's spacing format in orthography). This hypothesis was supported by the results of previous visual and auditory word recognition studies (Schmidtke et al., 2016, 2018), in which lexical decision response times were predicted by entropy of conceptual relations. Crucially, in these

experiments, the measure of entropy of conceptual relations affected visual and auditory word recognition responses despite the fact that the concatenated format of the presented stimuli in visual lexical decision, which is a feature that is also naturally absent in the auditory signal, was distinct from the spaced format of the possible relations task from which the entropy measure was computed. Thus, since the spacing measure was not an influential predictor variable in either analysis, we do not discuss this measure further.

Print exposure The results of the ART and MRT tests served as a measure of exposure to print, and as a proxy of language experience. The respective distributions of the ART and MRT test scores were strongly right skewed. To normalize the distribution, we applied the square root transformation. A Shapiro-Wilk test indicated that the transformed ART scores are normally distributed ($W = 0.987$, $p = 0.099$). After transformation, the MRT scores did not demonstrate any reliable influence on either of the dependent measures under consideration here. Because the MRT test did not demonstrate any effects on our dependent variables of interest, we no longer report any results pertaining to this variable. The correlation between raw scores of ART and lexical control variables was expectedly low (all rs < 0.001), and thus did not pose a problem of collinearity in statistical analyses. We henceforth refer to ART as Print Exposure.

Distributional characteristics of predictor variables are provided in Table 1 and a correlation matrix is provided in Table 2.

Statistical considerations

Generalized linear mixed-effects regression models and linear mixed effects regression models were fitted to the dependent variables (Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000). Individual relational diversity and entropy of conceptual relations necessitated a different error term distribution, and thus different distributional families were used for each analysis. Because individual relational

Table 2 Correlation matrix of lexical variables

	1.	2.	3.	4.
1. Compound frequency				
2. Modifier frequency	0.16***			
3. Head frequency	0.16***	0.08		
4. Modifier-whole LSA	0.15***	−0.07	0.05	
5. Head-whole LSA	0.22***	0.04	−0.07	0.29***

Lower triangle provides Spearman correlation coefficients

*** Correlation is significant at the .001 level.

** Correlation is significant at the .01 level.

* Correlation is significant at the .05 level.

diversity represents count data, its effect was tested using a generalized mixed-effects regression model with an underlying Poisson distribution (identity link function). Linear mixed-effects models with a Gaussian distribution were employed to test for the effect of experience-based entropy of conceptual relations (a continuous measure). We used the `lme4` (Bates, Maechler, Bolker, & Walker, 2013) package in the R statistical computing software program (R Core Team, 2014) to compute statistical models.

Print exposure scores (as a continuous variable in the first analysis, as a grouped categorical variable in the second analysis), compound frequency, as well as the separate frequencies of the compounds modifier and head, semantic similarity of the modifier and the whole word, and semantic similarity of the head and the whole word were treated as fixed effects. We used logarithmically transformed values of frequency-based predictor variables to account for their distributional skewness. All continuous independent variables were *z*-transformed to allow the predictors to be compared on the same scale. Collinearity between all lexical predictors was low (multicollinearity condition number, κ = 6.5), indicating that any overfitting problems that may have been caused by multi-collinearity between predictor variables are not an issue in our statistical analyses. We used effect coding for the experience-based grouping factor in the second analysis (low experience = -1; high experience = 1). Models also included by-item and by-participant random intercepts. Because the performance of each participant in our continuous measure of print exposure (the analysis of flexibility of conceptual relations) is represented by one value (and hence offers no variability), we did not model by-participant random slopes of individual-differences tests.

For each analysis, we refitted the final model after removing outliers from both data sets by excluding standardized residuals exceeding -2.5 or +2.5 standard deviations (Crawley, 2002; Baayen & Milin, 2010). In a separate preliminary analysis we tested the effect of experimental list on individual relational diversity and entropy separately. This variable did not reach significance and was thus not included as a fixed effect. We also tested the influence of trial order. Trial order had a main effect on both dependent measure (see results), but did not interact with any other critical predictor variable. In addition to providing estimates from the linear-mixed effects models, the significance of the fixed effects were also determined with type-II Wald tests using the ANOVA function provided by the `car` package (Fox & Weisberg, 2011). Finally, we also tested for a potential nonlinear nature of the critical main effect of print exposure and other predictors using generalized additive mixed effects regression (GAMM) models (Hastie & Tibshirani, 1990; Wood, 2006) using the `mgcv` package 1.8-17 (Wood, 2017). No effect showed an appreciable deviation from a linear/planar functional form.

Results and discussion

The initial data set included 18,600 critical trials, where a single critical trial is defined as a participant's set of responses to both subtasks of the possible relations experiment for a given compound word. The data cleaning process included the removal of all trials for which a response was not given. We also removed any selections made in the second sub-task that were duplicates of the selections made in the first sub-task, i.e., a participant's selection for 'another possible relational interpretation' sub-task was removed if that choice was the same as that participant's 'most likely relational interpretation' for the same compound word. We also considered only trials to all compound words for which a full set of lexical characteristics were available (see Independent variables above). In total, this process involved removing 844 (4.5%) trials. The final data source contained a total of 17,756 trials to 579 compound words. The number of judgments per compound ranged from 26 to 38, with a median of 30 judgments provided per compound word. The data, including the distribution of selections for each relation per compound across each group of reading experience is available as an online supplemental dataset.

Does print exposure influence flexibility of relational knowledge?

The generalized linear mixed effect model fitted to individual relational diversity indicated that a critical main effect of print exposure was reliably present. Print exposure was a significant predictor of individual diversity of conceptual relations [$\hat{\beta}$ = 0.24; SE = 0.1; z = 2.3; p = 0.02], such that participants with a relatively greater amount of exposure to print selected a greater number of possible conceptual relations per compound word. The top left panel of Fig. 1 presents this result visually. This finding confirms our hypothesized prediction that greater amounts of exposure to print results in greater flexibility of relational knowledge for compound words. Individuals who have greater reading experience, as indexed by a higher ART score, choose more possible relations for a compound word relative to individuals with less exposure to print.

In addition, compound frequency was a valid predictor of individual relational diversity [$\hat{\beta}$ = -0.057; SE = 0.021; z = -2.7; p = 0.007]. As can be seen in the bottom right panel of Fig. 1, participants selected fewer different possible relational interpretations for high-frequency compounds. We discuss the main effects of compound frequency and print exposure further in the General discussion.

