Applied Psycholinguistics, page 1 of 22, 2012
doi:10.1017/S0142716412000367

Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive

JAMES A. STREET and EWA DĄBROWSKA
Northumbria University

Received: September 27, 2010 Accepted for publication: December 22, 2011

ADDRESS FOR CORRESPONDENCE
James A. Street, Department of Humanities, School of Arts and Social Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK. E-mail: james.street@northumbria.ac.uk

ABSTRACT
This article provides experimental evidence for the role of lexically specific representations in the processing of passive sentences and considerable education-related differences in comprehension of the passive construction. The experiment measured response time and decision accuracy of participants with high and low academic attainment using an online task that compared processing and comprehension of active and passive sentences containing verbs strongly associated with the passive and active constructions, as determined by collostructional analysis. As predicted by usage-based accounts, participants' performance was influenced by frequency (both groups processed actives faster than passives; the low academic attainment participants also made significantly more errors on passive sentences) and lexical specificity (i.e., processing of passives was slower with verbs strongly associated with the active). Contra to proposals made by Dąbrowska and Street (2006), the results suggest that all participants have verb-specific as well as verb-general representations, but that the latter are not as entrenched in the participants with low academic attainment, resulting in less reliable performance. The results also show no evidence of a speed–accuracy trade-off, making alternative accounts of the results (e.g., those of two-stage processing models, such as Townsend & Bever, 2001) problematic.

Most linguists implicitly or explicitly assume that all first language learners converge on the same grammar (see, e.g., Birdsong, 2004, p. 83; Bley-Vroman, 2009, p. 179; Chomsky, 1975, p. 11; Crain & Lillo-Martin, 1999, p. 9; Lidz & Williams, 2009, p. 177; Nowak, Komarova, & Niyogi, 2001, p. 114; Seidenberg, 1997, p. 1600). Although it is well established that there are large individual differences in lexical knowledge and knowledge of archaic, very formal, or highly literary grammatical constructions (e.g., *Little did I know that . . .*), all speakers are believed to share the same "core" grammar. Linguists who assume convergence rarely justify it in any way: it is regarded as a self-evident, incontrovertible "fact" that linguistic theory must explain. Along with poverty of the stimulus, convergence is regarded as an incontestable argument for an innate universal grammar: if different learners are exposed to different input yet acquire essentially the same grammar, the construction of grammar must be guided by highly restrictive principles (cf. Chomsky, 1975; Lidz & Williams, 2009). Conversely, that second language learners typically do not converge has been used to argue that L2 learning is fundamentally different from first language acquisition (cf. Bley-Vroman, 1990, 2009).

However, there is a growing body of evidence that suggests there are considerable differences in native speakers' mastery of some basic grammatical constructions. Dąbrowska (2008b), for example, reports that different groups of Polish speakers appear to have learned different generalizations about the distribution of genitive masculine inflections: some rely on very general semantic cues (animacy), others use more specific semantic information (e.g., substance vs. object), and still others are sensitive primarily to phonological properties of the stem. These differences are not explainable in terms of regional or social variation, although they are not, perhaps, very surprising, as this part of the inflection system is highly irregular, and therefore it is not clear what the "correct" generalization would be. However, Dąbrowska (2008a) found large individual differences with the Polish dative inflection, which is almost completely regular. In a nonce word inflection experiment, participants were asked to supply the dative form of masculine, feminine, and neuter nouns. Within each gender, half the nouns belonged to densely populated neighborhoods (i.e., they resembled many existing nouns) and half to sparsely populated neighborhoods. Dąbrowska found that participants reliably supplied the target forms of nonce nouns belonging to large classes and densely populated neighborhoods, indicating that they were cooperative and had understood the task. However, results for nouns belonging to smaller classes and/or sparsely populated neighborhoods were much more varied, with individual scores ranging from 29% to 100% (4%–100% for nouns from low-density neighborhoods). Moreover, performance on the inflection task was strongly correlated with number of years in formal education, with more educated participants achieving higher scores. A smaller-scale study by Wolff (1981) reports a similar finding for German participles.

There is also evidence of considerable variation in native speakers' ability to understand complex syntactic constructions. Dąbrowska (1997) presented participants of varying educational backgrounds with complex syntactic structures such as the *tough* movement construction (e.g., *John will be hard to get his wife to vouch for*), sentences with parasitic gaps (e.g., *It was King Louis who the general convinced that this slave might speak to*) and complex noun phrases (NPs; e.g., *Paul noticed that the fact that the room was tidy surprised Shona*). Comprehension was tested using simple questions about the sentences (e.g., *What did Paul notice? What surprised Shona?* for the complex NPs). Dąbrowska found a strong relationship between educational attainment and performance on the comprehension task: university lecturers outperformed graduate students who in turn were better than undergraduates who were better than unskilled workers with no university education.

The complex syntactic structures used in this study place heavy processing demands on participants and, therefore, the less educated participants' failure to respond correctly may be attributable to working memory limitations rather than being a reflection of their linguistic competence. However, the results reported in Chipere (2001, 2003) suggest that this is not the case. Chipere tested high academic attainment (HAA) and low academic attainment (LAA) participants' ability to comprehend and recall complex NP sentences. In the first phase of the experiment, the LAA group performed significantly worse than the HAA group on both tasks. In the second phase, Chipere divided the LAA participants into two subgroups that were given different types of training. One group was given memory training (i.e., participants were asked to repeat complex NP sentences). The other group was given comprehension training (i.e., participants received explicit instruction about complex NP constructions followed by a practice session with feedback). Both groups were then tested again with new complex NP sentences. Chipere found that memory training resulted in improved performance on the recall task, but not the comprehension task, although comprehension training led to an improvement in performance on both tasks. These results suggest that the LAA group's poor performance on the initial comprehension test was not attributable to limitations in working memory capacity, but was instead due to lack of experience with the complex NP construction.

