Lingua 120 (2010), pp. 2080-2094

More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers?

James A. Street
Department of Humanities, School of Arts and Social Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST
james.street@northumbria.ac.uk

Ewa Dąbrowska
Department of Humanities, School of Arts and Social Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST
ewa.dabrowska@northumbria.ac.uk

This paper provides experimental evidence suggesting that there are considerable differences in native language attainment, and that these are at least partially attributable to individual speakers' experience. Experiment 1 tested high academic attainment (hereafter, HAA) and low academic attainment (LAA) participants' comprehension using a picture selection task. Test sentences comprised passives and two variants of the universal quantification construction. Active constructions were used as a control condition. HAA participants performed at ceiling in all conditions; LAA participants performed at ceiling only on actives. As predicted by usage-based accounts, the order of difficulty of the four sentence types mirrored their frequency. Experiment 2 tested whether the less educated participants' difficulties with these constructions are attributable to insufficient experience. After a screening test, low scoring participants were randomly assigned to two training groups. The passive training group were given a short training session on the passive construction; and the quantifier training group were trained on sentences with quantifiers. A series of post-training tests show that performance on the trained construction improved dramatically, and that the effect was long lasting.

Keywords: Individual differences, English passive construction, Universal quantification, Sentence comprehension, Linguistic competence, Linguistic performance

More Individual Differences in Language Attainment: How much do adult native speakers of English know about passives and quantifiers?

1. Introduction

Many linguists subscribe to the view that all native speakers acquire more or less the same grammar. Crain and Lillo-Martin, for example, state that "… children in the same linguistic community all learn the same grammar" (1999: 9); Seidenberg notes that "… children are exposed to different samples of utterances but converge on the same grammar" (1997: 1600); according to Nowak, Komarova and Niyogi, "children of the same speech community reliably learn the same grammar" (2001: 114); and Hermon (2002) argues that "language learning cannot be by trial and error, otherwise children would not all converge on the same grammar" . That all learners converge on the same grammar is taken as self-evident: indeed, this "fact" is often used as support for other claims, as illustrated by the last quotation (cf. also Chomsky 1975: 11).

To be sure, it is generally acknowledged that there are some individual differences in grammatical knowledge: Kayne (2000: 7), for instance, concedes that "[i]t is entirely likely that no two speakers of English have entirely the same syntactic judgments". However, such differences are generally believed to be relatively minor and of little theoretical significance: Chomsky, for instance, describes them as "marginal" and goes on to assert that they "can be safely ignored across a broad range of linguistic investigation" (1975: 18).

However, there is a growing body of evidence suggesting that such differences are in fact quite considerable and should not be ignored. Dąbrowska (2008a), for example, found that different groups of Polish speakers appear to have learned different generalizations about the distribution of genitive masculine inflections: while some used the most reliable cue, the semantic properties of the noun, others were sensitive to properties of the construction in which the noun occurred or its phonological properties. These differences were not explainable in terms of dialectal differences, although they were not, perhaps, very surprising, as this part of the inflection system is highly irregular, and therefore it is not clear what the "correct" generalization is. Another study (Dąbrowska 2008b), however, revealed large individual differences in Polish speakers' productivity with dative endings, which are almost completely regular. Adult native speakers of Polish reliably supplied the target form with nonce nouns belonging to densely populated neighbourhoods (i.e., nouns which are similar to many existing words) but had problems with nouns from sparsely populated neighbourhoods, in spite of the fact that they belonged to the same grammatical class and required the same inflection. Individual scores on nouns from sparsely populated neighbourhoods ranged from 4% to 100% and were strongly correlated with education. A smaller-scale study by Wolff (1981) reports similar finding for German participles.

Other studies report considerable differences in individual participants' ability to process certain syntactic structures. Dąbrowska (1997) found a very strong relationship between level of education and participants' ability to process the following sentence types: sentences containing complex NPs (e.g., *Paul noticed that the fact that the room was tidy surprised Shona*), the *tough* movement construction (e.g., *John will be hard to get his wife to vouch for*), and two types of sentences with parasitic gaps (e.g., *It was King Louis who the general convinced that this slave might speak to*, *The nervous-looking student that Chris met after being told his girlfriend wanted to jilt took the 11 o'clock train*). Participants had to answer simple questions about the sentences (e.g., *What did Paul notice? What surprised Shona?* for the complex NP sentence). Dąbrowska found that university lecturers outperformed university students and university students, in turn, outperformed the low academic attainment participants (in this case cleaners and porters).

It is unclear whether such differences reflect linguistic competence, or whether the low academic attainment (hereafter, LAA) participants' failure to respond correctly is attributable merely to performance factors such as working memory limitations. Although the study was designed to minimise the role of performance factors, it is undeniable that the sentences used in the study taxed the participants' processing capacities.

This issue was addressed by Chipere (2001), who tested two groups of 18-year-olds from the same school: a LAA group comprising pupils who scored ‘D’ or below in GCSE English, and a high academic attainment group (HAA) comprising pupils who scored ‘A’ in at least 5 GCSE subjects, including English. In the first phase of the experiment, participants were tested on comprehension and recall of complex NP sentences. The LAA group performed significantly worse than the HAA group on both tasks. Chipere then divided the LAA participants into two subgroups which were given different types of training. Half of the participants took part in a memory training programme in which they were asked to repeat complex NP sentences. The other half was given comprehension training which involved explicit instruction about the sentence type used in the experiment, followed by a practice session in which participants had to answer comprehension questions and were given feedback on their performance. 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, whilst 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 due to lack of experience with the particular grammatical structure used in the experiment rather than working memory capacity.