Moreover, the model revealed that individual diversity of conceptual relations was reliably predicted by modifier-whole semantic similarity [$\hat{\beta}$ = 0.16; SE = 0.02; z = 8; p <

0.0001; Fig. 1, top right panel] and modifier frequency [$\hat{\beta}$ = 0.06; SE = 0.019; z = 3.2; p = 0.001; Fig. 1, bottom left panel]. These effects demonstrate that participants selected more possible relational interpretations for compounds with greater semantic similarity between the whole compound word and its modifier (e.g., *thunder* in *thunderstorm*), and for compounds with more frequent modifiers. We did not observe any reliable interaction effects between print exposure and lexical variables on individual relational diversity. Table 3 summarizes the final model for individual relational diversity and Table 4 provides a summary of an analysis of variance for this model.

The results regarding modifier properties are consistent with previous findings (Gagné & Shoben, 1997; Spalding & Gagné, 2004) which showed modifiers seem to play a role in relation selection during compound word recognition. The RICE theory (Gagné & Spalding, 2014) accounts for these modifier-specific effects in an initial "relation suggestion stage" of conceptual combination, proposing that a person's recent and general experience with the way modifier's combine with relations is a first step of the process of conceptual combination. For example, the relational availability for *seasick* is initially based on an individual's knowledge of the relations used with *sea* specifically used in the modifier position in other compounds, independently from the joint meaning of the modifier and the head that is being processed (i.e., *seasick*). We interpret the modifier frequency and modifier-whole semantic similarity effects on the measure of individual diversity of conceptual relations as support for RICE's claim that the modifier plays a primary role in the proposal of possible semantic relations. The more frequent the modifier, the more likely it is to be used within other compound words and thus is associated with a greater diversity of conceptual relations. Furthermore, since the modifier plays a greater role in relational selection, greater attention is drawn to the modifier's semantic properties in relation to the whole compound word meaning. Here, less semantic similarity between the modifier and the meaning of the compound suppresses the amount of possible relational meaning selections because for compounds with weakly associated semantic relations between the meanings of their modifiers and the whole word (e.g., *brainchild*), the established meaning of the compound meaning more readily conflicts with the possible relational meanings.

Does print exposure influence precision of relational knowledge?

First, we found a significant main effect of print exposure on entropy of conceptual relations [$\hat{\beta}$ = 0.064; SE = 0.01; t = 6.5; p < 0.0001]. This main effect confirms that entropy of conceptual relations for the low print experience group
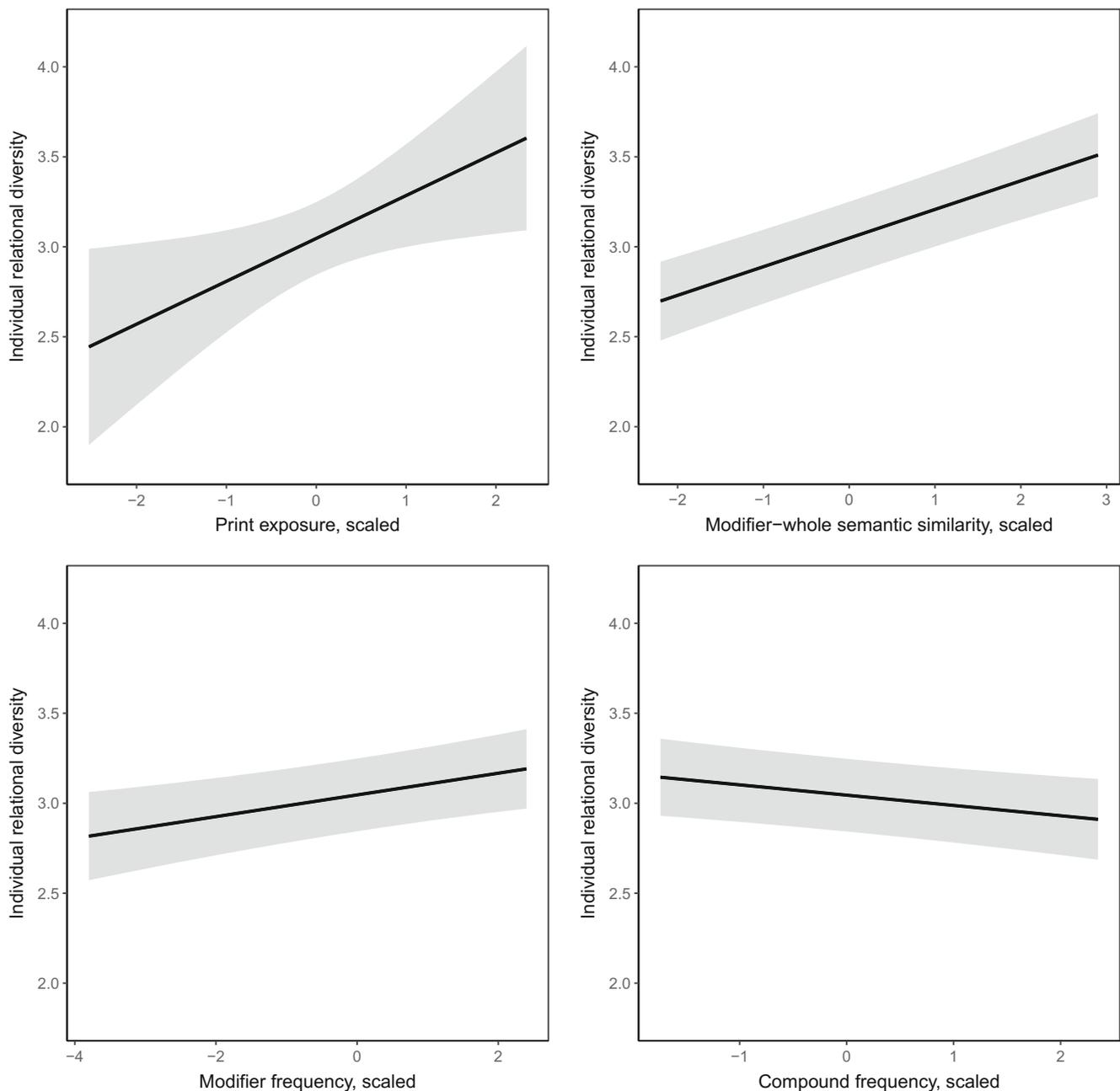
Fig. 1 Partial effects of print exposure (*top left panel*), modifier-whole semantic similarity (*top right panel*), modifier frequency (*bottom left panel*) and compound frequency (*bottom right panel*) on individual relational diversity

(M = 1.5, SD = 0.49) is significantly greater than that of the high print experience group (M = 1.42, SD = 0.51). We also found significant main effects of compound frequency [$\hat{\beta}$ = 0.049; *SE* = 0.022; *t* = 2.2; *p* = 0.029], modifier-whole semantic similarity [$\hat{\beta}$ = -0.1; *SE* = 0.022; *t* = -4.7; *p* < 0.0001], and head-whole semantic similarity [$\hat{\beta}$ = -0.044; *SE* = 0.022; *t* = -2; *p* = 0.049]. These main effects indicate that, irrespective of print exposure, there was a tendency for participants to be more random when selecting

a single most plausible conceptual relation for a compound word when the compound is more frequent. Moreover, participants were, on average, less random when selecting a single most plausible conceptual relation for compounds with greater semantic overlap between the whole compound word meaning and either of its constituents.