Similar conclusions can be drawn from another training study conducted by Wells, Christiansen, Race, Acheson, and MacDonald (2009). Wells et al. point out that people with low reading spans have particular difficulty with object relatives such as [*The reporter that the senator attacked*] *admitted the error.* This is traditionally attributed to low working memory capacity, which has a disproportionate effect on structures that place heavy demands on working memory (Just & Carpenter, 1992). Following MacDonald and Christiansen (2002), Wells et al. suggest that the differences might instead be attributable to differences in linguistic experience: people who read more have more experience with language, particularly with more difficult structures, and thus are faster and more accurate at interpreting object relatives; they also perform better on the reading span task because the task actually measures language processing skill, not working memory capacity. In their study, Wells et al. exposed undergraduate students to 160 sentences containing subject and object relatives in two training sessions. A posttest administered 4 days after the second training session revealed that reading times at the main verbs for object relatives decreased as a result of training, while there was no analogous effect in the control group who had been exposed to different types of sentences. This is consistent with the idea that the individual differences in the processing of object relatives observed in earlier studies are attributable to differences in the amount of experience with this structure.

Several recent studies (e.g., Dąbrowska & Street, 2006; Street, 2009; Street & Dąbrowska, 2010) have suggested that there are considerable education-related differences in speakers' mastery of syntactically less complex constructions such as the English passive. The passive, although undeniably part of "core" grammar, does not involve embedding and thus can be plausibly assumed to place a lighter burden on working memory than the constructions used by Dąbrowska (1997), Chipere (2001), and Wells et al. (2009). Further, because the full passive occurs predominantly in formal written texts,[1] more educated speakers (who tend to read more) have considerably more experience of this construction. According to usage-based accounts (e.g., Bybee, 2010; Langacker, 2000; MacDonald & Christiansen, 2002), more experience with a construction should result in greater entrenchment and hence improved performance. Dąbrowska and Street (2006) tested this prediction employing a modified version of a task developed by Ferreira (2003) in which participants were asked to identify the "do-er" (i.e., AGENT) of plausible and implausible reversible passive sentences (e.g., *The man was bitten by the dog*, *The dog was bitten by the man*). As anticipated, the HAA group performed at ceiling, while the LAA participants had considerable problems with implausible passives, suggesting that they processed the test sentences nonsyntactically (i.e., relied on world knowledge rather than grammatical knowledge).

Further evidence for individual differences in grammatical knowledge comes from research by Street (2009) and Street and Dąbrowska (2010), who tested three structures: reversible passives (e.g., *The girl was kissed by the boy*), locative sentences with the universal quantifier *every* (e.g., *Every fish is in a bowl*; [Q-*is*]), and possessive locatives with *every* (e.g., *Every bowl has a fish in it;* [Q-*has*]). In addition, reversible actives (*The boy kissed the girl*) were used as a control condition. The first experiment tested HAA (i.e., postgraduate) and LAA (i.e., nongraduate) native-English speakers using a picture selection task. For the active and passive sentences, the pictures depicted simple reversible transitive events (e.g., a boy kissing a girl and a girl kissing a boy). For the quantifier sentences, they depicted objects and containers in a partial one-to-one correspondence (i.e., three bowls containing a fish plus one empty bowl, and three bowls containing a fish plus one fish without a bowl). The results of Experiment 1 revealed that although the graduate participants performed at ceiling in all conditions, the nongraduates performed at ceiling only on the actives (97% correct). Their performance was somewhat worse on passives (88%), worse still on Q-*is* sentences (78%), and at chance on Q-*has* (43%). It is interesting that group performance reflected the order of frequency of the respective constructions in the British National Corpus (i.e., active > passive > Q-*is* > Q-*has*. This could be due to differences in entrenchment, although, because different constructions are involved, other interpretations are also possible.

The second experiment was a training study, in which adult literacy students were tested before and after training on sentences similar to those used in Experiment 1. A pretest was used to select low-scoring participants who were then randomly assigned to either a quantifier training group or passive training group. Training comprised a short "grammar lesson" on either the passive or the quantifier construction, followed by a practice session with feedback. Immediately after training, participants were tested again with a different version of the task. To see if any effects of training were long lasting, participants were retested 1 and 12 weeks after training, again with different versions of the task.

The pretest results were very similar to those of the LAA group in Experiment 1, with order of difficulty reflecting the relative frequency of the constructions. The posttraining results show that training resulted in selective improvement. Participants trained on the quantifier constructions showed a clear improvement in performance on the quantifier conditions (where the mean scores rose from 37% in the Q-*is* condition and 13% in the Q-*has* condition to 94% and 100%, respectively, on the first posttest), but not on the passive, whereas participants trained on the passive showed a clear improvement in performance on those sentences (from 48% to 98%) but not on the quantifier sentences. This pattern of results was observed in all the posttraining tests, showing that the effects of training were long lasting. Thus, performance improved dramatically as a result of additional experience with the relevant construction, indicating that the initial differences in test scores are attributable to differences in specific linguistic knowledge.

The differences observed in Dąbrowska and Street (2006) and Street and Dąbrowska (2010) indicate that some native speakers have problems in interpreting passive sentences and that the differences between speakers are at least partly related to educational level: the HAA participants' performance was invariably at ceiling, whereas a vast amount of variation was observed in the performance of LAA individuals. With particular regard to the LAA group's performance on passive constructions, such variation was observed with different types of passive sentences (i.e., implausible as well as unbiased) and across different testing paradigms (i.e., naming the "do-er" and picture selection).

How can we account for these education-related differences in the interpretation of passive sentences? One possibility is that the low-performing speakers have actually mastered this construction, and their inability to reliably provide the correct answer in an experimental situation is attributable to performance problems. It is clear that the findings summarized above cannot be explained away by appealing to linguistically irrelevant performance factors such as failure to attend or cooperate with the researcher or inability to understand instructions, because the same participants performed at ceiling on active sentences. They also cannot be attributed to working memory limitations, if working memory is conceptualized as a resource with a fixed capacity, because in the second experiment reported by Street and Dąbrowska (2010) the performance improved dramatically following a training session lasting only about 5 min. However, it is possible that the observed differences in performance are attributable to individual differences in parsing ability rather than linguistic knowledge. Specifically, the results could be accommodated by a two-stage processing theory such as late assignment of syntax theory (LAST; see Townsend & Bever, 2001) according to which sentence processing involves two distinct phases. In the first phase, the processing system constructs a "pseudoparse," a rough analysis based on superficial probabilistic cues and heuristics such as the so-called noun–verb–noun (NVN) strategy for assigning thematic roles to predicates (see Bever, 1970; Townsend & Bever, 2001).[2] The pseudoparse is then used to guide the true parse, an algorithmic process that accesses syntactic knowledge to construct a complete syntactic representation. Constructing the true parse is slower and computationally more demanding, and thus may not be carried out in certain circumstances (e.g., under time pressure, or when processing resources are limited). Under this approach, the LAA participants would be assumed to have the same grammatical knowledge as the HAA group but be less likely to perform the true parse.