Addressing the issue of working memory limitations from a different perspective, Dąbrowska and Street (2006) investigated whether such individual differences in comprehension could be shown with a less complex construction. Using a modified version of a task developed by Ferreira (2003), they tested highly educated and less educated speakers’ comprehension of passive sentences. The passive construction was chosen for two reasons. First, it does not place such a heavy burden on working memory as the constructions used in Dąbrowska (1997) or Chipere (2001), since it does not involve embedding. Secondly, while knowledge about passives is undeniably a part of ‘core’ grammar, individual speakers differ in the amount of experience they have had with this structure. Since full passives are used predominantly in written texts, the hypothesis was that more educated speakers would perform better than less educated speakers. Dąbrowska and Street also wanted to test whether the *type* of linguistic experience matters. They therefore tested non-native speakers, the logic being that whilst educated non-native speakers have the benefits of schooling, they have quantitatively less experience with passives than native speakers. All participants were tested on four types of sentences: plausible actives (e.g., *The dog bit the man*), implausible actives (e.g., *The man bit the dog*), plausible passives (e.g., *The man was bitten by the dog*) and implausible passives (e.g., *The dog was bitten by the man*).

The experiment generated two main findings. First, native speakers sometimes process sentences non-syntactically (i.e., relying on world knowledge rather than grammatical competence). Significantly, however, it is less educated speakers who are more likely to do this, presumably as a consequence of less experience with the passive. Second, some non-native speakers use syntactic cues more reliably than less-educated native speakers. This suggests that the amount of exposure is only one of several factors, since it is unlikely that the non-graduate non-native group had encountered more passives than the non-graduate native group. Thus, it seems that the type of linguistic experience also matters.

Further evidence that some native speakers may not fully master all the core constructions of their language is provided by Brooks and Sekerina's (2005/2006) study which tested comprehension of sentences with universal quantifiers. Participants (children aged from 7 to 9 and an adult control group) were shown pairs of pictures depicting distributional scenes with objects and containers in partial one-to-one correspondence (e.g., one picture depicts three alligators in bathtubs with two extra bathtubs whilst the other depicts three alligators in bathtubs with two extra alligators). They were then presented with a sentence containing a quantifier (e.g., *Every alligator is in a bathtub* or *Every bathtub has an alligator in it*) and had to choose the correct picture. All age groups performed at a relatively low level, with the adult participants choosing the correct picture only 79% of the time. Although this level of performance is above chance, an analysis of individual responses revealed that only half of the individual scores were above chance, that is to say, only half of the adults made 10 or more correct choices on 12 trials. Furthermore, some participants showed no sensitivity to the position of the universal quantifier at all, selecting the same picture for both sentences.

The results of these studies conflict with those reported in some language acquisition studies. Pinker, Lebeaux and Frost (1987), for instance, found that children as young as four performed at ceiling on passives; Crain et al. (1996) report excellent performance (88% to 98% correct responses) on tasks involving comprehension and production of sentences with universal quantifiers in children aged from 3;5 to 5;10. Crain et al. employed so-called Truth Value Judgment (TVJ) tasks in which short stories are acted out with toys and props and watched by the participants and a puppet. The puppet summarises each story and the child indicates whether the puppet's statement is true or false. If the child indicates that the statement is incorrect, s/he is asked to explain what actually happened. Crain et al. attribute the excellent performance demonstrated by the children in their study to the testing method, arguing that TVJ provides a more accurate reflection of the underlying competence than tasks such as act-out or picture verification (in which the child has to decide whether a particular stimulus sentence is an accurate description of a picture). This conclusion is disputed by other researchers (see e.g. Drozd 2004, Philip 2004 for critiques of their experiments and Gordon 1998 for a more general discussion of problems with TVJ tasks).

It should also be pointed out that that in most studies, children performed considerably worse than in the studies just mentioned: for instance, Maratsos et al. (1985) and Gordon and Chafetz (1990) report chance (and, in some conditions, below-chance) performance on passives in the same age group; Geurts (2003: 199), in his overview of research on the comprehension of sentences with universal quantifiers observes that "error rates in excess of 50% are quite common" even in children as old as 7. While it is possible that the observed differences are attributable to differences in methodology, it seems at least as likely that they reflect existing variability: while some language learners may know all there is to know about passives and quantifier constructions by age 4, others may need considerably more time – and, as we will see, some may never fully master the construction. Moreover, there is evidence that there is a great deal more variation among children from less privileged social backgrounds (see Ginsborg 2006, Hoff 2006, Huttenlocher 1998, Huttenlocher et al. 2002, Locke and Ginsborg 2003); and since children who participate in language acquisition experiments are usually recruited from middle- and upper-middle class backgrounds, it is likely that the results described in most existing research are not representative of the entire population.[1]

The aim of the present study is two-fold: to provide further evidence demonstrating the existence of individual difference in native language attainment, and to identify possible reasons for the differences. In particular, we investigate the possibility that they are attributable to differences in language experience. To do this we focus on three constructions: the full passive (e.g., *The boy was chased by the girl*); a variant of the quantifier construction in which the universal quantifier *every* modifies a noun referring to an object located in a container (e.g., *Every cat is in a basket;* hereafter, Q-*is*) and a quantifier construction in which the universal quantifier *every* modifies the noun referring to a container in which some object is located (e.g., *Every basket has a cat in it;* hereafter, Q-*has*). A fourth construction, the active (e.g., *The boy chased the girl*), is used as a control condition.

It is important to note that that these four constructions differ in frequency. The British National Corpus (2001), a 100 million word corpus of contemporary British English, contains no instances of the Q-*has* construction (*Every NOUN has a NOUN PREP it*), 8 instances of Q-*is*

[1] It is interesting to note in this connection that the children who participated in the Pinker et al. study were recruited from day care centres affiliated with Harvard University, while those tested by Maratsos et al. and Gordon and Chaftez came from more mixed backgrounds. Crain et al. do not provide any information about the children they tested apart from age.

(*Every NOUN is PREP a NOUN),* 5675 full passives, and over 120,000 active transitives.[2] This is of particular significance to usage-based models of language acquisition which posit that more experience with a particular construction results in entrenchment of and hence better performance on that construction (Abbot-Smith and Tomasello 2003, Bybee 2006, Tomasello 2003). On this basis and taking the results of the BNC search into account, we predict that participants will find actives easier than passives; passives easier than Q-*is*; and Q-*is* easier than Q-*has*.