Furthermore, in the same model we found significant interactions between print exposure and compound frequency [$\hat{\beta}$ = -0.03; *SE* = 0.002; *t* = -13; *p* < 0.0001], and

Table 3 Fixed effects of the generalized linear mixed-effects model fitted to individual relational diversity

	Estimate	Std. error	z	p
Intercept	3.048	0.103	29.578	0.000
Print exposure	0.238	0.102	2.328	0.020
Compound frequency	−0.057	0.021	−2.714	0.007
Modifier frequency	0.060	0.019	3.227	0.001
Head frequency	0.024	0.019	1.278	0.201
Modifier-whole semantic similarity	0.159	0.020	7.995	0.000
Head-whole semantic similarity	−0.008	0.020	−0.419	0.675
Trial order	−0.033	0.016	−2.139	0.032

Conditional R^2 = 0.37. SD of the residual = 0.62. SD for the random effect of compound = 1.38. SD for the random effect of participant = 0.33. N = 17,756. N after trimming = 17,472

print exposure and head-whole semantic similarity [$\hat{\beta}$ = 0.032; SE = 0.002; t = 15; p < 0.0001] in a model predicting entropy of conceptual relations. We compared this model to an initial model in which print exposure, compound frequency and semantic similarity were included as main effects, without interactions. We compared the quality of model fit using Akaike Information Criteria (AIC) estimates as a diagnostic of model deviance. Comparison of AIC values after the inclusion of the interaction effects indicated that model deviance was significantly reduced, signifying that the inclusion of both interaction effects was justified (χ^2 = 150.96, p < 0.0001, AIC difference = 147). Finally, we did not observe a significant triple interaction of frequency x semantic similarity x print exposure. Table 5 summarizes the final model for entropy of conceptual relations and Table 6 provides a summary of an analysis of variance for this model.

The nature of both of the critical interaction effects is visualized in the plots in Fig. 2. The plots show that differences in entropy of conceptual relations scores between low and high reading experience groups were greatest for low-frequency compound words (upper panel), and compounds with greater head-whole semantic similarity (lower panel). Furthermore, both panels of Fig. 2 both show that the

effects of both compound frequency and head-whole semantic similarity were weaker among participants with low print exposure.

As the upper panel of Fig. 2 shows, less frequent compounds were associated with lower entropy values for participants with high print exposure relative to participants with low print exposure. This finding suggests that participants with more reading experience were less random, i.e., were more precise, in selecting their first-choice relational interpretation for less frequent compounds. This interaction pattern is exemplified in the raw data patterns presented in Fig. 3. Figure 3 is a plot of the distribution of conceptual relation selections for an example high-frequency compound and an example low-frequency compound broken down by reading experience. For the low-frequency compound—*saddlebag*—participants belonging to the low print exposure group (upper left panel) selected five relational interpretations almost the same number of times each, yielding a near uniform probability distribution of conceptual relation selections and the entropy value H = 2.3. In contrast, for the same compound, participants belonging to the high print exposure group (lower left panel) selected fewer conceptual relations, with one interpretation ("bag HAS saddle") being the single most frequently

Table 4 Analysis of variance table for the fixed effects of the generalized linear mixed-effects model fitted to individual relational diversity

Fixed effects	χ^2	df	p
Print exposure	5.421	1.000	0.020
Compound frequency	7.367	1.000	0.007
Modifier frequency	10.411	1.000	0.001
Head frequency	1.633	1.000	0.201
Modifier-whole semantic similarity	63.915	1.000	0.000
Head-whole semantic similarity	0.175	1.000	0.675
Trial order	4.576	1.000	0.032

Table 5 Fixed effects of the linear mixed-effects model fitted to entropy of conceptual relations

	Estimate	Std. error	t	p
Intercept	1.404	0.023	61.314	0.000
Print exposure	0.064	0.010	6.493	0.000
Compound frequency	0.049	0.022	2.190	0.029
Modifier frequency	−0.014	0.021	−0.641	0.522
Head frequency	0.036	0.021	1.719	0.086
Modifier-whole semantic similarity	−0.104	0.022	−4.685	0.000
Head-whole semantic similarity	−0.044	0.022	−1.973	0.049
Trial order	0.001	0.004	0.299	0.765
Print exposure * Compound frequency	−0.030	0.002	−12.653	0.000
Print exposure * Head-whole semantic similarity	0.032	0.002	14.702	0.000

Conditional $R^2 = 0.78$. SD of the residual = 0.24. SD for the random effect of compound = 0.43. SD for the random effect of participant is 0.11. $N = 17{,}756$. N after trimming = 17,431

selected one, yielding a relatively lower entropy value ($H = 0.78$). The right panels of Fig. 3 exemplify that this counter-entropic effect of print exposure was less dramatic for highly frequent compounds. For the example high-frequency compound—*sunlight*—the high print exposure group (lower right panel) was only slightly less random than the low print exposure group (upper right panel) in selecting a first-choice conceptual relation (low print exposure group, $H = 1.92$; high print exposure group, $H = 1.75$).