An alternative explanation, suggested by Dąbrowska and Street (2006), is that LAA participants lack a well-entrenched verb-general passive construction and therefore rely predominantly on lexically specific representations for individual verbs. Lexically specific representations are known to play an important role in language acquisition. Young children's syntactic representations have been shown to be tied to particular lexical items, typically verbs (Abbot-Smith & Tomasello, 2006; Behrens, 2009; Tomasello, 2003). Later in development, children acquire more abstract units by generalizing over the early lexically specific patterns. However, as demonstrated by the large body of work in the constraint-based lexicalist framework (Garnsey et al., 1997; Holmes et al., 1989; MacDonald, Pearlmutter, & Seidenberg, 1994; MacDonald & Seidenberg, 2006; Trueswell & Tanenhaus, 1994; Trueswell, Tanenhaus, & Kello, 1993), adults are very sensitive to the distributional peculiarities of individual words, suggesting that lexically specific representations survive into adulthood.

More abstract patterns are more difficult to acquire and process (Dąbrowska, 2010), and hence their full mastery requires exposure to a relatively large number of exemplars. Because passive sentences occur predominantly in formal written language, less educated speakers have comparatively little experience of this construction and may develop only a weak passive schema that cannot be reliably accessed during language use, or they may fail to develop a verb-general passive construction at all. Such individuals would still be able to produce and understand some passive sentences by relying on verb-specific representations (*NP1 was known by NP2, NP1 is based on NP2,* etc.) for verbs that frequently occur in the passive. However, if they encountered a passive sentence for which they lacked a verb-specific schema, they would have to either rely on analogy, which is costly in processing terms, or process it nonsyntactically (i.e., using a processing heuristic such as the NVN strategy), which leads to errors (cf. Dąbrowska & Street, 2006). Thus, according to this account, inconsistent individual performance on passive sentences is attributable to a lack of a well-entrenched verb-general passive construction.

This article examines the role of lexically specific representations and processing heuristics in the comprehension of passives by LAA and HAA adults. We use a version of a task developed by Ferreira (2003) in which participants are presented with active and passive sentences and asked to identify the "do-er" (i.e., AGENT) or the "acted-on" (i.e., the PATIENT). Ferreira's study compared performance on plausible passives such as *The man was bitten by the dog*, symmetrical passives such as *The man was visited by the woman*, and implausible passives like *The dog way bitten by the man.* The participants (undergraduate students) performed above chance in all conditions, indicating that they had the relevant syntactic knowledge. However, error rates were relatively high: about 12% for plausible sentences, 18% for symmetrical sentences, and 23% for implausible sentences.[3] This, according to Ferreira, indicates that they did not always compute the full parse, but often relied on a "quick and dirty" pseudoparse, for instance, assumed that NVN = AGENT VERB PATIENT.

The aim of our study is twofold: (a) to determine whether the education-related differences observed in earlier studies can be attributed to differences in the use of processing heuristics, specifically, that some participants may not compute a full parse; and (b) to evaluate the role of lexically specific representations in the processing of passive sentences by HAA and LAA participants. To do this, we compare participants' comprehension of passive sentences containing verbs that are strongly associated with active sentences (PV_A) and verbs that are strongly associated with passives (PV_P). We also compare participants' comprehension of active sentences containing verbs strongly associated with passive constructions (AV_P) and verbs strongly associated with active constructions (AV_A). The active sentences serve as a control condition: all speakers should have a well-entrenched verb-general active construction, and therefore no education-related differences are expected.

PREDICTIONS

Our experiment was designed to test two usage-based predictions and an additional prediction derived from the LAST. According to usage-based theories (Bybee, 2010; Langacker, 2000), language use is influenced by two main factors: frequency and lexical specificity. Repeated experience with a particular construction leads to greater entrenchment, which in turn results in faster and more accurate processing. Therefore, we expect active sentences to be easier than passive sentences for all participants. Moreover, because full passives occur predominantly in formal written texts, differences between active and passive sentences should be particularly pronounced in less educated participants, who have relatively little experience of such texts. With regard to lexical specificity, usage-based theories maintain that more abstract (i.e., lexically underspecified) constructions are more difficult to access. Thus, all other things being equal, participants should do better (i.e., process more quickly and interpret more accurately) in those sentences that contain verbs for which they have lexically specific representations; that is, they do better in passive sentences containing verbs strongly associated with the passive (PV_P) than in passive sentences containing verbs strongly associated with the active (PV_A). An analogous difference between active–attracting and passive–attracting verbs in active sentences would also be compatible with a usage-based model; however, because active sentences are known to be processed very quickly and accurately, it is likely that these differences will be masked by ceiling effects. Thus the predicted order of difficulty is that active–attracting verbs used in active sentences (AV_A) and passive–attracting verbs used in active sentences (AV_P) will be easier than passive–attracting verbs in passive sentences (PV_P) that in turn will be easier than active–attracting verbs in passive sentences (PV_A):

Prediction 1: AV_A, $AV_P > PV_P > PV_A$

We expect that the HAA participants will perform at ceiling, so the differences will show up only in reaction time data. For the LAA group, the effects are likely to be visible on both measures.

Furthermore, if, as hypothesized by Dąbrowska and Street (2006), LAA participants rely more on lexically specific knowledge about passives, lexical effects should be more pronounced for the LAA group; in other words, there should be an interaction among construction, verb, and group.

Prediction 2: For the HAA group, $PV_P > PV_A$; for the LAA group, $PV_P >> PV_A$

Two-stage processing theories such as LAST do not predict differences between active- and passive–attracting verbs in either construction because the same general rules and parsing routines are thought to apply in both cases. However, they do predict a speed–accuracy trade-off for passives: because the true parse requires additional processing time, participants who are more accurate should respond more slowly than the less accurate participants.

Prediction 3: There should be a positive correlation between accuracy and reaction time for passives.

