The role of frequency in the association between verbs and argument structure constructions

Online Proceedings of UK-CLA Meetings
http://uk-cla.org.uk/proceedings
Vol 4: 48 – 68

ELIZABETH ANDERSON, PATRICIA COWELL AND RUTH HERBERT

Department of Human Communication Sciences, University of Sheffield, UK

ecanderson1@sheffield.ac.uk

Abstract

In Goldberg's (1995) Construction Grammar, argument structure constructions encode sentence structure: they associate syntactic forms with event-level meanings. Argument structure constructions are defined as a set of mappings between grammatical and semantic roles, independent of particular verbs. Ellis, Brook O'Donnell and Römer (2014) reported that properties of language use affected how often speakers named verbs in the intransitive construction. The authors found that verbs' lexical frequency and the frequency of verbs' occurrence in the construction were both significantly correlated to how often speakers named verbs in two verb generation tasks. The present study explored the effect of these two frequency measures on verbs produced in response to eight unique argument structure constructions. Twenty native speakers of British English took part in a verbal fluency task in which they named verbs in response to argument structure constructions. Two measures of frequency for the verbs produced in response to each argument structure construction were derived: the lexical frequency of verbs as single words, and the construction frequency of each verb in a particular argument structure construction. The number of times participants generated particular verbs in response to each construction was significantly correlated with verbs' lexical frequency and construction frequency. These results demonstrate the effect of language experience on language use. Findings are consistent with exemplar-based language representation (e.g. Bybee, 2006), where the strength of the association between verbs and argument structure constructions varies based on frequency.

Key words: argument structure, Construction Grammar, frequency effects, verbal fluency, verbs

1. Introduction

1.1 Usage-based approaches to grammar

In a usage-based approach, language structure is viewed as an emergent phenomenon that arises from domain-general capacities of human cognition (Diessel & Hilpert, 2016). This approach contrasts with theories of grammar that distinguish between speakers' implicit grammatical knowledge and their behaviour, which is considered a difference between competence and performance (e.g. Chomsky, 1965).

The frequency with which a language user encounters particular linguistic units is one of the main determinants of grammar. Frequency of occurrence drives important cognitive mechanisms that shape language structure, including exemplar learning. Exemplar theory (Bybee, 2006) states that every token of experience is classified and organised as part of a network. Incoming tokens that are identical to those already stored in memory serve to strengthen the representation of that exemplar. Incoming tokens that are similar to those already stored in memory are stored near them in the network. Exemplar theory allows specific instances of use to form part of the mental representation of an exemplar, and the collection of similar exemplars forms categories. Within a collection of related exemplars, or category, prototypical exemplars constitute central members and membership of an item in a category is graded with respect to its similarity to other items in the category (Goldberg, 2006).

1.2 Construction Grammar

The linguistic units that arise from exemplar learning constitute constructions, which are language representations that arise from generalisation across exemplars. Constructionist approaches to grammar are built on the notion that grammar comprises an inventory of linguistic constructions. A construction is any unique pairing of a linguistic form and particular meaning (Goldberg, 2003; Kay & Fillmore, 1999). A variety of constructionist approaches have been articulated over the past three decades (e.g. Croft, 2001; Goldberg, 2006; Langacker, 1987; Stefanowitsch & Gries, 2003). All constructionist approaches agree on a number of principles (Langacker, 2005), including a commitment to the construction as the primary object of linguistic description, and recognition that the lexicon and grammar form a continuum of constructions.

The continuum between the lexicon and grammar can be seen in how constructions vary in their complexity, abstraction and lexical specification (Goldberg, 2013). At one end of the spectrum, constructions like morphemes, such as the plural *-s* in English, and lexemes, like the word *cat*, are fully lexically-specified. Idioms, like *let the cat out of the bag*, are fully lexically-specified but more complex than morphemes or lexemes. At the other end of the spectrum, constructions like the ditransitive, as in *the cook handed the chef the scissors*, or the passive, as in *the scissors were handed to the chef by the cook*, are abstract and can occur with a variety of lexical items. All these structures can be considered constructions because they pair a specific linguistic form with a particular meaning.

Goldberg (1995) identified argument structure constructions as the means for sentence expression in Construction Grammar. Argument structure constructions encode event-level meaning by specifying the relationship between syntactic phrases and their semantic roles. For example, the caused motion construction can be defined as a linguistic form that contains a subject, verb, object and oblique. The construction refers to an event in which an agent causes a patient to move along a path. The sentences in (1) below are both instances of the caused motion construction, despite containing different verbs and a different set of lexemes (Goldberg, 1995, p. 162).

(1a) Sam helped him into the car
(1b) Sam guided him through the terrain

The sentences in (1) both refer to an event in which an agent causes an object to move along a path. The sentences have the same syntactic form, and therefore the caused motion can be considered a construction, because it pairs a syntactic form with event-level meaning.