For the modulation of the effect of head-whole semantic similarity by print exposure, the lower panel of Fig. 2 shows that participants with more reading experience tended to be more precise in selecting a first-choice relational interpretation for a compound when there is greater semantic similarity between that compound word and its head constituent. To illustrate this effect, Fig. 4 provides participants' selections of relational interpretations for two example compounds, one with low semantic similarity between the head constituent and the compound word—*slipstream*—and one with high semantic overlap between the head con-

stituent and compound word, *surfboard*. As can be seen for the distribution of selections for the compound *slipstream*, there is little difference between the low (upper left panel) and high (lower left panel) exposure to print groups with respect to the amount of uncertainty in selecting a plausible relational interpretation. Individuals with more exposure to print were on average slightly less random in selecting a relational interpretation ($H = 1.89$) compared to individuals with less exposure to print ($H = 1.92$). However, for the compound *surfboard*, the difference in entropy across both groups was much greater. Individuals with greater exposure to print unanimously selected the interpretation "board FOR surfing", yielding the minimum value of entropy $H = 0$, while the selections for individuals with less exposure to print were spread more evenly across three possible interpretations (producing an entropy value of $H = 1.53$) We discuss both interaction effects described above and their relation to the results concerning flexibility of relational knowledge in the General discussion.

Table 6 Analysis of variance table for the fixed effects of the linear mixed-effects model fitted to entropy of conceptual relations

| Fixed effects	χ^2	df	p
Print exposure	43.044	1.000	0.000
Compound frequency	4.600	1.000	0.032
Modifier frequency	0.411	1.000	0.521
Head frequency	2.954	1.000	0.086
Modifier-whole semantic similarity	21.953	1.000	0.000
Head-whole semantic similarity	3.911	1.000	0.048
Trial order	0.089	1.000	0.765
Print exposure * Compound frequency	160.098	1.000	0.000
Print exposure * Head-whole semantic similarity	216.150	1.000	0.000
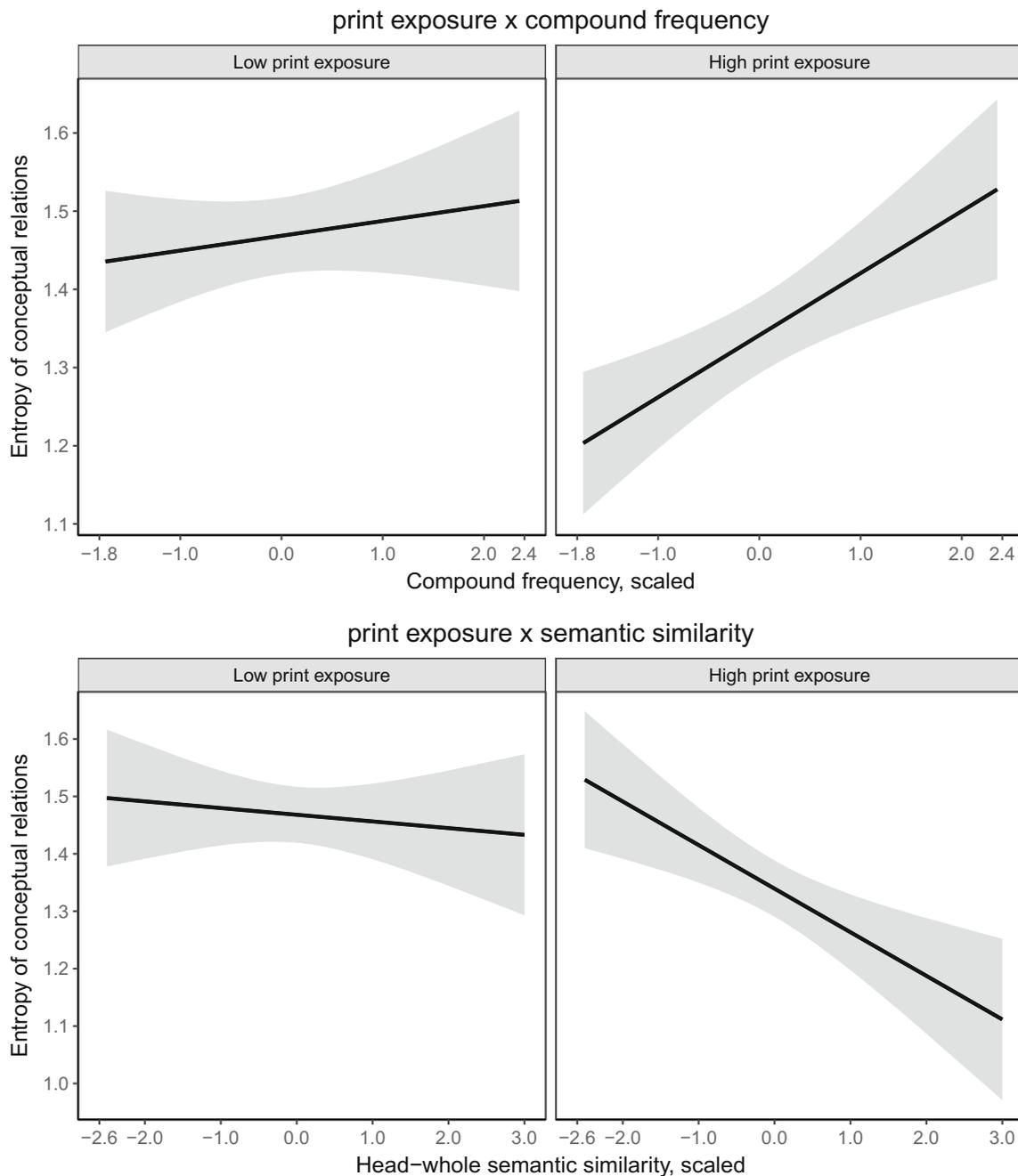
Fig. 2 A plot of the partial interaction effects between compound frequency and print exposure (*upper panels*), and head-whole semantic similarity and print exposure (*lower panels*)

General discussion

The current study asked whether and how experience with language shapes a person's relational knowledge about compound words. The key result is that individuals with greater print exposure provided a greater number of possible relational meanings for compound words, and, at the same time, were less uncertain about the most dominant relational