METHOD

We employed a 2 × 2 × 2 mixed design with one between-group independent variable: level of academic attainment (HAA vs. LAA) and two within-group independent variables: construction (active vs. passive) and verb (active–attracting vs. passive–attracting). The dependent variables are decision accuracy and response time. Participants read simple active transitive and full passive sentences, and at the end of each they identified a named participant in the sentence as either the "do-er" (i.e., AGENT) or the "acted on" (i.e., PATIENT; for a similar naming task, see Ferreira, 2003).

Participants

Sixty-four 17- to 50-year-old adults (25 males, 39 females) participated in the experiment. The HAA group comprised 31 postgraduate students or recent graduates. All participants had at least an MA degree (i.e., at least 17 years of formal education); some were studying for or had recently completed a PhD. They came from a variety of academic disciplines (arts and humanities, social sciences, and life sciences). The remaining 33 participants (the nongraduate, LAA group) had at most 11 years of formal education and were employed as packers, cleaners, and hairdressers. All participants were native speakers of English. One LAA participant did not fully engage with the task and talked throughout the trials and was therefore excluded from the data analyses. Thus, the final sample comprised 63 participants.

Materials

There were 24 test sentences containing 12 verbs.[4] Six of the verbs were associated with the active transitive construction; the other six were associated with the passive (based on the British International Corpus of English [ICE-GB]). To determine which verbs are strongly associated with active and passive constructions, we employed collostructional analysis (Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003). Collostructional analysis measures the association of words to syntactic constructions rather than words to words (as in traditional collocational analysis). This involves two steps. First, the analyst establishes the frequency of the target word in the target construction (e.g., *kick* in the passive), the target word in a contrasting construction (i.e., *kick* in the active), the frequency of all other words in the target construction (i.e., all verbs other than *kick* in the passive), and the frequency of all other words in the contrasting construction (in this case, all other verbs in the active). Then the observed frequency of the word in the construction is compared to one expected by chance, given the frequency of the verb and the frequency of the other constructions. The strength of attraction or repulsion of the word to the construction is measured using the *p* value of a Fisher–Yates exact test. Although this does not give an effect size, there are several advantages to using the *p* value of a Fisher–Yates exact test. For example, in addition to incorporating the size of the effect in any cross-tabulation (normally obtained using η^2, *d*, or r^2), it also weighs the effect on the basis of the observed frequencies with the added advantage of not overestimating the association of the strength of infrequent data. By way of example, Stefanowitsch and Gries observe that in the ICE-GB corpus, 14 out of 35 occurrences (i.e., 40%) of the N *waiting to happen* construction involve the lexeme *accident*. This yields a Fisher–Yates *p* value of 2.12E-34. By contrast, if only 8 instances of *accident* in a total of 20 cases in the N *waiting to happen* construction had been observed (again 40%), this would yield a *p* value of 3.22E-20. The lower *p* value in this particular example reflects that the attraction is considered more noteworthy if it is observed for a greater number of occurrences (i.e., in this case the occurrence of *accident* in the noun slot of the N *waiting to happen* construction). As Stefanowitsch and Gries note, "this sensitivity to frequency seems a desirable property for a measure of collostruction strength, given that frequency plays an important role for the degree to which constructions are entrenched and the likelihood of the production of lexemes in individual constructions" (2003, pp. 218–219). It is also worth noting that significance level (e.g., .05, .01) is considered to be less important than the ranking of collostructions.

Many of the most distinctive collexemes (i.e., lexemes which can occur in either the active transitive or passive but appear in one of the constructions significantly more often than would be expected by chance) of the two constructions are psychological verbs such as *hate, impress,* and *remember.* However, pilot studies revealed that a number of participants were unsure about assigning the labels "do-er" and "acted-on" to participants in sentences containing such verbs; and some participants were inclined to identify the PATIENT as the "do-er" in some cases. For example, in a sentence such as *Daniel hated Peter*, some participants thought that Peter was the "do-er" of some action that had caused Daniel to hate him. To avoid such ambiguities, all the verbs used in the experiment designated dynamic transitive events that readily allow two animate arguments. As a consequence, although the verbs are distinct collexemes for their respective constructions, they are not the most distinct collexemes for these constructions. A list of all the verbs used in the experiment, their frequencies in the active and passive in the ICE-GB corpus, and the Fisher–Yates *p* value computed on the basis of these figures are given in Table 1.[5] Note that because attraction to a construction is calculated relative to the construction's frequency, and actives are much more frequent than passives, verbs attracted to the passive are not necessarily more frequent in the passive. *Kick*, for example, is more frequent in actives than passives but occurs in the passive more than one would expect by chance given the frequency of the verb and the frequency of the other constructions. Because the absolute frequency of

Table 1. *Relative frequencies in the active and passive, overall frequencies and p value of attraction to construction of the verbs used in the experiment*

	Frequency					
Verb	Active ICE-GB	Passive ICE-GB	Active BNC	Passive BNC	Overall BNC	Fisher–Yates *p*
			Type A: Active Attracting			Attraction to Actives
Touch	30	3	90	10	10,836	.121
Beat	38	8	60	30	9,440	.526
Shake	28	2	98	2	8,903	.068
Bite	6	1	76	24	3,414	.623
Punch	4	0	74	26	2,319	.445
Grab	7	0	88	12	2,936	.242
Mean frequency	18.8	2.3	81.0	17.3	6,308	
			Type P: Passive Attracting			Attraction to Passives
Kick	6	3	55	45	5,167	.219
Feed	26	17	40	60	7,479	.001
Injure	4	19	17	83	2,508	<.001
Attack	35	21	45	55	16,411	.001
Flick	5	5	70	30	1,254	.023
Hurt	15	8	40	50	5,239	.046
Mean frequency	15.2	12.2	44.5	53.8	6,343	

Note: ICE-BG, British ICE corpus; BNC, British National Corpus.

these verbs in the ICE-GB is relatively small, we have also included estimates of the frequency of these verbs in each construction in the British National Corpus. These were computed on the basis of a random sample of 100 tokens of each verb.