The meaning of an utterance arises from the combination of all the constructions it contains, at all levels of complexity. In Goldberg's (2003) words, 'it's constructions all the way down' (p. 223). The interpretation of a sentence is built up from the combination of constructions it contains, including morphemes, lexemes and its argument structure. The influence of the meaning of an argument structure construction on the interpretation of sentence was observed by Kako (2006, Experiment 5). Readers rated the subject and object of sentences in terms of semantic properties usually associated with agents and patients. When responding to intransitive verbs presented in the transitive argument structure construction, they rated the subjects as agent-like and objects as patient-like, showing that they interpreted the intransitive verbs as more transitive-like when they appeared in the transitive construction.

In addition to referring to a unique form-meaning pairing, constructions can be defined if they are highly frequent linguistic patterns (Goldberg, 2003). Two investigations identified linguistic units that qualify as constructions for this reason. Bybee and Schiebman (1999) measured the amount of phonetic reduction in the articulation of *don't* in various contexts. They observed the greatest degree of phonetic reduction in contexts where *don't* was most frequent. Vogel Sosa and MacFarlane (2002) used a word-monitoring paradigm and found that participants took longer to identify the word *of* in two-word contexts where it was most frequent. All four authors concluded that the observed frequency effects demonstrate that the phrases under consideration represent entire processing units. Such a basic unit of language can be termed a construction. Construction Grammar recognises constructions as meaningful or highly frequent patterns in language.

1.3 Frequency effects in language processing

Because frequency of occurrence influences the processes that give rise to the shape of grammar, it is unsurprising that frequency effects also have also been identified in online language processing. It has long been known that speakers process single words with high frequency more quickly than single words with low frequency, as measured by picture naming (Oldfield & Wingfield, 1965), word reading (Forster & Chambers, 1973) and spoken word duration (Wright, 1979). Recent research has investigated frequency effects in constructions more complex than the single word.

Frequency effects at this level have been attested in language production and comprehension. Janssen and Barber (2012) examined the production of multi-word noun phrases. They found that speakers produced high frequency multi-word phrases more quickly than phrases with lower frequency. Arnon and Snider (2010) explored the effect of frequency on input processing. They

asked participants to decide whether four-word phrases were allowable sentences, in the manner of a lexical decision task. Phrases were minimal pairs that differed in the final word, such as *don't have to worry* and *don't have to wait*. The phrases differed in surface frequency but were matched for the frequency of the final word, bigram and trigram. Participants' response times were faster to high frequency phrases than low frequency phrases. This research demonstrates that frequency of occurrence affects the processing of multi-word constructions.

Verb biases result when frequency effects occur with linguistic units greater than the single word. Verb bias reflects the frequency of a verb within a particular context: given a particular verb, the verb bias is a frequency-based effect of the surrounding syntactic context. Verb bias is taken to be a lexical property of verbs that arises from frequent exposure to verbs in certain syntactic contexts over others (Lapata, Keller & Schulte im Walde, 2001). For example, the verb *remember* can occur in the two syntactic contexts in (2).

(2a) The teacher remembered
(2b) The teacher remembered to talk

Connine, Ferreira, Jones, Clifton and Frazier (1984) found that writers were most likely to produce sentences containing the verb *remember* in a sentence with a structure like (2b) than (2a), showing that *remember* has a bias for occurring with an infinitive complement. Similarly, other verbs are more likely to precede a direct object or embedded clause (Garnsey, Pearlmutter, Myers & Lotocky, 1997). Verb biases affect language processing, as readers' interpretations of sentences are influenced by how often a particular verb occurs in the syntactic context. Generally, reading times are longer when verbs appear in a syntactic context that conflicts with their bias (Garnsey et al., 1997; Trueswell & Kim, 1998; Trueswell, Tanenhaus & Kello, 1993; Wilson & Garnsey, 2009).

In addition to verb biases, language users are sensitive to the frequency of verbs given a particular syntactic context. Ellis, Brook O'Donnell and Römer (2014) explored usage-based properties of verbs that participants named in response to instances of the intransitive motion construction. In their first experiment, participants produced a single word to complete sentence frames like '*he ___ across the…*'. In their second experiment, participants were given one minute to generate multiple words to complete the sentence frame in a computerised verbal fluency task. The authors reported significant effects of three usage-based properties on how often participants named verbs in response to the constructions. These effects included the frequency of the verb in each sentence frame, the faithfulness of a verb to the form of the particular sentence frame and the semantic prototypicality of the verb in the sentence frame. The authors also reported significant correlations between the number of times verbs were generated in response to a sentence frame and verbs' lexical frequency.

The present study builds on the findings of Ellis et al. (2014) and examines the effects of the lexical frequency of verbs and verbs' frequency in particular syntactic contexts on verb retrieval. We employed a spoken verbal fluency task and asked participants to respond to eight unique syntactic constructions

identified by Goldberg (1995; Johnson & Goldberg, 2013). In this study, we aimed:

- to explore the association between verbs and argument structure constructions in terms of the frequency of verbs; and
- to examine differences in verb frequency among unique argument structure constructions.

2. Method

2.1 Participants

Twenty native adult speakers of British English took part in the study. The group included ten men and ten women. Participants reported no history of speech or language difficulties or psychiatric condition. All participants reported normal or corrected-to-normal vision and hearing. Because this group served as the control in a related investigation of language in acquired aphasia (Anderson, Herbert & Cowell, 2015), participants were aged between fifty and eighty years. Background characteristics of participants are shown in Table 1.