2. Experiment 1

Dąbrowska and Street (2006) tested comprehension of plausible and implausible active and passive sentences by adult speakers, and found that the less educated participants performed extremely poorly (below chance) on implausible passives. However, some participants also made errors on implausible actives, which suggests that their problems with passive sentences were partly attributable to reliance on pragmatic rather than syntactic cues. Moreover, the task used to test comprehension (identify the 'doer', i.e., agent) relies to some extent on metalinguistic abilities, which could disadvantage the less educated participants. The present study was designed to determine whether the education-related differences in the comprehension of passive sentences can be replicated using pragmatically neutral sentences and a different testing method (i.e., picture selection). In addition, we tested comprehension of the two quantifier constructions described earlier using the same task.

We used picture selection rather than truth value judgment since earlier research on this topic suggests that sentences with universal quantifiers are particularly prone to misinterpretation when using truth value judgment (see Brooks and Sekerina 2005/2006 for discussion). Picture selection is widely regarded as a very simple task which places minimal demands on participants' processing system, and is thus suitable for use even with children as young as two (Gerken and Shady 1998). Furthermore, studies which directly compared performance on the tasks have

[2] The last figure is an estimate based on data provided by Roland et al. (2007).

demonstrated either similar or slightly better performance on picture selection. For instance, Salis and Edwards (2009) tested the comprehension of passive sentences by aphasic patients (who are known to be very prone to variable performance) and found that they averaged 51% correct answers when tested using picture selection and 43% correct on the truth value judgment task (65% correct for matches, 20% for mismatches). They observe a similar advantage for PS on other sentence types (actives, subject clefts, and object clefts). Baauw and Zuckerman (2008) tested children using both methods and also report considerably better performance on picture selection. For instance, for simple sentences with pronouns, children gave the correct response 80% of the time when tested using picture selection and only 50% on the time when tested using truth value judgment. Thus, the picture selection task appears to be least dependent on linguistically irrelevant performance factors and hence offers the most accurate reflection of our participants' linguistic abilities.

2.1 Method

2.1.1 Participants

Fifty adults (27 males and 23 females) ranging in age from 18 to 50 participated in the experiment. 19 participants were postgraduate students at the University of Sheffield who had at least 17 years of formal education. The remaining 31 participants (the non-graduate, LAA group) had had at most 11 years of formal education and were employed as shelf-stackers, packers, assemblers, or clerical workers. All participants were native speakers of English.

2.1.2 Materials

There were three experimental conditions (passive, Q-*is* and Q-*has*) plus one control condition (active). The active and passive sentences described simple transitive events (e.g., *The sailor hit the soldier*, *The soldier was hit by the sailor*). Both sentence types are semantically reversible so that the NPs depicting the subject/object or agent/patient can be switched to give sentences describing events of similar probability (*The soldier hit the sailor*, *The sailor was hit by the soldier*). The Q-*is* and Q-*has* sentences described locative scenes in which various objects were situated in containers (e.g., *Every dog is in a basket*, *Every basket has a dog in it*).

The visual stimuli comprised twenty-four pairs of pictures. Twelve pairs of pictures depicted a simple transitive event (e.g., a sailor hitting a soldier and a soldier hitting a sailor: see Figure 1). The other twelve pairs depicted various entities arranged in containers (e.g., flowers in vases, fish in bowls, dogs in baskets). These pictures showed distributive arrangements with the entities and containers in partial one-to-one correspondence with each other. For example, one picture depicted four baskets, three of which had a dog in them whilst the remaining basket didn't; and the other picture depicted four dogs, three of which were in baskets whilst the remaining dog wasn't (see Figure 2). The location of the extra object and extra container was counterbalanced across stimuli: in other words, in half of the picture pairs the picture with the empty container appeared on the right and in the other half it appeared on the left.

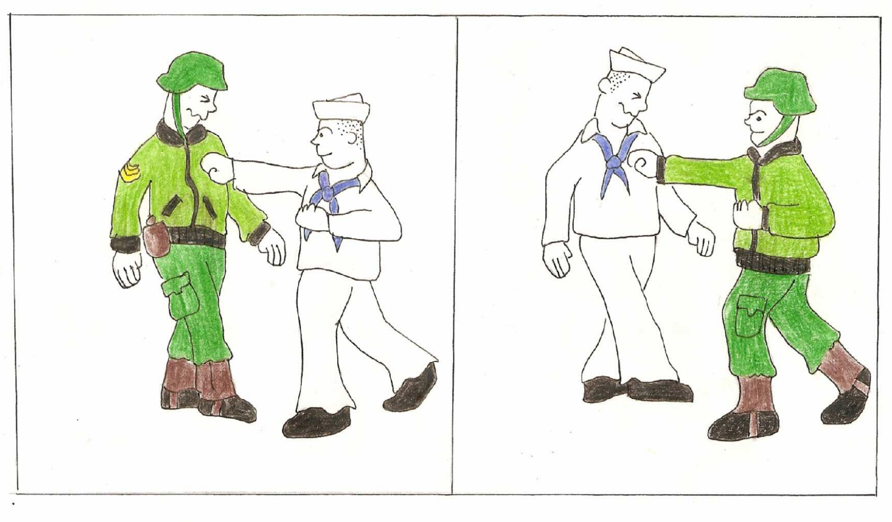

Figure 1: *Example of pictures used to test comprehension of active and passive sentences*