interpretation for the same set of compounds. Using an adaptation of the possible relations task (Spalding & Gagné, 2014a), the present study extends the theoretical base of the RICE hypothesis (Spalding et al., 2010) by demonstrating for the first time that a person's exposure to print modulates their relational knowledge about compound words. Our findings are also consistent with the theorizing of the Lexical Quality Hypothesis (Perfetti, 2007), in that they
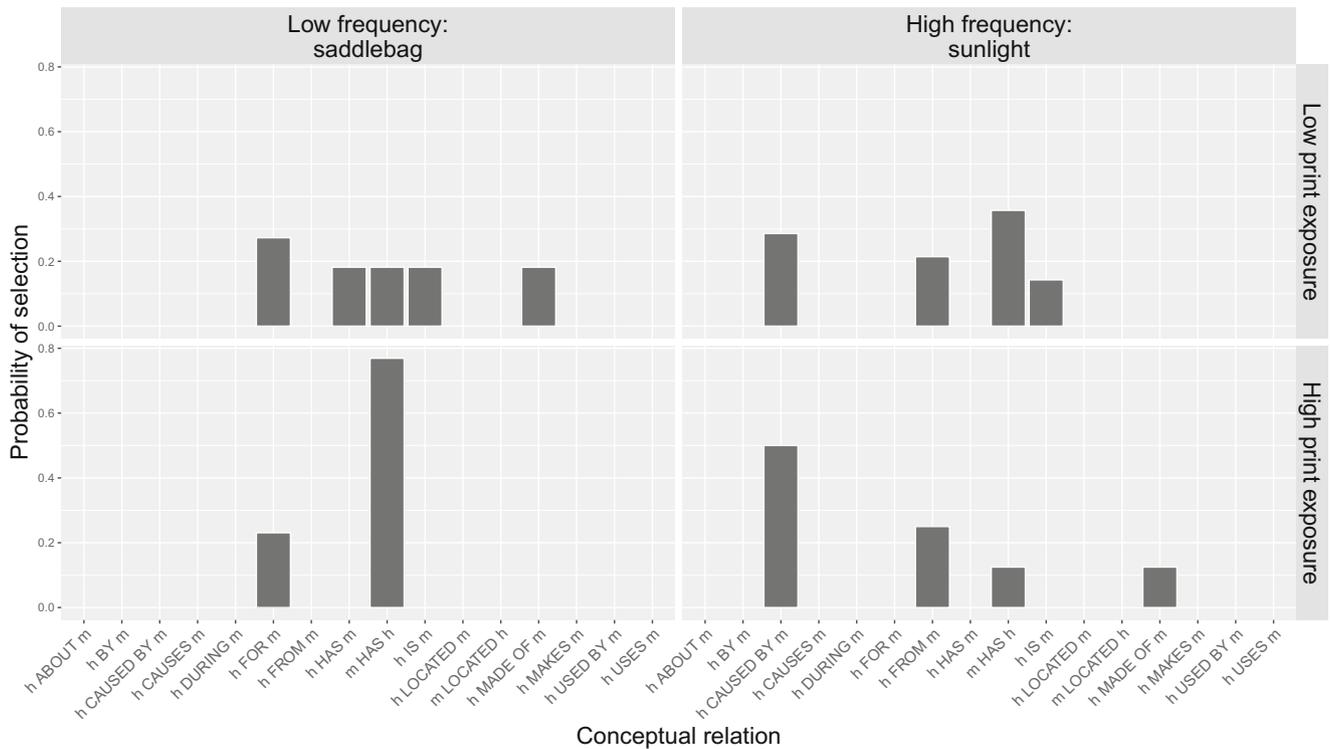
Fig. 3 Example data from the top relation subtask. Distributions of possible relations selections for a low-frequency and high-frequency compound broken down by exposure to print. h = head; m = modifier
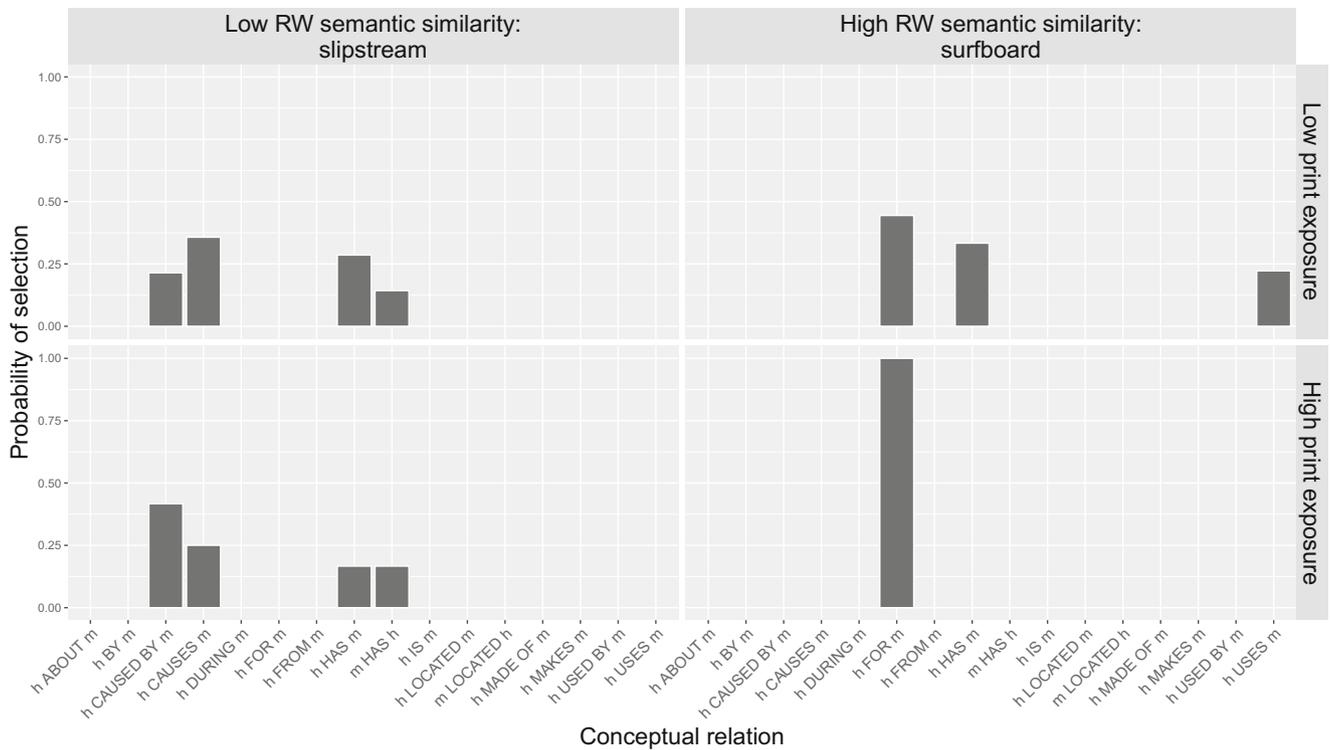

Fig. 4 Example data from the top relation subtask. Distributions of possible relations selections for a semantically opaque and semantically transparent compound broken down by exposure to print. h = head; m = modifier

show that semantic knowledge about words becomes both more flexible and precise as a consequence of increased exposure to printed materials.