Table 1

Age and education of participants

	M	*SD*	minimum	maximum
Age in years	62.9	6.4	52	74
Years in education	16.3	3.6	11	23

Note. Age in years (decimal point); years in full-time education, including primary, secondary and higher education.

2.2 Ethics, recruitment and consent

This study received ethical approval from the University Research Ethics Committee in the Department of Human Communication Sciences at the University of Sheffield. Participants were recruited via mailing lists to members of the university and community groups for individuals over the age of fifty. Participants gave written informed consent before taking part in the study.

2.3 Materials

Eight argument structure constructions were included in the study. Goldberg (1995; Johnson & Goldberg, 2013) identified these constructions as independent linguistic units in which a variety of verbs can occur. The constructions were presented to participants as sentences in which noun phrases were encoded as pronouns. Therefore, the meaning of each stimulus sentence can be attributed to the event-level meaning of the argument structure construction, rather than to any activation of lexical semantics from

the words contained in each sentence. A blank space stood in place of the verb. The constructions, their meanings and the stimulus sentences included in the study for each construction are shown in Table 2.

Table 2

Argument structure constructions included in materials

Construction	Semantics	Stimulus sentences
Caused motion	X causes Y to move to Z	You __ it to us I __ it over there
Conative	X directs motion at Y	You __ at us They __ at it
Ditransitive	X causes Y to receive Z	They __ us some things I __ you something
Intransitive motion	X moves to/from Y	We __ through there It __ through there
Passive	X is acted on by Y	We were __ by them It was __ by them
Removal	X causes Y to move from Z	I __ it from you You __ it from there
Transitive	X acts on Y	We __ them You __ it
Way	X moves through path	They __ their way to it We __ our way there

Two versions of each of the eight constructions were included in the task, giving a total of sixteen stimuli. Stimuli were presented in a pseudo-random order following the constraint that stimulus sentences that instantiated the same construction were separated by at least three other constructions. Two list orders were created and assigned randomly to participants, with half of the participants experiencing each list order.

A written and auditory version of each stimulus sentence was created. Written sentences were presented individually in black, bold, point-42 Helvetica type centred on A5-sized paper. The audio version of the sentences consisted of a female native speaker of British English recording the sentences on a Marantz PMD670 solid-state audio recorder, recording at stereo pulse code modulation with a sample frequency of 22.1 KHz via a Sennheiser MD425 microphone. Five hundred milliseconds of white noise with an amplitude of 0.01 was inserted on the software Audacity to represent the missing verb in the auditory version of each sentence.

2.4 Procedure

Participants met with one researcher in a quiet room in a university building. The verbal fluency task began with one example item and two practice items. On example and practice items, participants were asked to generate verbs in response to sentences in which blank spaces occurred after modal verbs or infinitival *to*, such as *we can* ___ and *they want to* ___. This ensured that over the course of the entire verbal fluency task participants were exposed to the same number of each of the eight argument structure constructions of interest.

On each trial, participants were presented with the written form of a sentence stimulus. At the same time, the recording of the sentence was played from a laptop computer through USB speakers. The researcher prompted participants to respond by saying 'please say as many single words as you can think of that can fit in the sentence'. Participants were given thirty seconds to name words. Each trial ended when a timer sounded and the researcher said 'please stop'.

Participants' responses were transcribed in situ, and each session was audio recorded on a Marantz PMD670 solid-state audio recorder, recording at stereo pulse code modulation with a sample frequency of 22.1 KHz via a Sennheiser MD425 microphone. Written transcriptions were checked against the audio recording by the first author to ensure that transcriptions were accurate.

2.5 Data analysis

2.5.1 Response coding

Participants' responses were coded in one of five categories. The response codes, their definitions and an example of each type of response are shown in Table 3. Only real verbs and rejections, which were verbs produced but immediately rejected by participants, were entered into the dataset for analysis.

Verb type refers to a unique verb lemma. For example, the responses *run* and *ran* are different forms of the same verb type, the verb *run*.

Table 3

Response codes for analysis of verbal fluency task

Response code	Definition	Example
Real verb	Any form of a real English verb	give, gave, given
Rejection	Response was rejected by participant	like - no sorry
Repetition	Verb type was previously produced in response to same stimulus sentence	*run* (real verb), *ran* (repetition)
Non-verb	Real words of English that were	late, dark, often

	not verbs	
Multiple words	Multiple words corresponded to the blank space in stimulus sentence	fed up, running at

2.5.2 Inter- and intra-rater reliability

The entire dataset was transcribed and coded by the first author. Additionally, two randomly selected trials from each participant were subject to inter- and intra-rater reliability testing of transcription and coding. This equated to 12.5% of the entire dataset. Two different subsets of trials were subject to inter- and intra-rater testing.

Inter-rater reliability judgements were performed by a native speaker of British English who held a university degree in English language. Agreement on the transcriptions was 95%. Coding agreement was 99%, which corresponded to a Kappa value of $\kappa = 0.92$ (Cohen, 1960). This value indicates substantial agreement between the two coders (Landis & Koch, 1977).