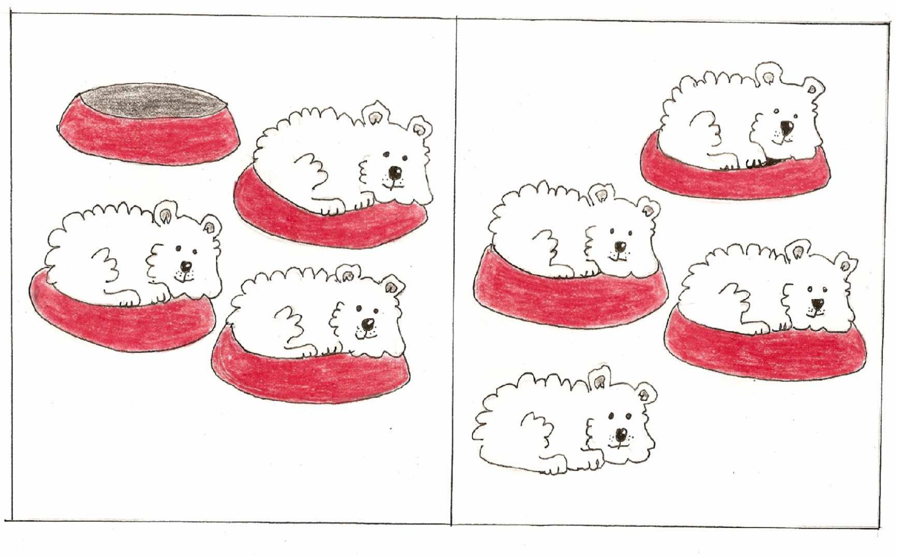

Figure 2: *Example of pictures used to test comprehension of quantifier sentences*

There were four versions of the test, each containing six sentences for each of the four conditions. For simple transitive events there are four possible descriptions (see above); therefore each of the four appeared in a different version. For descriptions of entities in containers there are only two constructions (see above); the sentences were therefore divided so that that if the Q-has variant appeared in version 1 and 3, the Q-*is* variant appeared in version 2 and 4, and vice versa. Within any one version there were no repeats of the same action involving the same participants or the same entities in containers. The sentences in each version were presented in a semi-random order (i.e., no items belonging to the same condition appeared immediately next to each other). A complete list of sentences used in one version of the test is given in the appendix.

2.1.3 Procedure

Participants were tested individually in a familiar setting at the place where they worked or studied. They were asked to listen carefully to each sentence and then select the matching picture. In the active and passive conditions, the participant was shown a pair of pictures such as those in Figure 1, and then heard one of the following sentences: *The soldier hit the sailor, The sailor hit the soldier, The soldier was hit by the sailor,* or *The sailor was hit by the soldier.* For the quantifier sentences, the participant was shown a pair of pictures depicting entities and containers in partial one-to-one correspondence such as those in Figure 2, and then heard either *Every dog is in a basket* or *Every basket has a dog in it*. If a participant said that neither picture matched the sentence, they were asked to "choose the one that fits the sentence better than the other one"; if they continued to insist that neither picture corresponded to the test sentence, the response was scored as incorrect. Since the distractor picture depicted a scene which definitely did *not* match the scene described in the stimulus sentence, this method ensured that participants should choose the target picture if their grammars allowed this interpretation, even if it was a dispreferred one. For instance, even if a speaker prefers the symmetrical interpretation of *Every dog is in a basket* (i.e., one in which there is a one-to-one mapping between dogs and baskets), on being confronted with this sentence and the two pictures Figure 2, s/he ought to choose the picture on the left, since the picture on the right clearly does not match the sentence. Approximately one-quarter of the participants in each group were tested with each version of the test.

2.1 Results

The results of the experiment are summarised in Table 1. As can be seen from the table the graduate (HAA) group performed at ceiling in all conditions and therefore their results will not be analysed any further. The non-graduate (LAA) participants also performed at ceiling in the active condition, but had considerably lower scores in the other conditions. Since the results are clearly not normally distributed, the data were analysed using non-parametric tests (Friedman's ANOVA, Wilcoxon Signed Ranks). All reported significance levels have been corrected for multiple comparisons.

Table 1: *Proportion of correct responses (Experiment 1)*

Condition	High academic attainment (*N*=19)	Low academic attainment (*N*=31)
Active		
Mean (SD)	100 (0)	97 (6)
Median	100	100
Min-Max	100-100	83-100
Passive		
Mean (SD)	100 (0)	88 (18)
Median	100	100
Min-Max	100-100	33-100
Q-*is*		
Mean (SD)	100 (0)	78 (24)
Median	100	83
Min-max	100-100	0-100
Q-*has*		
Mean (SD)	100 (0)	43 (30)
Median	100	33
Min-Max	100-100	0-100

A Friedman's ANOVA was calculated for the LAA participants' scores on the four test conditions. The results were significant: χ^2 (3) = 54.03; p <0.001. Follow up pairwise Wilcoxon tests revealed that the LAA participants were at ceiling on actives and at chance on Q-*has* sentences. Performance on passives was significantly worse than on actives (z = –2.62, p = 0.026, N=31, r = 0.24) but better than on Q-*is* (z = - 4.28, p < 0.001, N=31, r = 0.38); and performance on Q-*is* was better than on Q-*has* (z = –4.18, p < 0.001, N=31, r = 0.37). This order of difficulty reflects the corpus frequencies of the four constructions, and hence supports the usage-based view that mastery is a function of amount of experience.

As a group, the LAA participants performed relatively well in the passive condition. However, if we consider individual performance, a somewhat different picture emerges. According to the binomial distribution ($p < 0.05$), above chance performance requires 6 out of 6 correct picture selections. At this criterion, 19 LAA participants (61%) performed above chance. Adopting a more lenient criterion (at least 5 out of 6), 24 LAA participants (77%) performed above chance (see Figure 3).

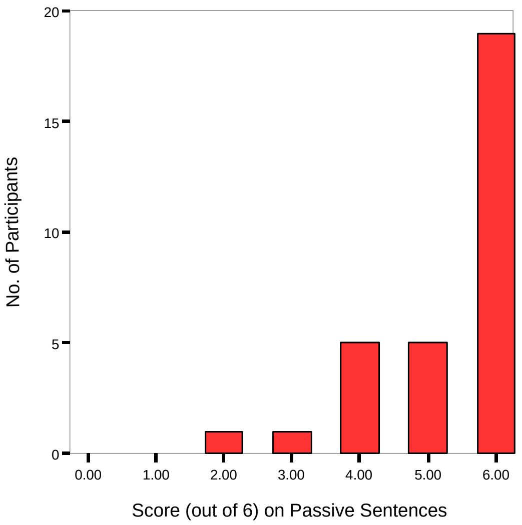

Figure 3: *Distribution of individual LAA participants' scores in the passive condition*

A similar picture emerges with the Q-*is* construction. While group performance was quite good, only 11 participants (35.5%) performed above chance, i.e., answered all six questions correctly. If we use the more lenient criterion (at least 5 out of 6 correct), 21 participants, i.e., 68%, were above chance (see Figure 4).