So far, studies which have examined compound word processing as a window into conceptual combination have inferred that individuals implicitly absorb information about how concepts combine from patterns of natural language usage (Gagné and Shoben, 1997; Marelli et al., 2017). This study provides empirical support for this claim by showing that language experience simultaneously increases (1) an individual's diversity of knowledge about compositional meanings, and (2) that individual's capacity to refine the relational knowledge that is necessary for combining concepts. These findings therefore extend the RICE theory to include explicit predictions about how relational knowledge develops with language exposure.

The results of this study are also in line with the broader strand of research which concerns the nature of word learning and lexical entrenchment. Firstly, our results regarding the flexibility of relational knowledge conform with a specific prediction of the Lexical Quality Hypothesis, which is that lexical representations become richer and more flexible as individuals gain more exposure to words (Perfetti, 2007). Secondly, the results concerning modifications to precision of relational knowledge also chime with the predictions of the Lexical Quality Hypothesis, and do so in a way that converges with the individual differences literature emerging from compound and monomorphemic word processing. We discuss these findings below.

Compound frequency and individual differences in relational knowledge Our results show that relational information associated with more frequent compounds tends to be less flexible, i.e., high-frequency compounds are on average associated with fewer potential underlying relational interpretations (Fig. 1; bottom right panel). This finding, combined with the results regarding precision of relational knowledge suggests that, irrespective of print exposure, high-frequency compounds yield fewer possible relational meanings, but a relatively stable and high degree of competition between the most plausible of them (Fig. 2; upper panels). We interpret this pattern as a reflection of the cognitive entrenchment of highly frequent compounds. That is, more frequent compounds have more entrenched meaning representations, and therefore, the relational meanings that are suggested by participants are fewer and more competitive.

The pattern of results for high-frequency compounds contrasts with that of low-frequency compounds. For low-frequency compounds, all participants, on average, selected a more heterogenous set of conceptual relations. Because infrequent compounds are less likely to be encountered in naturalistic contexts by readers, it is possible that semantic knowledge about them is more diffuse and fuzzy, which

lead participants in the current study to associate with these compounds a more diverse set of possible meanings. As a result of this greater potential for semantic flexibility, we propose that the challenge of refining the precision of a relational meaning for low-frequency compounds requires a greater number of exposures to that word in context. Thus, it is only with sufficient increased exposure to these infrequent compounds that readers are able to narrow down the rich and heterogenous space of possible relational interpretations to a more precise relational meaning.

The observed interaction between compound frequency and language experience dovetails with prior research which shows that an individual's sensitivity to certain distributional characteristics of a word is contingent on the extent of entrenchment of the lexical form itself (Kuperman & Van Dyke, 2013b; Falkauskas & Kuperman, 2015; Rahmanian & Kuperman, 2017). Because individuals with less reading experience sample fewer texts, they are especially less likely to encounter low-frequency words relative to high-frequency words. Since less experienced readers have less opportunities to strengthen the quality of lexical representations for these low-frequency words, they suffer a greater processing penalty when recognizing these items (e.g., Ashby, Rayner, & Clifton; 2005; Davies, Arnell, Birchenough, Grimmond, & Houlson, 2017; Hawelka, Gagl, & Wimmer, 2010; Kuperman & Van Dyke, 2011, 2013b; Yap, Balota, Sibley, & Ratcliff, 2012), or are impervious to specific distributional properties that are associated with their infrequent orthographic forms (e.g., the likelihood of the compound word occurring in a spaced format, Falkauskas & Kuperman, 2015). The results of the current study add to this body of literature by presenting a case where deficits in specifically *semantic* representations arise from insufficient exposure to written word forms. Specifically, our results complement this literature by showing that it is only with increased experience with written language that individuals are able to more effectively overcome the demanding compositional burden that is posed by low-frequency compounds.

Semantic similarity and individual differences in relational knowledge The results of the present study show that there is a differential effect of semantic transparency on the precision of relational knowledge across individuals with less and more print exposure. Specifically, individuals with more experience with printed materials tend to exhibit greater precision in the relational interpretation of compounds with high semantic similarity between the head and whole word meanings. On the contrary, for less experienced readers, precision of relational knowledge is weakly affected by semantic similarity.

It is expected that knowledge about a possible conceptual interpretation for transparent compounds is more precise

because these compounds denote more clearly defined relational meanings. For example, a transparent compound such as *carwash* will naturally be easier to compute as opposed to an opaque compound such as *hogwash*, because constituents of transparent compounds are more easily combinable with a conceptual relation e.g., 'wash FOR car'. The results of the present study show that only individuals with greater exposure to print are able to benefit from this effect of transparency. We argue that, as a consequence of greater experience with language, individuals become better attuned to the conjunctive activation of the concepts denoted by compound words and their embedded constituents. This speculation is supported by studies which report facilitative effects of semantic transparency on compound word processing among experienced readers (El-Bialy, Gagné, & Spalding, 2013; Juhasz, 2007; Juhasz, Lai, & Woodcock, 2015; Libben, 1998; Libben, Gibson, Yoon, & Sandra, 2003; Schmidtke et al., 2017; Zwitserlood, 1994). Therefore, because individuals with greater reading experience benefit from greater sensitivity to conjunctive activation of semantic concepts, the problem of deciding upon a conceptual relation for a compound word is aided by the presence of semantic co-activation.

Our results show that less experienced readers are less sensitive to the additional benefit of semantic transparency for relational knowledge of compound words. As is depicted by the less experienced readers' flatter regression slope (Fig. 2, lower panels), this population remains impervious to knowledge about how semantically related the constituents of compounds are, and this impacts their ability to form a stable and precision relational meaning for the compound word. In other words, the semantic similarity between the head and compound for *hogwash* and *carwash* offer equally diffuse clues about a potential relational meaning. Following from the line of reasoning outlined above, we argue that this pattern is indicative of individual deficits in the conjunctive activation of semantically related concepts, which ultimately is a consequence of inexperience with printed language. This speculation aligns with Perfetti's hypothesis that reduced precision of lexical representations among less skilled readers is characterized by "less coherence of word identities" (2007, p. 362).