Each verb occurred in both the active and passive construction. Thus, of the 24 test sentences, 6 were active transitive containing an active–attracting verb (V_A), 6 were active transitive containing a passive–attracting verb (V_P), 6 were passives containing a V_A, and 6 were passives containing a V_P. Twelve pairs of names (6 male, 6 female) were used for the AGENT and PATIENT thematic roles in the test sentences. Twelve sentences contained male names, and the other 12 sentences contained female names. Thus, for each of the 24 test sentences, the NPs occupying the AGENT and PATIENT roles were of the same gender. These were used instead of common nouns in order to reduce noise by eliminating the effects of pragmatics (e.g., participants may feel that a man is more likely to attack a woman than vice versa). Each name within each pair appears four times: once as AGENT of an active sentence; once as AGENT of a passive sentence; once as PATIENT in an active transitive; and once as PATIENT in a passive. Because passive sentences contain

Table 2. *Examples of test sentences*

	Construction	
Verb	Active	Passive
Active attracting	*James grabbed Peter*	*Sally was bitten by Rachel*
Passive attracting	*Jane injured Emma*	*Robert was attacked by John*

additional morphological markers (the auxiliary *be* and the preposition *by*), they were slightly longer than the actives; however, within the same construction, all the sentences were of approximately the same length (i.e., contained approximately the same number of letters and same number of syllables). Example sentences of each type are given in Table 2.

There were four versions of the test, each containing six sentences for each of the four conditions; and within any one version there were no repeats of the same action involving the same participants (i.e., in any one version no name appears with the same verb twice). For descriptions of simple transitive events there are four possible descriptions (e.g., *Mary punched Sandra*, *Sandra punched Mary*, *Mary was punched by Sandra* and *Sandra was punched by Mary*); each of the four possible descriptions appeared in a different version of the test. Ultimately, each participant saw 24 sentences: 6 AV_A, 6 AV_P, 6 PV_P, and 6 PV_A with no repeats of the same names with the same verbs and each name appearing twice in the AGENT role and twice in the PATIENT role. The order of sentences was randomized for each participant. A complete list of sentences used in one version of the test is given in Appendix A.

Procedure

The experimental session began with participants reading written instructions displayed on a laptop screen. Participants were informed that they would be presented with a series of sentences followed by a name, for instance,

Justin scratched Steven
Justin

and that their task was to identify the person named as either the "do-er" (i.e., AGENT) or the "acted-on" (i.e., PATIENT). This was followed by a brief nontechnical explanation of the terms "do-er" and "acted-on." A complete transcript of the written instructions is given in Appendix B.

Participants responded by pressing "D" (for "do-er") or "A" (for "acted-on") on the keypad. Both these keys were highlighted (using white stickers situated beneath the letter at the bottom of the key). The sentence remained on screen until a participant pressed either "D" or "A"; it then disappeared and was replaced by a short sign (+++) and shortly afterward, the next sentence.

Before the test trials began, all instructions were clarified (and concept checked) verbally by the experimenter. Participants then completed four practice trials (two active, two passive). These were supervised by the experimenter to ensure that participants understood the task. Participants were tested individually, with each testing session lasting approximately 5 min. The stimuli were presented using E-prime software (Psychology Software Tool, Pittsburgh, PA), which also recorded the participants' decision accuracy and reaction times.

RESULTS

Decision accuracy

The accuracy data are summarized in Table 3. As can be seen from the table, the HAA group performed at ceiling in all conditions. The LAA group also performed at ceiling on active sentences but had considerably lower scores for passives. As the results are not normally distributed, the data were analyzed using nonparametric tests (Kruskal–Wallis). These show that there are no significant differences between groups on decision accuracy for the active sentences: for AV_A, H (1) = 1.017, p = .313; for AV_P, H (1) = 0.153, p = .696. However, there are significant differences between groups in decision accuracy for the passives: for PV_A, H (1) = 4.587, p = .032; for PV_P, H (1) = 9.843, p = .002.

In order to further examine the effects of construction and the effects of verb, we conducted two sets of planned comparisons for each group. To examine the effect of construction we compared participants' scores for active and passive sentences with each type of verb (i.e., AV_A vs. PV_A and AV_P vs. PV_P). To examine the effect of verb type we compared participants' scores for active–attracting and passive–attracting verbs in each construction (i.e., AV_A vs. AV_P and PV_A vs. PV_P). Again, as the results are not normally distributed, the scores were compared using a nonparametric equivalent of the paired t test: the Wilcoxon signed ranks. Comparisons adjusted using the Bonferroni correction give a probability value of p = .016.

Effect of construction. At this revised level, the results show that for the HAA group there is a small but significant difference between AV_A and PV_A: $z = -2.530$, p = .011, N = 31, r = .22. However, there is no significant difference between AV_P and PV_P: z = 1.000, p = 1.000, N = 31, r = .08. By contrast, Wilcoxon tests for the LAA group reveal that both comparisons are highly significant at the corrected probability level of p = .016 and that the difference is much greater than that of the HAA group: for AV_A versus PV_A, $z = -3.384$, p = .001, N = 32, r = .29; for AV_P versus PV_P, z = –2.810, p = .005, N = 32, r = .25. Thus, the LAA group made more errors on both types of passive sentences. This is also evident if results for the two verb types are collapsed (cf. active and passive columns in Table 3). Wilcoxon tests reveal that, although differences in decision accuracy for actives and passives is not significant for the HAA group ($z = -1.890$, p = .059, N = 31, r = .16), they are highly significant for the LAA participants (z = 3.532, p < .001, N = 32, r = .31).

Table 3. *Proportion of correct responses for each condition by group*

Group	AV_A	AV_P	PV_A	PV_P	AV_A and AV_P	PV_A and PV_P
HAA ($N = 31$)						
Mean (*SD*)	99 (3)	98 (6)	95 (8)	98 (6)	99 (3)	98 (6)
Median (range)	100 (83–100)	100 (83–100)	100 (83–100)	100 (83–100)	100 (92–100)	100 (83–100)
LAA ($N = 32$)						
Mean (*SD*)	98 (7)	98 (7)	86 (20)	85 (20)	98 (6)	86 (18)
Median (range)	100 (67–100)	100 (67–100)	91 (0–100)	100 (17–100)	100 (67–100)	91 (9–100)

Note: AV_A, active-attracting verb in active sentence; AV_P, passive-attracting verb in active sentence; PV_A, active-attracting verb in passive sentence; PV_P, passive-attracting verb in passive sentence; AV_A and AV_P, active (collapsed); PV_A and PV_P, passive (collapsed).