Intra-rater reliability judgements were performed eight weeks following the original. Agreement on the transcriptions was 99%, and agreement on the coding was 100%, which corresponded to a Kappa value of $\kappa = 1.00$.

2.5.3 Frequency values for verbs produced in verbal fluency task

For each verb in the dataset, two frequency measures were extracted from the British National Corpus (Davies, 2004-; Leech, Rayson & Wilson, 2001). The lexical frequency of a verb was the frequency of the verb lemma, in instances per million, as listed in the British National Corpus (Leech et al., 2001). The construction frequency of a verb was the raw frequency of a verb occurring in the argument structure construction to which it was produced. To derive construction frequency values, each argument structure construction was defined as a query using the UCREL tagset (UCREL, 1993) and searched on Brigham Young University's interface to the British National Corpus (Davies, 2004-). Results were manually filtered to ensure that they matched the target construction. The search queries for each argument structure construction are provided in the Appendix.

The dependent variable was the total number of times participants generated verbs in response to an argument structure construction. The maximum possible value for each verb was 40, which equated to each of the 20 participants in the study producing the verb in response to each of the two stimuli versions of an argument structure construction.

2.5.4 Statistical analysis

For each construction, we examined the correlations between the number of times all participants generated a verb in the verbal fluency task, and the verb's lexical frequency and construction frequency. We used Baguley's (2012) implementation of Zou's (2007) method for investigating the difference

between these two correlations. Zou (2007) advanced a method for calculating the significance between two dependent overlapping correlations, where two correlations share a common variable. Zou's method is based on the construction of confidence intervals, and significance can be determined when confidence intervals do not contain zero.

3. Results

3.1 Summary of participants' responses

The majority of participants' responses were coded as real verbs. Repetitions accounted for a mean of 4% of participants' responses, followed by non-verb responses, multiple word responses and rejections. The proportion of participants' responses that were coded as each category is presented in Table 4.

Table 4

Proportion of participants' responses for each response code

	M	*SD*	minimum	maximum
Real verb	.94	.04	.86	1.00
Rejection	.004	.008	0	.02
Repetition	.04	.03	0	.14
Non-verb	.01	.02	0	.09
Multiple words	.003	.006	0	.02

With only one exception, all participants named verbs in response to each stimulus sentence. On one trial, one participant was unable to name any verbs in response to a stimulus sentence, which was an instance of the passive construction. However, the participant succeeded in naming verbs in response to the other version of the passive construction included in the verbal fluency task.

Table 5 shows the number of verbs that participants produced in response to each construction. The total number is the number of verbs all participants produced, and the mean, standard deviation minimum and maximum refer to the number of verbs produced per participant.

Table 5

Number of verbs produced in response to each construction

	Total	*M*	*SD*	minimum	maximum
Caused motion	105	12	3	6	18
Conative	71	8	3	4	14

Ditransitive	64	9	2	4	16
Intransitive motion	106	13	3	8	19
Passive	171	14	6	3	23
Removal	72	9	2	5	13
Transitive	164	17	5	9	27
Way	85	9	4	1	18

3.2 Verb generation to constructions

Figure 1 shows the frequency distributions of the number of times participants generated verbs in response to each of the eight argument structure constructions. Each bar in the graphs represents one verb. The y-axis shows the frequency that the verb was produced by all participants, which has a maximum value of 40. Labels on the x-axis are collapsed; for example, the frequency distribution for the caused motion construction shows results for 105 verbs, but only 27 are labelled on the x-axis. The eight distributions are all unimodal and heavily skewed to the right.

For all eight constructions, a few verb types accounted for a large proportion of participants' responses, and a large number of verb types were produced only once. For example, participants generated a total of 72 unique verb types in response to the removal construction. The five most frequently generated verb types accounted for 44% of responses, and 41 verb types were produced only once.

3.3 Correlations between verb generation, lexical and construction frequency

Spearman's correlations between the number of times participants generated verbs in response to each construction in the verbal fluency task, and verbs' lexical and construction frequencies are shown in Table 6. A significant difference between the strength of the correlations to lexical and construction frequency was determined following Zou (2007).

For six of the eight constructions, the correlation to construction frequency was stronger than the correlation to lexical frequency. Only the caused motion construction and the passive construction did not exhibit this pattern. For four constructions, the difference in the strength of the correlation between the number of times participants generated verbs in the verbal fluency task and verbs' lexical and construction frequency reached significance. For the ditransitive, intransitive motion, removal and transitive constructions, the correlation to construction frequency was significantly greater than the correlation to lexical frequency.

3.4 Further investigation of the caused motion and passive constructions

Unlike the other argument structure constructions included in the verbal fluency task, the caused motion construction and passive construction did not show a stronger relationship between the number of times participants generated verbs in the verbal fluency task and verbs' construction frequency rather than lexical frequency. In order to further explore this finding, Spearman's correlations were calculated separately for each of the two stimulus sentences included in the verbal fluency task for the constructions. The results of this investigation are show in Table 7.