The exploration and exploitation of semantic information
The overarching pattern of results points to a relationship between flexibility and precision of relational knowledge. More experienced readers are individuals who tend to produce more diverse sets of relations, and are also the ones who more readily converge on the most probable relation for a compound. Is there a link between these two patterns? And if so, why might greater semantic flexibility also be associated with greater semantic precision? We propose that an answer to this might be found in the 'exploration versus

exploitation' trade-off hypothesis within the Reinforcement Learning framework (Sutton & Barto, 1998). The exploration versus exploitation hypothesis can be expressed in terms of a decision making dilemma. At the outset of skill learning, an individual's decision to either explore many opportunities, rather than conservatively exploit a narrow set of potentially suboptimal cues, may impact the developmental trajectory of that skill in that individual. Results from psychological tasks, including decision making in video games (see Stafford & Dewar, 2014b and references therein) and word reading (Milin, Divjak, & Baayen, 2017) hint that greater initial variability in performance of the task is linked to higher subsequent performance. Thus, these studies arrive at the conclusion that greater initial exploration results in higher task performance later in the span of learning.

Adapting the exploration and exploitation hypothesis to the present results, it is a possibility that diversity of relational knowledge reflects 'explored' conceptual knowledge that an individual has accumulated through exposure to print. In other words, the observed breadth in diversity of conceptual relations among experienced readers is possible evidence of past exploration of potential semantic candidates that have been proposed during previous encounters with compound words. As a consequence of enriching the range of semantic possibilities, more experienced readers may be rewarded later in learning by gaining greater certainty about the most likely relational interpretation for that compound. In sum, we propose that within a person's developmental trajectory, greater flexibility of relational knowledge early in the lifespan grants greater subsequent precision of that knowledge.

Furthermore, regarding the flexibility of relational knowledge, the sub-task assessing this skill required participants to select as many conceivable possible conceptual relations from an array of discrete relational categories. The motivation for this task design was to specifically assess 'productivity' of relational knowledge, i.e., the ability to form new complex concepts from conceptual subunits using a defined set of relational links. This task therefore requires that participants are provided with a representative set of possible relational phrases from which they must select as many possible interpretable meanings. As a result, participants may choose as many possible relational interpretations without assessing the semantic plausibility of each interpretation. For example, while it is possible to interpret 'sick MADE FROM sea' as a relational paraphrase of the compound *seasick*, it is semantically less plausible than 'sick CAUSED BY sea'. While the absence of a measure of semantic plausibility does not alter the interpretation of the diversity effect as an index of productivity of relational knowledge, future work could tap into this dimension of relational knowledge. For example, future studies could assess the quantity of possible relational interpretations by

also asking the separate question of the degree of the semantic plausibility of each relational interpretation. In addition, it may also be possible to ask participants to provide free-form responses of possible combinable meanings which may also provide insight into individual creativity in synthesizing semantic entities. In sum, though the present results demonstrate a key finding about the productivity of relational knowledge, assessing the plausibility and creativity of diversity of relational knowledge may provide additional information about the quality of individual semantic representations of compound words.

Relevance of findings for online compound word processing

Finally, the present findings may also inform current processing accounts of conceptual combination. The apparent counter-entropic effect of language experience on individual relational knowledge can be accommodated within the theoretical framework of the RICE theory (Spalding et al., 2010), and the more recent Compounding as an Abstract Operation in Semantic Space (CAOSS) proposal (Marelli et al., 2017). While both accounts predict that experience with language influences the availability of relational information during the processing of combined concepts, the consequences of learning on compound word processing is explicitly outlined in the CAOSS model. This model proposes that humans implicitly learn relational information about compounds from distributional patterns of language usage and automatically apply this acquired knowledge when processing new compound words. Thus, relational knowledge is learned from and is fine-tuned by exposure to compound words in a diverse set of contexts. As this knowledge becomes more entrenched, conceptual relations maybe more easily proposed and selected during online word recognition.

Further interrogation of this hypothesis may be reliant upon combining the present method of assessing individual relational knowledge of compound words with measures which tap into the same individual's online word recognition processes. Indeed, previous word recognition studies (Schmidtke et al., 2016, 2018) have shown that, at the aggregate level of the population, the relative difficulty of converging on any one relational interpretation of a compound translates into increased processing effort. The present results add to this finding by demonstrating that competition among relational candidates is susceptible to adjustment as experience with printed language increases. Thus, one may speculate that during online word processing individuals with more precise and flexible relational knowledge are able to more easily narrow down the space of possible relational meanings, and arrive at a compound meaning more quickly (especially less frequent and more transparent compounds).

To conclude, this study of compound word knowledge takes a step towards explaining how language experience determines the way that humans combine and construct new meanings from smaller conceptual units. As the capacity to produce complex meanings using conceptual relations increases with language experience, so does the precision about which candidate meaning is the most plausible.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*(2), 299.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). The CELEX database. *Nijmegen: Center for Lexical Information, Max Planck Institute for Psycholinguistics, CD-ROM*.

Baayen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. *R package version 1.0-4*.

Beech, J. (2002). Individual differences in mature readers in reading, spelling, and grapheme-phoneme conversion. *Current Psychology*, *21*(2), 121–132.

Brysbaert, M., & New, B. (2009). Moving beyond kučera 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.

Burt, J. S., & Fury, M. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing*, *13*(1), 1–30.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition*, *28*(1), 143–153.

Coolen, R., Van Jaarsveld, H. J., & schreuder, R. (1991). The interpretation of isolated novel nominal compounds. *Memory & Cognition*, *19*(4), 341–352.

Crawley, M. J. (2002). *Statistical computing: an introduction to data analysis using S-PLUS*. Chichester: Wiley Chichester.

Davies, R. A. I., Arnell, R., Birchenough, J., Grimmond, D., & Houlson, S. (2017). Reading through the life span: Individual differences in psycholinguistic effects. *Journal of Experimental*

Psychology: Learning Memory, and Cognition, Online first article: accessed September 14 2017.

Downing, P. (1977). On the creation and use of English compound nouns. *Language*, *53*(4), 810–842.

El-Bialy, R., Gagné, C. L., & Spalding, T. (2013). Processing of English compounds is sensitive to the constituents' semantic transparency. *The Mental Lexicon*, *8*(1), 75–95.

Estes, Z. (2003). Attributive and relational processes in nominal combination. *Journal of Memory and Language*, *48*(2), 304–319.

Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *41*(6), 1607–1627.

Firth, J. (1957). *Papers in linguistics*. Oxford: Oxford University Press.

Fox, J., & Weisberg, S. (2011). Car: companion to applied regression. available at: https://CRAN.R-project.org/package=car. Accessed 20th September 2017.

Gagné, C. L., & Shoben, E. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 71.

Gagné, C. L., & Spalding, T. (2004). Effect of discourse context and modifier relation frequency on conceptual combination. *Journal of Memory and Language*, *50*(4), 444–455.