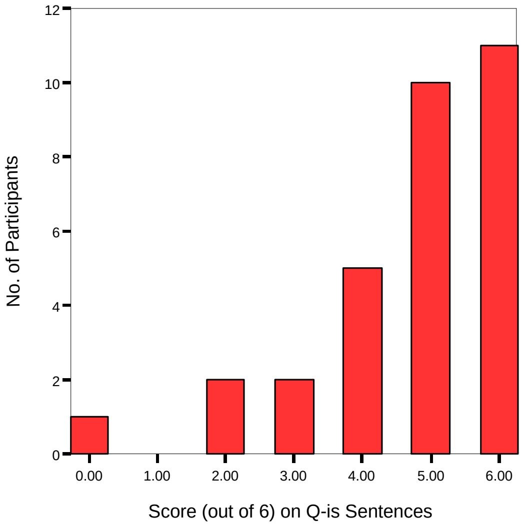

Figure 4: *Distribution of individual LAA participants' scores in the Q-*is *condition*

The LAA group had the most problems with the Q-*has* construction, performing as a group at about chance level. Only three participants (9.7%) performed above chance. If we adopt the more lenient criterion, six participants (19.3%) were above chance (see Figure 5).

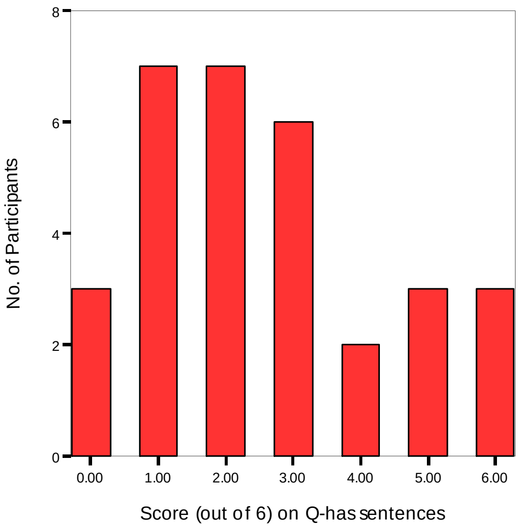

Figure 5: *Distribution of individual LAA participants' scores in the Q-has condition*

It is worth observing that four participants (one in the Q-*is* condition and three in the Q-*has* condition) refused to select either picture for the quantifier sentences, even after the second prompt. These participants probably had a different understanding of such structures, one which requires a one-to-one mapping between objects and containers. In this special case, *Every OBJECT is in a CONTAINER* is synonymous with *Every CONTAINER has an OBJECT in it.*

Clearly LAA participants have some knowledge of passive and Q-*is* constructions: their performance as a group is above chance. However, such knowledge is not as well entrenched as that of actives. In addition, even when group results are relatively good, some individuals are performing at chance, or even below chance. Experiment 1 therefore confirms the existence of individual differences with regard to passives and quantifier constructions. Furthermore, the difficulties the LAA participants experienced are fully compatible with the order predicted on the basis of frequency.

2.3 Discussion

Our results suggest large education-related differences in the ability to correctly interpret passive sentences and sentences with the universal quantifier *every*, with the highly educated participants consistently interpreting all sentences correctly, and considerable individual variation in the less-educated group. Do these differences reflect differences in underlying linguistic knowledge, 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 results are highly unsatisfactory. The interviews were conducted at the place where the participants worked and were as informal as possible. Participants had as much time as was necessary to answer the questions (although most completed the task in less than five minutes), and were all extremely co-operative. Second, issues surrounding ‘test-wiseness’ should be evident across all constructions; yet the LAA group performed at ceiling on the control condition (i.e., the actives). Thirdly, as argued earlier, a picture selection task places minimal demands on linguistically irrelevant performance factors, and in fact children as young as two generally succeed on this task when presented with simple sentences that they can understand. Of course, one can always argue that the participants' difficulties were due to non-linguistic task demands and that they might succeed if tested using a different method; however, such an argument is vacuous until it is demonstrated that participants who fail on picture selection succeed on some other task tapping knowledge of the same constructions. We conclude, therefore, that the participants' relatively poor performance on passives and quantifiers reveals incomplete knowledge of these constructions.

3. Experiment 2

According to usage-based models of language (Barlow and Kemmer 2000, Langacker 2000, Bybee 2006), structure emerges from use: in other words, linguistic knowledge is shaped by usage factors such as the frequency of a particular unit and speakers' perception of its similarity to other units. The input that language learners are exposed to contains many recurrent patterns – that is to say, specific forms are associated with specific meanings – and learners extract schemas capturing these patterns. These become entrenched (and hence more easily accessible) through repeated use. Thus, inability to produce or understand a particular construction may be attributable to insufficient experience, and therefore, additional experience with the construction should result in an improvement in performance. Experiment 2 was designed to test this prediction.

The experiment tested comprehension of the same constructions as Experiment 1 before and after training: half of the participants received training on the passive and the other half on the more difficult of the two quantifier constructions, Q-*has*. An earlier training study by Chipere (2001: see Introduction) demonstrated that the low academic attainment students who had problems with complex NP sentences performed at ceiling on the same construction after just a few minutes' training. However, Chipere only tested his subjects immediately after the practice session, so we do not know whether the effects lasted more than a few hours. Our design incorporates a series of post-tests in order to determine whether the effects of additional experience are long-lasting.

Our hypothesis is that training will lead to selective improvement in performance: participants in the passive training group will improve in performance on passive sentences but not on sentences with quantifiers, whilst participants in the quantifier training group will improve in performance on quantifier sentences but not on passives. As the quantifier training group were only instructed on the Q-*has* construction, a further question is whether the effects of training will generalise to the other quantifier construction. If learners are highly conservative, they might show improvement on the trained construction only; otherwise, performance will improve on both variants of the construction.