Figure 1

Frequency distributions of the number of times participants generated verbs in response to each argument structure construction

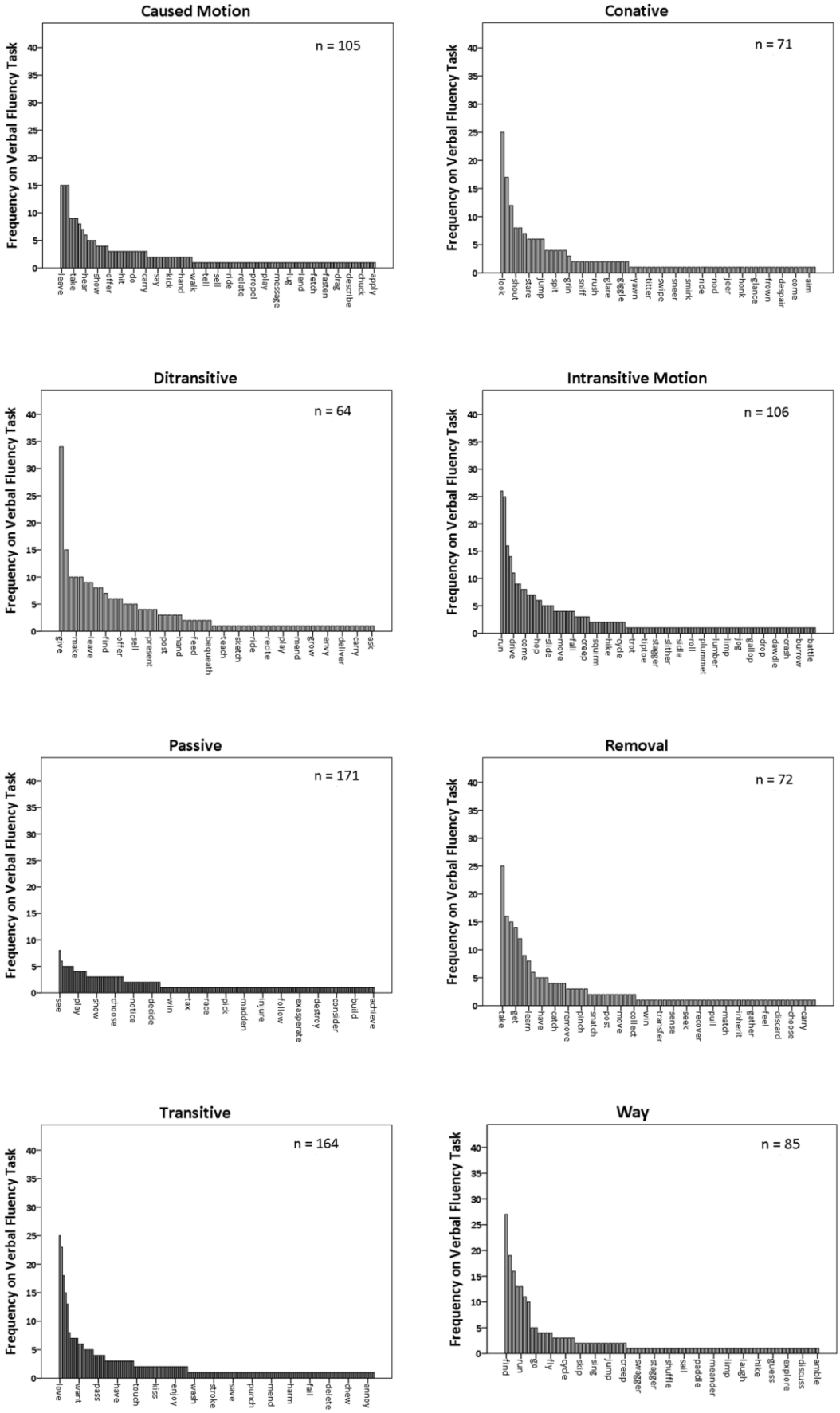

Table 6

Correlations between number of times verbs were generated in the verbal fluency task and verbs' lexical frequency and construction frequency

	n	Lexical frequency	Construction frequency	Significant difference between lexical and construction frequency
Caused motion	105	0.46***	0.42***	
Conative	71	0.29**	0.49***	
Ditransitive	64	0.32**	0.62***	•
Intransitive motion	106	0.29***	0.51***	•
Passive	171	0.40***	0.29***	
Removal	72	0.23*	0.57***	•
Transitive	164	0.39***	0.54***	•
Way	85	0.28**	0.34***	

*$p \leq 0.05$. **$p \leq 0.01$. ***$p \leq 0.001$.

Table 7

Correlations between number of times verbs generated on verbal fluency task and verbs' lexical frequency and construction frequency for versions of the caused motion and passive constructions

	n	Lexical frequency	Construction frequency	Significant difference between lexical and construction frequency
		Caused motion		
You __ it to us	47	0.35**	0.50***	
I __ it over there	71	0.54***	0.28**	•
		Passive		
It was __ by them	95	0.48***	0.37***	
We were __ by them	97	0.17	0.15	

*$p \leq 0.05$. **$p \leq 0.01$. ***$p \leq 0.001$.