Gagné, C. L., & Spalding, T. (2009a). Constituent integration during the processing of compound words: Does it involve the use of relational structures?. *Journal of Memory and Language*, *60*(1), 20–35.

Gagné, C. L., Spalding, T. L., Figueredo, L., & Mullaly, A. (2009b). Does snow man prime plastic snow? The effect of constituent position in using relational information during the interpretation of modifier-noun phrases. *The Mental Lexicon*, *4*(1), 41–76.

Gagné, C. L., & Spalding, T. L. (2014). Conceptual composition: the role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. In Ross, B. H. (Ed.) *The psychology of learning and motivation*, (pp. 97–130). San Diego: Academic Press.

Gerrig, R. J., & Murphy, G. (1992). Contextual influences on the comprehension of complex concepts. *Language and Cognitive Processes*, *7*(3-4), 205–230.

Gleitman, L. R., & Gleitman, H. (1970). *Phrase and paraphrase: Some innovative uses of language*. New York: Norton.

Gordon, P. C., Lowder, M., & Hoedemaker, R. (2016). Reading in normally aging adults. In Wright, H. (Ed.) *Cognitive-linguistic processes and aging*, (pp. 165–191). Amsterdam: John Benjamins Publishing.

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. UK: Wiley Online Library.

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, *115*(3), 367–379.

Juhasz, B. J. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In van Gompel, R. P., Fischer, M., Murray, W. S., & Hill, R. L. (Eds.) *Eye movements: A window on mind and brain*, (pp. 373–389). Amsterdam: Elsevier.

Juhasz, B. J. (2015). A database of 629 English compound words: ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior Research Methods*, *47*(4), 1004–1019.

Kuperman, V., & Van Dyke, J. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73.

Kuperman, V., & Bertram, R. (2013a). Moving spaces: Spelling alternation in English noun-noun compounds. *Language and Cognitive Processes*, *28*(7), 939–966.

Kuperman, V., & Van Dyke, J. (2013b). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 802–823.

Landauer, T. K., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Levi, J. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, *122*(3), 316.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain & Language*, *61*(1), 30–44.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain & Language*, *84*(1), 50–64.

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57–78.

Marelli, M., Gagné, C. L., & Spalding, T. (2017). Compounding as abstract operation in semantic space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition*, *166*, 207–224.

Martin-Chang, S. L., & Gould, O. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*(3), 273–284.

Milin, P., Divjak, D., & Baayen, R. (2017). A learning perspective on individual differences in skilled reading: Exploring and exploiting orthographic and semantic discrimination cues. *Journal of Experimental Psychology, Learning, Memory, and Cognition*.

Miller, G. A., & Charles, W. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, *6*(1), 1–28.

Mol, S. E., & Bus, A. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267.

Moore, M., & Gordon, P. (2015). Reading ability and print exposure: item response theory analysis of the author recognition test. *Behavior Research Methods*, *47*(4), 1095–1109.

Perfetti, C. A. (1985). *Reading ability*. Oxford: Oxford University Press.

Perfetti, C. (1992). The representation problem in reading acquisition. In Gough, P. B., Ehri, L. C., & Treiman, R.. (Eds.) *Reading acquisition*, (pp. 145–174). Hillsdale: Lawrence Erlbaum Associates.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In Vehoeve, L., Elbron, C., & Reitsma, P. (Eds.) *Precursors of functional literacy*, (pp. 189–213). Amsterdam: John Benjamins.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383.

Perfetti, C., & Adlof, S. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. In Sabatini, J. P., Albro, E., & O'Reilly, T.. (Eds.) *Measuring up: Advances in how we assess reading ability*, (pp. 3–20). Lanham: Rowman & Littlefield Education.

Pham, H., & Baayen, R. (2013). Semantic relations and compound transparency: A regression study in CARINtheory. *Psihologija*, *46*(4), 455–478.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

R Core Team (2014). R: A language and environment for statistical computing. Vienna, Austria: http://www.R-project.org/

Rahmanian, S., & Kuperman, V. (2017). Spelling errors impede recognition of correctly spelled word forms. *Scientific Studies of Reading*, 1–13.

Schmidtke, D., Kuperman, V., Gagné, C. L., & Spalding, T. (2016). Competition between conceptual relations affects compound recognition: the role of entropy. *Psychonomic Bulletin & Review*, *23*(2), 556–570.

Schmidtke, D., Gagné, C. L., Kuperman, V., Spalding, T. L., & Tucker, B. V. (2017). Conceptual relations compete during auditory and visual compound word recognition. Language, Cognition and Neuroscience: Online first. Available at: https://doi.org/10.1080/23273798.2018.1437192

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2018). Individual variability in the semantic processing of English compound words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *44*(3), 421–439.

Sears, C. R., Campbell, C. R., & Lupker, S. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1040.

Spalding, T. L., Gagné, C. L., Mullaly, A. C., & Ji, H. (2010). Relation-based interpretation of noun–noun phrases: A new theoretical approach. *Linguistische Berichte Sonderheft*, *17*, 283–315.

Spalding, T. L., & Gagné, C. (2014a). Relational diversity affects ease of processing even for opaque English compounds. *The Mental Lexicon*, *9*(1), 48–66.

Stafford, T., & Dewar, M. (2014b). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science*, *25*(2), 511–518.

Stanovich, K. E., & West, R. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402–433.

Stanovich, K. E., & Cunningham, A. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, *20*(1), 51–68.

Stanovich, K. E., West, R. F., & Harrison, M. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, *31*(5), 811.

Štekauer, P. (2005). *Meaning predictability in word formation*. Amsterdam: John Benjamins Publishing Company.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction (Vol. 1) (No. 1)*. Cambridge: MIT Press.

Taylor, J. N., & Perfetti, C. (2016). Eye movements reveal readers? Lexical quality and reading experience. *Reading and Writing*, *29*(6), 1069–1103.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Warren, B. (1978). Semantic patterns of noun–noun compounds. Acta Universitatis Gothoburgensis. *Gothenburg Studies in English Goteborg*, *41*, 1–266.

West, R. F., & Stanovich, K. (1991). The incidental acquisition of information from reading. *Psychological Science*, *2*(5), 325–330.

Wood, S. (2006). *An introduction to generalized additive models with R*. Boca Raton: Chapman.

Wood, S. (2017). Package mgcv. *R package version*, 1–7.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, *9*(3), 341–368.