Finally, in order to collect additional data on the possible causes for individual differences in language attainment, the study also included a reading questionnaire and a need-for-cognition questionnaire.

3.1 Method

3.1.1 Participants

54 adults (33 women, 21 men) aged 16-50 participated in the experiment. Participants were recruited from a local college with the help of their teachers. All were enrolled in Skills For Life classes aimed at adults who have problems with literacy and numeracy. Skills For Life programmes comprise 5 levels (3 entry levels plus Levels 1 and 2). Entry levels 1–3 are intended to teach basic skills such as reading a newspaper article or instruction manual and writing a letter to a utility company. Level 1 is equivalent to a GCSE (General Certificate of Secondary Education) pass, and Level 2 to a good pass (C or above) at GCSE.[3] As only individuals with no

---

[3] GCSE (General Certificate of Secondary Education) is the name of a set of British examinations, usually taken by secondary school students in England, Wales and Northern

formal qualifications are allowed to enrol for Skills for Life, all participants are considered to have low academic attainment. However, all participants who had been identified by their teachers as having learning difficulties were excluded from the sample.

3.1.2 Materials

Four versions of the test were created. Version 1 contained the same sentences as the test used in Experiment 1. The other three versions contained the same verbs and universal quantifiers but with different NPs. For example, the test sentence *The boy chased the girl* was replaced with *The man chased the woman*, *The dancer chased the chef,* and *The soldier chased the fireman* in versions 2, 3 and 4; and *Every apple is in a dish* was replaced with *Every orange is in a dish, Every banana is in a dish,* and *Every pear is in a dish.*

3.1.3 Procedure

As in Experiment 1, participants were tested individually in a familiar setting at the place where they studied. The experiment comprised six stages:

(1) pre-test

(2) training (conducted one week after the pre-test);

(3) post-test 1 (conducted immediately after training);

(4) post-test 2 (conducted one week after training);

Ireland at age 15-16. A different examination is taken for each area of study, but school students are obliged to take examinations for “core subjects” (English language, English literature, mathematics, and science) along with several optional subjects. At the end of the two-year GCSE course, each student receives a grade for each subject. The pass grades, from best to worst, are: A*, then A – G. Receiving five or more A*-C grades is often a requirement for taking A-levels after leaving secondary school. Most universities typically require a C or better in English and Mathematics, regardless of a student's performance in their A-level.

(5) post-test 3 (conducted approximately 12 weeks after training); and

(6) reading and need-for-cognition questionnaire.

The pre-test was used to select participants who achieved low scores (no more than 4 correct out of a maximum score of 6) in all three experimental conditions. 17 of the participants tested met this criterion and were randomly assigned to a passive training group or a quantifier training group. Participants were then given a brief non-technical explanation of the construction in question. Participants in the passive training group were shown two sentence types: Type 1 was active (*The boy chased the girl*); type 2 was passive (*The boy was chased by the girl*). Participants then heard the following explanation:

> "In the first sentence there are two people involved in an activity. The person who does the action (the doer) appears first (before the action word *chase*). The person affected by the action appears second after the action word *chase*. This sentence matches the picture in which a boy is chasing a girl and the girl is running away. The second sentence is different because in this sentence the person who does the action (the doer) appears after the word *by* but the person who is affected by the action appears first, before the action word *chase*. This sentence matches the picture in which the girl is chasing the boy and the boy is running away."

Participants in the quantifier training group were trained on the Q-*has* variant only. Participants were presented with a sentence such as *Every basket has a dog in it* and heard the following explanation:

> "In this type of sentence there are two things: a basket and a dog. The word *every* refers only to the thing which follows it; in this case *basket*. It is similar to saying *All the baskets have a dog in them*. If all the baskets in a picture have a dog in them, then this picture

matches the sentence. But if just one basket does not have a dog in it, then the picture does not match the sentence."

Participants were then asked to redo the six items from the pre-test exemplifying the construction that had just been explained (either the passives or the *Q-has* sentences, depending on which training group they were in). If a participant made an error, the structure was explained to him/her again and the correct picture was indicated. Immediately after the training session, the experimenter administered post-test 1. No feedback was provided during this or the following two post-tests.

Each participant selected for the training study completed all four versions of the test, with the order of the versions counterbalanced across participants and stages using a balanced Latin square design in which each version precedes and follows every other version equally often.

All participants who took part in the pre-test were also asked to complete a brief reading questionnaire (described more fully in Street in preparation) and the short version of the need for cognition questionnaire (Cacioppo, Petty and Kao 1984) which measures how much people enjoy undertaking cognitively challenging tasks. The key question on the reading questionnaire was "How much time did you spend reading books, newspapers, magazines, etc. for pleasure (not college work) last week?". Participants were asked to choose from the following options: no more than 15 minutes; 15-60 minutes; 1-4 hours; 4-10 hours; more than 10 hours. Participants who took part in the training study completed these immediately after post-test 3; the remaining participants did so in a separate session about three months after the pre-test.

3.2 Results and discussion

3.2.1 Pre-test

The results of the pre-test are summarised in Table 2 and mirror the results of the LAA group from Experiment 1. All participants are at or close to ceiling on the active (control) condition but there are vast individual differences on the passive and quantifier constructions. As in Experiment 1, the data are not normally distributed, so they were analysed using nonparametric tests (Friedman's ANOVA , Wilcoxon Signed Ranks). All reported significance levels in this and the following section have been corrected for multiple comparisons.

As in Experiment 1, a Friedman's ANOVA was calculated for the LAA participants' scores on the four test conditions. The results were significant: χ^2 (3) = 88.46; $p < 0.001$. Follow-up pairwise Wilcoxon tests revealed that performance on passives was significantly worse than on actives ($z = -4.92$, $p < 0.001$, $N=54$, $r = 0.33$) but better than on Q-*is* ($z = -4.28$, $p < 0.001$, $N=54$, $r = 0.29$); and performance in the latter condition was significantly better than with Q-*has* sentences ($z = -3.68$, $p < 0.001$, $N=54$, $r = 0.25$).