Table 4. *Mean (standard deviation) response times (ms) of sentences by group*

Group	AV_A	AV_P	PV_A	PV_P	Overall Mean
HAA ($N = 31$)					
Mean (*SD*)	3495 (863)	3484 (982)	4720 (1492)	4436 (1396)	4034
Range	1928–5227	2084–6543	2240–8182	2398–7098	
LAA ($N = 32$)					
Mean (*SD*)	4170 (1225)	4156 (1316)	5688 (1496)	5385 (1415)	4850
Range	229–5962	2329–6918	2937–9828	2897–9248	
Overall mean	3833	3820	5204	4910	

Note: AV_A, active-attracting verb in active sentence; AV_P, passive-attracting verb in active sentence; PV_A, active-attracting verb in passive sentence; PV_P, passive-attracting verb in passive sentence; HAA, high academic attainment; LAA, low academic attainment.

Effect of verb type. With regard to the effect of verb type, as anticipated, there was no significant difference between AV_A and AV_P for either group: for the HAA participants, $z = -1.342$, $p = .180$, $N = 31$, $r = .12$; for the LAA participants, $z = 2.000$, $p = 1.000$, $N = 32$, $r = .17$. However, the predicted difference between PV_A and PV_P was not observed: for HAA participants, $z = -1.890$; $p = .059$, $N = 31$, $r = .16$; and for the LAA group, $z = 0.053$, $p = .958$, $N = 32$, $r = 0$.

Individual differences. As can be seen from the standard deviations and ranges in Table 3, there were considerable individual differences in the LAA participants' performance on passive sentences. There were a total of 12 passive sentences on the test. According to the binomial distribution ($p < .05$), above chance performance requires 10 out of 12 correct responses, and a score of 2 or less would be below chance. At this criterion, 24 of the LAA participants (i.e., 75%) performed above chance; 7 were at chance (scores between 3 and 10); and 1 participant performed below chance, supplying the correct response on only one of the passive trials. (This participant responded correctly to 11/12 actives, so she or he consistently interpreted the first NP in the sentence as the agent.) All of the HAA participants performed above chance.

Response time

Mean response times and standard deviations for all conditions by groups are summarized in Table 4. The data were analyzed by means of a 2 (construction: active vs. passive) × 2 (verb: active–attracting versus passive–attracting) × 2 (group: HAA vs. LAA) mixed-design factorial analysis of variance (ANOVA).[6] The dependent measure was response time. All effects are reported as significant at $p < .05$. The analysis revealed a main effect of construction, $F_1\ (1, 61) = 111.681$, $p < .001$, partial $\eta^2 = 0.647$; $F_2\ (1, 10) = 118.362$, $p < .001$, partial $\eta^2 = 0.922$, indicating that, as expected, both groups were faster at processing active constructions than passive constructions. There was also a main effect for group, F_1 $(1, 61) = 7.913$, $p < .007$, partial $\eta^2 = 0.115$; F_2 $(1, 10) = 61.644$, $p < .001$, partial $\eta^2 = 0.860$, indicating that the HAA participants are significantly faster than the LAA participants. Finally, the analysis by subject revealed a main effect of verb, F_1 $(1, 61) = 5.612$, $p < .0211$, partial $\eta^2 = 0.084$, qualified by a Construction × Verb interaction, F_1 $(1, 61) = 4.554$, $p < .037$, partial $\eta^2 = 0.069$; however, neither of these was significant in the analysis by item. This suggests that using a verb in a construction with which it is not normally associated incurs a higher cost for passive sentences than for actives. As shown in Table 4, differences in response times between AV_A (active constructions containing active–attracting verbs) and AV_P (active constructions containing passive–attracting verbs) conditions are negligible for both groups (response times are faster on the AV_P condition). However, paired *t* tests reveal that both groups' response times were significantly faster for the PV_P (passive sentences containing passive–attracting verbs) condition than the PV_A (passive sentences containing active–attracting verbs) condition: for HAA participants, $t\,(30) = 1.996$, $p = .050$, $d = 0.2$; for the LAA group, $t\,(31) = 2.049$, $p = .049$, $d = 0.2$.

The reaction time results show that the HAA group was faster than the LAA group overall. This is to be expected and could be due to a number of factors (e.g., greater overall processing speed, greater familiarity with computers, or simply better literacy skills). As expected, both groups had slower reaction times for passives. This could be attributed to entrenchment (actives are much more frequent than passives), sentence length (passives are slightly longer than actives), morphological complexity (passives contain additional morphological markers), or, in a generative framework, the belief that passives involve movement. The processing of passives was speeded up with verbs that are attracted by the construction, showing that participants have lexically specific knowledge about passives. As anticipated, there was no corresponding effect for actives, which is likely to be a ceiling effect: processing times for the active are so short that there is no room for other influences to manifest themselves.

Relationship between speed and accuracy

To determine whether there is a relationship between speed and accuracy, we computed Spearman correlations between these two variables for all sentences and for actives and passives separately. The relevant figures are presented in Table 5. The correlation coefficients range from weak and statistically insignificant to moderately strong and highly significant. It is crucial, however, that they are all negative, indicating that the more accurate participants also responded faster than less accurate participants. Thus, our data show no evidence of a speed–accuracy trade-off.

DISCUSSION

The accuracy data reported here replicate earlier research suggesting that many speakers with relatively little schooling have problems interpreting passives. The overall performance on passives in the LAA group (86% correct) is similar to

Table 5. *Correlations between decision accuracy and response time*

	Academic Attainment		
Condition	High ($N = 31$)	Low ($N = 32$)	All Participants ($N = 63$)
Actives	−.28	−.28	−.38**
Passives	−.09	−.41*	−.42***
All sentences	−.16	−.30	−.40***

*$p = .05$. **$p = .01$. ***$p = .001$.

that observed in Street and Dąbrowska (2010), even though the testing procedure (naming the "do-er" or "acted-on") was much more demanding than the picture selection task used in that study.

Do these differences in decision accuracy reflect differences in underlying linguistic representations, or could they be attributed to linguistically irrelevant factors such as willingness to cooperate with the experimenter, amount of experience with formal testing, or ability to perform the experimental task? In our view, appeals to performance factors as an explanation of the decision accuracy results are highly unsatisfactory. First, issues surrounding "testwiseness" should be evident across all constructions; yet the decision accuracy of the LAA group was at ceiling in the active conditions. Second, the test trials were preceded by a supervised practice trial conducted to establish that participants did understand the task. Third and finally, all participants were extremely cooperative.