The correlation between the number of times participants generated verbs in response to the caused motion stimulus sentence containing the preposition

phrase *to us* and verbs' construction frequency was stronger than the correlation for lexical frequency. This difference did not reach significance. In contrast, the number of times participants generated verbs in response to the caused motion stimulus sentence containing the preposition phrase *over there* and verbs' lexical frequency was significantly stronger than the correlation for construction frequency.

The number of times participants generated verbs in response to the passive stimulus sentence in which the grammatical subject was encoded by the pronoun *it* was significantly correlated with verbs' lexical frequency and construction frequency. There was no significant difference between the strength of these correlations. In contrast, there was no relationship between the number of times participants generated verbs in response to the passive stimulus sentence in which the grammatical subject was encoded by the pronoun *we* and verbs' lexical frequency or construction frequency.

4. Discussion

This study employed a verbal fluency task to investigate the effect of lexical frequency and construction frequency on the association between verbs and argument structure constructions. There were significant correlations of lexical and construction frequency for the eight argument structure constructions included in the study. For six of the eight argument structure constructions, the effect of construction frequency was greater than the effect of lexical frequency, and this difference reached significance for four of the six constructions.

The findings from this study replicate the results from the study by Ellis et al. (2014), who reported the significant effect of lexical frequency and construction frequency on the number of times participants generated verbs in response to versions of the intransitive motion construction. Ellis et al. observed a greater effect of construction frequency in most instances. The present study found that construction frequency had a significantly greater effect than lexical frequency on participants' responses to the intransitive motion construction.

The present work further establishes the primacy of construction frequency over lexical frequency on the processing of five other argument structure constructions in addition to the intransitive motion construction, including the conative, ditransitive, removal, transitive and way constructions. For these constructions, construction frequency had a greater effect than lexical frequency on participants' verb responses.

Participants responded to argument structure constructions as meaningful units of language, despite the fact stimulus sentences were devoid of lexical semantic content. The effect of construction frequency on verb generation suggests not only that frequency affects language processing, but language users are sensitive to the context in which lexemes occur, for a wide range of structures. Findings will be discussed in this section with reference to the semantics of argument structure constructions and their lexical specification.

4.1 Argument structure constructions are semantic categories

The frequency distributions of the number of times participants generated verbs in response to individual argument structure constructions (Figure 1) can be described as Zipfian (Zipf, 1935). Zipf (1935) advanced a power law that related a linguistic unit's frequency of occurrence and frequency rank. Zipfian distributions - or near-Zipfian, in cases that approximate this situation - can be described as those in which few high frequency items account for most of the tokens in a set of linguistic units, and many items occur with a very low frequency. This type of distribution is a property of natural language, and many forms of linguistic units have been observed to follow this pattern, including terms for animals (Henley, 1969), number words (Dehaene & Mehler, 1992), two-word phrases (Williams et al., 2015) and pairs of phonemes (Siew, 2013).

The fact that natural language phenomena pattern in this way is a nontrivial aspect of language (Piantadosi, 2015). When considering why verbs named in response to argument structure constructions should display this distribution, we follow Bybee's (2010) reasoning that schematic components of constructions result from the process of categorisation. Central members of categories are highly frequent exemplars that serve as prototypes. Many early investigations of semantic memory explored category membership, for semantic categories such as types of birds, units of time, tools and occupations. Verbal fluency tasks, like that used in the present study, were used to identify members of the category, and the most frequently named exemplars were taken to represent the most prototypical members of the category (Chang, 1986). For example, Battig and Montague (1969) employed a written verbal fluency task with a time limit of thirty seconds to obtain norms for 56 semantic categories. Results from this type of task were interpreted as indicating the strength of the semantic relationship between a category and an exemplar of that category (Loftus, 1973; Smith, Shoben & Rips, 1974).

In the current study, the most frequently named verbs can be considered central members of argument structure construction categories. The idea that linguistic units such as argument structure constructions form categories supports two assumptions of the Construction Grammar framework and related usage-based approaches to language: first, that argument structure constructions are meaningful units of language, and, second, that language users apply the domain-general processing mechanisms to syntax.

4.2 Lexical specification of argument structure constructions

For six of the eight argument structure constructions included in this study, participants' verb responses showed a greater effect of construction frequency than lexical frequency. Four of these six constructions were defined in part by specific lexical items (see Table 2). The conative construction includes a preposition phrase headed by the lexeme at, the removal construction by a preposition phrase headed by the lexeme from and the way construction by an object noun phrase consisting of a possessive pronoun and the noun way. In this study, the form of the intransitive motion construction was defined by a preposition phrase headed by the lexeme through. Therefore, it remains

possible that the frequency effects related to the argument structure constructions included in this study may be attributable to the bigram frequency of a verb and the following preposition or noun phrase, rather than the more complex argument structure construction containing a subject, verb and other phrases.

This possibility is corroborated by findings on the caused motion construction. The form of the caused motion construction that contained the preposition phrase headed by the lexeme to showed a greater effect of construction than lexical frequency, like the constructions mentioned above. However, the form of the caused motion construction containing the preposition over did not. The different effect of construction frequency between the two forms of the caused motion construction suggest that the forms are in fact two distinct constructions, and the identity of the lexemes contained in the stimulus sentences played a defining role in the construction.