Table 2: *Proportion of correct responses on the pre-test (Experiment 2; N=54)*

Condition	% correct
Active	
Mean (SD)	99 (4)
Median	100
Min-Max	83-100
Passive	
Mean (SD)	79 (21)
Median	83
Min-Max	17-100
Q-*is*	
Mean (SD)	71 (33)
Median	83
Min-max	0-100
Q-*has*	
Mean (SD)	53 (36)
Median	50
Min-Max	0-100

3.2.2 Post-tests

Information about performance on the three post-tests, as well as the selected participants' performance on the pre-test, is given in Table 3.

Table 3: *Proportion of correct responses before and after training (Experiment 2)*

	Passive training group (N=8)				Quantifier training group (N=9)			
	Pre	Post-1	Post-2	Post-3	Pre	Post-1	Post-2	Post-3
Active								
Mean (SD)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)
Median	100	100	100	100	100	100	100	100
Min-Max	100-100	100-100	100-100	100-100	100-100	100-100	100-100	100-100
Passive								
Mean (SD)	48 (16)	98 (5)	94 (9)	93 (9)	61 (12)	59 (14)	67 (8)	62 (8)
Median	50	100	100	100	66	67	67	67
Min-Max	17-67	83-100	83-100	83-100	33-67	33-83	50-83	50-67
Q-*is*								
Mean (SD)	44 (38)	29 (32)	35 (33)	38 (27)	37 (32)	94 (12)	94 (8)	97 (6)
Median	42	25	41	50	50	100	100	100
Min-max	0-100	0-83	0-83	0-67	0-67	67-100	83-100	83-100
Q-*has*								
Mean (SD)	15 (16)	15 (19)	20 (23)	28 (23)	13 (20)	100 (0)	91 (12)	92 (9)
Median	17	9	16	33	0	100	100	100
Min-Max	0-50	0-50	0-50	0-66	0-50	100-100	67-100	83-100

Note: Three participants (two from the quantifier group and one from the passive group) withdrew from the experiment after post-test 2. Thus there were only seven participants in each group in post-test 3.

The results of the post-tests indicate a clear improvement in performance on the trained construction. In post-test 1, the passive training group performed at ceiling in that condition but showed no improvement on the quantifier constructions. Conversely, participants in the quantifier training group were at ceiling in that condition but showed no improvement on the passive construction. Interestingly, these participants were also at ceiling on the Q-*is* construction in which they received no instruction. This suggests that participants have generalised from one quantifier construction to the other. Wilcoxon Signed Ranks tests confirm this interpretation. Participants in the passive training group still performed significantly better on active sentences than on Q-*is* ($z = –2.53$, $p = 0.033$, $N=8$, $r = 0.44$) and Q-*has* sentences ($z = –2.55$, $p = 0.033$, $N=8$, $r = 0.45$). However, there is no significant difference between their performance on actives and passives ($z = –1$, $p = 0.951$, $N=8$, $r = 0.17$). By contrast, participants in the quantifier training group still perform significantly better on active sentences than on passives ($z = –2.69$, $p = 0.021$, $N=9$, $r = 0.44$), but there is no significant difference between their performance on actives and Q-*is* ($z = –1.34$, $p = 0.54$, $N=9$, $r = 0.22$) or Q-*has* sentences ($z = 0.00$, $p = 1$, $N=9$, $r = 0$).

The results of the later post-tests show a similar pattern, with only a very small decline in performance between post test 1 and post-tests 2 and 3. In post-test 2, participants in the passive training group still performed significantly better on active constructions than Q-*is* ($z = –2.53$, $p = 0.033$, $N=8$, $r = 0.44$) and Q-*has* constructions ($z = –2.55$, $p = 0.033$, $N=8$, $r = 0.45$). However, there were no significant difference between their performance on active and passive constructions ($z = –1.73$, $p = 1$ $N = 8$, $r = 0.30$). Similarly, participants in the quantifier training group performed significantly better on active constructions than passive constructions ($z = -2.80$, $p = 0.015$, $N=9$, $r = 0.46$), but there was no significant difference between their performance on active constructions and Q-*is* constructions ($z = –1.73$, $p < 0.25$, $N=9$, $r = 0.28$) or Q-*has* constructions ($z = -1.89$, $p = 0.18$, $N=9$, $r = 0.31$).

In post-test 3, participants in the passive training group still performed significantly better on actives than Q-*is* ($z = –2.41$, $p = 0.048$, $N=7$, $r = 0.45$) and Q-*has* sentences ($z = –2.41$, $p =$

0.048, N=7, r = 0.45) whilst there was no significant difference between their performance on actives and passives (z = –1.73, p = 1, N=7, r = 0.32). Participants in the quantifier training group continued to perform significantly better on actives than on passives (z = –2.46, p = 0.042, N=7, r = 0.46), whilst there was no difference in performance on active and Q-*is* (z = –1.00, p = 0.93, N=7, r = 0.18) or Q-*has* sentences (z = -1.73, p = 0.95, N=7, r = 0.32). These results indicate that the effects of training were long-lasting: participants are still close to ceiling on the construction in which they had been trained even 12 weeks post-training.

Table 4: *Changes in performance during Experiment 2* (Friedman's ANOVA)

Condition	Passive training group		Quantifier training group	
	χ^2	*p*	χ^2	*p*
Passive	18.00	<0.001	3.21	0.347
Q-*is*	0.58	0.901	18.00	<0.001
Q-*has*	2.61	0.456	16.65	0.001

Performance on the four tests (for each group separately) was further analysed using a Friedman test. The results are summarised in Table 4. In the passive training group there was a significant change in performance on the passive but not in the Q-*is* or Q-*has* condition. In the quantifier training group, on the other hand, there was a significant change in performance on **both** types of quantifier sentences but not on the passive. Thus, participants in the quantifier training group generalized to the other quantifier sentence type, in spite of the fact that they had not received training for sentences of the form *Every NOUN is in a NOUN.* A possible explanation for this is that the two sentence types are variants of the same construction – or, more precisely, combinations of the quantifier construction with the locative construction (NP BE PREP NP) and the 'possessive locative' (NP HAVE NP PREP *it*). However, the differences in performance on Q-*is* and Q-*has* on the pre-test would argue against such an interpretation.