As explained in the introductory section, the main purpose of the study was to determine whether the less educated speakers' difficulties with the passive can be attributed to reliance on lexically specific representations or processing heuristics. Specifically, we tested two predictions derived from usage-based models (that there should be a processing advantage for passive–attracting verbs used in passive sentences and that this advantage should be stronger for the LAA group) and a further prediction derived from two-stage processing models such as LAST (that faster reaction times should be associated with lower accuracy).

Our results support the idea that speakers make use of lexically specific representations. As we have seen, both groups responded faster to passive sentences containing passive–attracting verbs than to passive sentences with active–attracting verbs. This is consistent with earlier research suggesting that lexically specific knowledge survives into adulthood (Garnsey et al., 1997; Holmes et al., 1989; MacDonald et al., 1994; MacDonald & Seidenberg, 2006; Trueswell et al., 1993; Trueswell & Tanenhaus, 1994).

However, Prediction 2 (for the HAA group, $PV_P > PV_A$; for the LAA group, $PV_P >> PV_A$) was not confirmed. There was no interaction between construction, verb type, and group: although the LAA participants were somewhat slower overall than the HAA group, they did not experience a particular difficulty with active–attracting verbs in the passive. Moreover, they were equally accurate with passive sentences with passive–attracting verbs and passive sentences with active–attracting verbs (85.4% and 85.9% correct, respectively). Thus, our results do not support the proposal made by Dąbrowska and Street (2006) that education-related differences in the comprehension of passive sentences arise as a consequence of LAA participants relying on verb-specific schemas to a greater extent than more educated participants.

Finally, we found that the HAA participants were significantly faster as well as more accurate at processing passive sentences; and in both groups the faster participants were also more accurate. Thus, Prediction 3, that there should be a positive correlation between accuracy and reaction time for passives, was also not confirmed. This finding is inconsistent with the proposal that individual differences in the comprehension of passive sentences are attributable to differences in parsing ability, specifically the contention that mistakes in the comprehension of passives arise when listeners abort processing after the pseudoparse and never compute the true parse. Such an account would predict a speed–accuracy trade-off; but there is no evidence of such a trade-off in our data; we actually found the opposite.

Thus, neither of the two explanations proposed in the introduction can account for the observed results. Our findings suggest that both groups of participants have verb-general as well as verb-specific representations and that both process the sentences in the same way (although the LAA participants are somewhat slower). The difference, we suggest, lies in the degree of entrenchment of these representations, which in turn is a function of the amount of experience with the passive construction. Passive sentences are considerably more frequent in written texts (particularly formal, written texts) than in speech; because more educated participants tend to read more, their passive constructions are better entrenched, and hence accessed more reliably, which results in faster and more accurate performance. This is true for verb-specific constructions (because they have more experience with individual passive–attracting verbs) as well as verb-general constructions (because they have experienced more verb types in the passive). Entrenchment is clearly a matter of degree; hence, performance on relatively infrequent structures such as the passive varies considerably, particularly in the LAA group, where individual scores on passive sentences ranged from 8% to 100%.

The results reported here support two fundamental claims of usage-based models: that much of our linguistic knowledge is lexically specific, and that frequency plays a crucial role in shaping speakers' mental grammars. More experience with a particular construction results in greater entrenchment, and hence, more reliable performance. Our results are also compatible with constraint-based lexicalist (CBL) models (MacDonald et al., 1994; MacDonald & Seidenberg, 2006; Trueswell & Tanenhaus, 1994). CBL models share many assumptions with usage-based models, and indeed can be regarded as a type of usage-based model. Proponents of such approaches see sentence comprehension as involving rapid integration of a variety of probabilistic constraints emerging from the lexical properties of individual words, the relative frequency of the verb in different constructions, the frequency of the constructions themselves, thematic fit of the verb's arguments, and information derived from the preceding discourse and the nonlinguistic context. According to CBL models, lexical information, including grammatical preferences of individual verbs, is a particularly good cue, in that it is frequently available and more reliable than most other cues. Our results indicate that speakers know which verbs are strongly associated with the active and the passive and are able to use this information on-line; consequently, passive sentences with passive attracting verbs are processed faster than passives with active–attracting verbs.

CONCLUSION

As in previous studies (e.g., Dąbrowska & Street, 2006; Street & Dąbrowska, 2010), the LAA participants performed well below ceiling (and, in some cases, at or even below chance) on a task tapping comprehension of passive sentences. Their relatively poor performance cannot be explained by task demands because both groups performed at ceiling on actives.

Both groups processed passive sentences with passive–attracting verbs faster than passives with active–attracting verbs, suggesting that both groups have lexically specific knowledge about which verbs are used in the passive. This finding adds to the substantial body of research within the CBL framework, suggesting that speakers are highly sensitive to the distributional characteristics of individual words. Significantly, the processing advantage for passives with passive–attracting verbs was equally strong in both groups, but there were no differences between the number of correct responses for passives with passive–attracting and passive–repelling verbs in either group. This shows that the LAA group's relatively poor performance on passives cannot be explained by appealing to differences in the availability of a verb-general passive construction (as proposed by Dąbrowska & Street, 2006). Instead the results suggest that both groups have verb-specific and verb-general constructions. It is significant, however, that the LAA group's knowledge about the passive is less entrenched, and hence accessed less reliably.

As pointed out in the introductory section, the claim that all first language learners acquire essentially the same grammar is one of the strongest arguments for an innate language faculty and one of the "facts" that the universal grammar hypothesis was supposed to explain. The results discussed in this paper add to a growing body of evidence suggesting that the convergence argument is based on a false premise: there are actually substantial differences in native speakers' mastery of a number of grammatical constructions, and these differences appear to be related to differences in linguistic experience. It does not necessarily follow from this that universal grammar does not exist: one can argue for strong innate constraints on language learning on other grounds (e.g., poverty of the stimulus). It is also possible that the individual differences observed in this and other studies reflect differences in the innate biases for language learning. Given that there are considerable individual differences in most genetically determined traits, we would expect variation in the innate endowment for language as well.