Recall from Section 1.2 that both examples in (1), repeated below, are instances of the caused motion construction.

(1a) Sam helped him into the car
(1b) Sam guided him through the terrain

These two sentences are similar at the structural level in that they both contain a subject, verb, object and oblique. This structure is abstract, because no particular lexical item is shared between the sentences. However, this abstract level of similarity does not recognise the crucial contribution of the individual verbs to the sentences' meanings. To illustrate, the verb *hear* can occur grammatically in sentence (1b), but not in sentence (1a). However, the semantics of the caused motion construction is not evident in the sentence *Sam heard him through the terrain*. In this event, no movement occurs: rather, the preposition phrase encodes a barrier, rather than the trajectory of movement, as in the caused motion construction. This distinction remains despite the status of the preposition phrase as argument or adjunct, as in both instances the preposition phrase qualifies as an adjunct (Needham & Toivonen, 2011).

We suggest that in the current study a similar dissimilarity existed for verbs and prepositions included in the two stimulus sentences for the caused motion construction. While the preposition phrase *to us* encoded the movement associated with the semantics of the caused motion construction, the preposition phrase *over there* did not, because it is polysemous between encoding trajectory of movement and static locations. The sentences in (3) both contain the preposition phrase *over there* in the syntactic form of the caused motion construction, but only (3a) can refer to an event which pertains to movement.

(3a) The athlete kicked the ball over there
(3b) The professor saw the paper over there

The frequency of the combinations between verbs and prepositions may explain the effect of construction frequency observed in this study. Indeed, Ellis et al. (2014) defined verb-argument constructions as the combination of

a verb and subsequent preposition; constructions in their study differed in the identity of the preposition. In contrast, the current study took abstract, sentence-level structures as the starting point for investigation of verbs and argument structure. This discussion raises the more general question of how abstract grammar can be. Previous research on verb behaviour suggests that some syntactic frequency effects are lexically independent. For example, Gries and Stefanowitsch (2004) identified verbs that are strongly associated with the ditransitive construction, containing two post-verbal arguments, and the transitive construction, containing one post-verbal argument. Current findings suggest that lexemes play a pivotal role in the interpretation of sentence structure. The nature of the relationship between the contribution of lexemes to more complex constructions is a direction for future research.

A final point on the role of lexemes in the interpretation of argument structure constructions concerns the findings from the two forms of the passive construction. Construction frequency proved to be significantly related to verbs that participants generated in response to the passive sentence containing the subject pronoun *it*, but not the subject pronoun *we*. Like the caused motion construction, the lexemes contained in the stimulus sentences influenced the verbs that were generated in response. In the case of the passive construction, two accounts may contribute to an explanation of this finding. First, the bigram frequency between verbs and the subject pronoun contained in the stimulus sentence may have had an effect on the verbs that participants generated. Second, the passive has a discourse function (Goldberg, 2003), rather than a specific event-level meaning like the other argument structure constructions included in the study. In this instance, the lack of discourse context may have influenced participants' responses.

5. Conclusion

This study demonstrated that frequency plays a major role in language processing at multiple levels of language structure, including the lexical frequency of single words, and the frequency with which single words combine with larger structures, such as argument structure constructions. Results are compatible with usage-based approaches to language that recognise frequency as a driving force behind both the shape of grammar and human language processing. The work raises the question of whether linguistic constructions can be defined as structures that are entirely devoid of lexical content, and queries the relationship between argument structure constructions and the lexical items they contain.

Appendix

Table A1 shows the search queries that were used to derive values of construction frequency from the British National Corpus.

Table A1

Construction	Corpus query	Gloss
Caused motion	[vv*] [pp*] to|over -[v*]	lexical verb - pronoun - *to* or *over* - not a verb
Conative	[vv*] at [pp*]	lexical verb - *at* - pronoun
Ditransitive	[vv*] [pp*] [d*] [nn*]	lexical verb - pronoun - determiner - noun
Intransitive motion	[vv*] through	lexical verb - *through*
Passive	[vb*] [vvn] by	form of *be* - past participle of lexical verb - *by*
Removal	[vv*] [pp*] from	lexical verb - pronoun - *from*
Transitive	[vv*] [pp*] .|,|;	lexical verb - pronoun - clause-final punctuation
Way	[vv*] [appge] way [i*]	lexical verb - possessive pronoun - *way* - preposition

For each verb in the dataset, the verb lemma was specified in the query for the construction to which the verb was produced. For example, the verb *leave* was produced in response to the caused motion construction. In order to ascertain the construction frequency of the verb *leave* in the caused motion construction, Brigham Young University's interface to the British National Corpus (Davies, 2004-) was searched with the query shown in (A1).

(A1) [leave].[vv*] [pp*] to|over -[v*]

To summarise, this search returned text strings in the corpus that contained any form of the verb *leave* tagged as a lexical verb, followed by a pronoun, followed by the word *to* or *over* and finally by any word that was not tagged as a verb. The first author reviewed results to the query to ensure that they were genuine instances of the caused motion construction. This process was repeated for all 105 verbs that participants produced in response to the caused motion construction.