Many linguists will argue that our less educated participants' difficulties in interpreting passive sentences are facts about linguistic performance rather than competence (and hence irrelevant to the truth or falsity of the universal grammar hypothesis). Whether we accept this argument depends, of course, on where we draw the line between competence and performance: in a usage-based theory, systematic differences in comprehension accuracy and processing speed are thought to reflect differences in entrenchment, and hence facts about speakers' linguistic representations, not just their use of these representations in processing. Wherever we draw the line, it is clear that some adults perform at chance, or even below chance on a task tapping comprehension of passive sentences, which is problematic for theories that claim all first language learners master the basic constructions of their language at a young age. Achieving full mastery of the passive (and other constructions, see introductory section) requires a considerable amount of experience with this construction.

APPENDIX A

Sentences used in one version of the test

AV_A

Thomas bit Roger
James grabbed Peter
Mary punched Sandra
Emma touched Jane
Robert shook John
Rachel beat Sally

PV_A

Sally was bitten by Rachel
Sandra was shaken by Mary
Peter was beaten by James
John was punched by Robert
Jane was grabbed by Emma
Roger was touched by Thomas

AV_P

Roger fed Thomas
Peter kicked James
John flicked Robert
Sandra attacked Mary
Jane injured Emma
Sally hurt Rachel

PV_P

Thomas was hurt by Roger
James was injured by Peter
Robert was attacked by John
Mary was flicked by Sandra
Emma was kicked by Jane
Rachel was fed by Sally

APPENDIX B
Written instructions

You will see sentences like this one: Bill hit Jamie.

In this sentence Bill is the "doer"—the person doing the action (he hits Jamie).
Jamie is the "acted on"—the person affected by the action (somebody hits him).

Your task is to read each sentence and then name either the "do-er" by pressing "D" on the keypad OR the "acted on" by pressing "A" on the keypad.

Example: for this sentence you would press "D" because Justin is the "do-er."

Justin scratched Steven
    Justin

but for this sentence you would press "A" because Jacky is "acted on."
Tracey hit Jacky
    Jacky

You have as much time as you need to answer: the sentence will remain on screen until you select either "D" or "A".

After your selection, you will see this sign +++ for 2 seconds. Then the next sentence will appear.

The first four sentences are to practice.

Press SPACEBAR to begin Practice Session.

ACKNOWLEDGMENTS
This research was partially supported by AHRC Award 06/125939 (to J.A.S.).

NOTES
1. Passives are four to five times more frequent in writing than in speech (Roland, Dick, & Elman, 2007). The difference in frequency is even more pronounced for full passives (i.e., passives with an agentive adjunct). In the British National Corpus, for example, the average frequency of the full passive is 63 per million words for written texts and about 9 per million in speech, a sevenfold difference in frequency.
2. The NVN strategy refers to the comprehender's tendency to assume that the subject of a sentence is also the AGENT of a particular action and the object of the same action is THEME. It is argued that this occurs because the majority of sentences in English conform to this pattern. Consequently, those sentences that require the PATIENT thematic role to be assigned before the AGENT thematic role (e.g., passives, object-clefts) prove more difficult to process than those that conform to the more common NVN pattern (see Ferreira, 2003, p. 25).
3. These figures were obtained by averaging across several different conditions in two experiments.
4. The number of items per condition (six) was kept relatively small to keep the task as short as possible in order to ensure that the low-education participants remained

engaged; however, it is in line with previous studies using a similar methodology (e.g., Ferreira, 2003).

5. The Fisher–Yates *p* values and frequencies in the ICE-GB corpus are taken from unpublished data collected by Stefan Gries and Anatol Stefanowitsch. We thank these researchers for sharing this information with us.
6. To ensure that the results reported here are not unduly influenced by outliers, we conducted a second ANOVA in which all observations that were more than 1.5 *SD* from the mean were removed. The results were very similar to those reported in the text.

REFERENCES

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *Linguistic Review, 23*, 275–290.

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics, 47*, 383–411.

Bever, Thomas G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 279–362). New York: Wiley.

Birdsong, D. (2004). Second language acquisition and ultimate attainment. In A. Davies & C. Elder (Eds.), *Handbook of applied linguistics* (pp. 82–105). London: Blackwell.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis, 20*, 3–49.

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition, 31*, 175–198.

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

Chipere, N. (2001). Native speaker variations in syntactic competence: Implications for first language teaching. *Language Awareness, 10*, 107–124.

Chipere, N. (2003). *Understanding complex sentences: Native speaker variation in syntactic competence*. New York: Palgrave Macmillan.

Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.

Crain, S., & Lillo-Martin, D. (1999). *An introduction to linguistic theory and language acquisition*. Malden, MA: Blackwell.

Dąbrowska, E. (1997). The LAD goes to school: A cautionary tale for nativists. *Linguistics, 35*, 735–766.

Dąbrowska, E. (2008a). The effects of frequency and neighborhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language, 58*, 931–951.

Dąbrowska, E. (2008b). The later development of an early-emerging system: The curious case of the Polish genitive. *Linguistics, 46*, 629–650.

Dąbrowska, E. (2010). The mean lean grammar machine meets the human mind: Empirical investigations of the mental status of rules. In H.-J. Schmid, S. Handl, & S. Handl (Eds.), *Empirical approaches to cognitive linguistics* (pp. 151–170). Berlin: Mouton de Gruyter.

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*, 604–615.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203.

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

Gries, S. Th., & Stefanowitsch, A. (2004). Extending collostructional analysis. A corpus-based perspective on "alternations." *International Journal of Corpus Linguistics, 9*, 97–129.

Holmes, V. M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language 28*, 668–689.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). Stanford, CA: CSLI.

Lidz, J., & Williams, A. (2009). Constructions on holiday. *Cognitive Linguistics, 20*, 177–189.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review, 109*, 35–54.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 581–611). London: Elsevier.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science, 291*, 114–118.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*, 348–379.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275*, 1599–1603.

Stefanowitsch, A., & Gries, S. Th. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics, 8*, 209–243.

Street, J. A. (2009). *Individual differences in comprehension of passives and universal quantifiers by adult native speakers of English.* Unpublished doctoral thesis, University of Sheffield.

Street, J. A., & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua, 120*, 2080–2094.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of child language acquisition*. Cambridge, MA: Harvard University Press.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, Jr., & L. Frazier (Eds.), *Perspectives in sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.

Trueswell, J. C., Tanenhaus, M. K., & 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.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58*, 250–271.

Wolff, R. A. (1981). German past participles and the simplicity metric. *Linguistics, 19*, 3–13.