A similar process was undertaken to derive construction frequency values for all verbs produced in response to the remaining seven argument structure constructions.

References

Anderson, E., Herbert, R., & Cowell, P. (2015). *Argument structure constructions elicit verbs from speakers with aphasia*. Paper presented at the British Aphasiology Society Biennial International Conference, University College London.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67-82.

Audacity (version 2.0.3) [Computer software]. Retrieved from http://audacity.sourceforge.net/

Baguley, T. (2012). Comparing correlations: independent or dependent (overlapping or non-overlapping). Retrieved from https://seriousstats.wordpress.com/2012/02/05/comparing-correlations/

Battig, W., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology, 80*(3), 1-46.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language, 82*(4), 711-733.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: the reducation of *don't* in English. *Linguistics, 37*(4), 575-596.

Chang, T. M. (1986). Semantic memory: facts and models. *Psychological Bulletin, 99*(2), 199-220.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Connine, C., Ferreira, F., Jones, C., Clifton, C., & Frazier, L. (1984). Verb frame preferences: descriptive norms. *Journal of Psycholinguistic Research, 13*(4), 307-319.

Croft, W. (2001). *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Davies, M. (2004-). BYU-BNC. (Based on the British National Corpus from Oxford University Press.), from http://corpus.byu.edu/bnc

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43*(1), 1-29.

Diessel, H., & Hilpert, M. (2016). Frequency effects in grammar. In M. Aronoff (Ed.), Oxford Research Encyclopedias: Linguistics Oxford University Press. Retrieved from http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-120. doi: DOI: 10.1093/acrefore/9780199384655.013.120

Ellis, N., O'Donnell, M., & Römer, U. (2014). The processing of verb-argument constructions is sensitive to form, function, frequency, contingency and prototypicality. *Cognitive Linguistics, 25*(1), 55-98.

Forster, K., & Chambers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627-635.

Garnsey, S., Pearlmutter, N., Myers, E., & Lotocky, M. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58-93.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Goldberg, A. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences, 7*(5), 219-224.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, A. (2013). Argument structure constructions versus lexical rules or derivational verb templates. *Mind and Language, 28*(4), 435-465.

Henley, N. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*, 176-184.

Janssen, N., & Barber, H. (2012). Phrase frequency effects in language production. *PLoS One, 7*(3), e33202. doi: 10.1371/journal.pone.0033202

Johnson, M., & Goldberg, A. (2013). Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes, 28*(10), 1439-1452.

Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition, 101*, 1-42.

Kay, P., & Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language, 75*(1), 1-33.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Langacker, R. (1987). *Foundations of Cognitive Grammar* (Vol. 1). Stanford: Stanford University Press.

Langacker, R. (2005). Integration, grammaticization, and constructional meaning. In M. Fried & H. C. Boas (Eds.), *Grammatical Constructions: Back to the roots* (Vol. 4, pp. 157-189). Amsterdam: John Benjamins Publishing Company.

Lapata, M., Keller, F., & Schulte im Walde, S. (2001). Verb frame frequency as a predictor of verb bias. *Journal of Psycholinguistic Research, 30*(4), 419-435.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: based on the British National Corpus*. Abingdon: Routedge

Loftus, E. (1973). Category dominance, instance dominance, and categorization of time. *Journal of Experimental Psychology, 97*(1), 70-74.

Needham, S., & Toinonen, I. (2011). *Derived Arguments*. Paper presented at the LGF11 Conference, University of Hong Kong.

Oldfield, R., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*(4), 273-281.

Piantadosi, S. (2014). Zipf's word frequency law in language language: A critical review and future directions. *Psychonomic Bulletin and Review, 21*, 1112-1130.

Siew, C. (2013). Community structure in the phonological network. *Frontiers in Psychology, 4*(Aug), Article 553.

Smith, E., Shoben, E., & Rips, L. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*(3), 214-241.

Stefanowitsch, A., & Gries, S. (2003). Collostructions: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics, 8*(2), 209-243.

Trueswell, J., & Kim, A. (1998). How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. *Journal of Memory and Language, 39*, 102-123.

Trueswell, J., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 528-553.

UCREL. (1993). UCREL CLAWS7 Tagset. Retrieved from http://ucrel.lancs.ac.uk/claws7tag.html

Vogel Sosa, A., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: collocations involving the word *of*. *Brain and Language, 83*, 227-236.

Williams, J., Lessard, P., Desu, S., Clark, E., Bagrow, J., Danforth, C., & Dodds, P. (2015). Zipf's law holds for phrases, not words. *Scientific Reports, 5*, Article number 12209.

Wilson, M., & Garnsey, S. (2009). Making simple sentences hard: verb bias effects in simple direct object sentences. *Journal of Memory and Language, 60*, 368-392.

Wright, C. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory and Cognition, 7*(6), 411-419.

Zipf, G. (1935). *The psychobiology of language*. Boston: Houghton Mifflin Co.

Zou, G. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12*(4), 399-413.