Another possibility is that the participants were able to draw inferences about the meaning of the untrained construction because it was implicitly contrasted with Q-*has* in the test stage.

3.2.3 Reading and need for cognition

As explained in the Method section, we also asked participants to complete reading and need-for-cognition questionnaires, and computed correlations between these measures and overall performance on the test (see table 5). The reading measure which correlated most robustly with the overall test score was amount of reading (rho = 0.551, p <0.001, N = 47). Need for cognition had a similar effect on the overall test score (rho = 0.576, p <0.001, N = 47). A correlation between scores on the need for cognition test and the amount of reading was also computed (rho = 0.370; p = 0.010, N = 47).

Table 5: *Correlations (Spearman's rho) between performance on experimental task and reading and need for cognition*

Comprehension measure	Amount of reading	Need for cognition
Overall score	0.551***	0.576***
Passive sentences	0.529***	0.404**
Quantifier sentences	0.520***	0.606***

** $p \leq 0.01$; *** $p \leq 0.001$

We also computed correlations between these two measures and performance on the passive and the two types of quantifier sentences. This analysis revealed that performance on the passive was more strongly correlated with amount of reading (rho = 0.529, p < 0.001, N = 47) than with need for cognition (rho = 0.404, p = 0.005, N = 47), whilst for sentences with quantifiers, need for cognition was the better predictor (rho = 0.606, $p < 0.001$, N = 47; for quantifier score and amount of reading, rho = 0.520, $p < 0.001$, N = 47). Z-tests for two correlation coefficients (Kanji 2006: 42) show that these differences between the correlation coefficients are significant (passive score: $z = 3.53$, $p < 0.001$; quantifier score: $z = 2.78$, $p = 0.005$).

These figures suggest that exposure to written language may be a factor that contributes to the development of the passive construction (presumably because passives are relatively frequent in written texts) whilst need for cognition is more relevant for the development of knowledge about quantifiers (possibly because quantifiers play an important role in logical reasoning: see Braine and O'Brien 1998).

4. General discussion

Experiment 1 demonstrated the existence of considerable individual differences in adult native speaker's comprehension of passive and universal quantifier constructions. Furthermore, such individual differences are strongly associated with educational attainment. Whereas the HAA participants performed at ceiling in all conditions, the LAA group performed at ceiling only in the active condition. Their performance on passive sentences – 88% correct – was also relatively good, but well below ceiling. Performance on Q-*is* sentences was considerably worse (78% correct), and performance in the Q-*has* condition (43% correct) was at chance. Nevertheless, even in this group some participants performed at ceiling in all conditions. This indicates that, while there is a strong relationship language attainment and education, education cannot be the only factor. Significantly, the order of difficulty of these constructions (active > passive > Q-*is* > Q-*has*) reflects their frequency of occurrence. This is consistent with the idea that the differences between structures are due to amount of experience.

These findings were replicated with a different group of participants in Experiment 2. Experiment 2 also showed that training resulted in significant improvement on the structure trained but not on the other (although there was transfer from Q-*has* to Q-*is*). This provides very strong evidence that constructional schemas underlying our ability to produce and understand new sentences emerge as a result of experience. It also confirms that the poor performance on the pre-test was *not* due to inability or unwillingness to attend to the task, working memory capacity, or some other linguistically-irrelevant factor.

How can we explain the education-related differences observed in the first study? One possibility is that less educated speakers have less relevant experience than the more educated group. This is a plausible explanation for their problems with passives, since full passives occur more frequently in formal written texts and more educated speakers have more experience with such texts.[4] However, it is not clear that this also applies to sentences with quantifiers. The two specific structures tested (*Every NOUN is in a NOUN, Every NOUN has a NOUN in it*) are too rare to allow any meaningful comparisons across varieties; but the overall frequency of the word *every* is, if anything, higher in speech than in writing.[5]

A second possibility is that more educated speakers have more exposure to language overall. Most graduate students come from middle-class backgrounds, and there is evidence that middle-class parents talk more to their children than working-class parents (see, for example, Hart and Risley 1995, Hoff 2006). It is also possible that graduate students have richer linguistic experience in later life – in particular, they are likely to have had more exposure to written

[4] According to Roland et al. (2007), "a verb phrase in a written corpus such as Brown or Wall Street Journal is four or five times more likely to be passive than one in a spoken corpus such as Switchboard" (p.17). Our own searches of the BNC indicate that mean frequency of the full passive is about 63 per million words in written texts, compared to about 9 per million in speech – a sevenfold difference in frequency.

[5] There are, on average, 475 occurrences of *every* per million words in the spoken part of the BNC. For written texts, the figures vary from 241 (academic prose) to 490 (fiction).

language, and they are more likely to be skilled readers. In normal adult conversation, about 150 words are spoken every minute (Huggins 1964, Maclay and Osgood 1959); skilled readers, on the other hand, are able to process between 200 and 400 words per minute (Pressley 2006, Rayner and Pollatsek 1989). Thus, skilled readers are able to absorb more language per unit of time than skilled conversationalists.

It is also possible the participants' linguistic experience was qualitatively different: specifically, the more educated participants may have been exposed to more explicit explanations of various linguistic phenomena, which may have had an effect on their linguistic development. We know from work on second language acquisition that explicit instruction can 'jump start' implicit learning (see Ellis 2005). Similar process may also be at work in first language acquisition. The training stage in Experiment 2 involved a very brief non-technical explanation of the construction in question (see section 3.1.3), followed by practice with feedback. And yet after this minimal amount of focussing, many participants did not make a single error during the practice session. Indeed several participants reported a 'eureka' experience as soon as the particular construction was explained during the 'grammar lesson'.[6] These participants claimed that whereas in the pre-test they had simply guessed, they now knew what the correct answer was – and their performance corroborates this. It is conceivable that participants in the more educated group had parents (or teachers) who were more likely to draw their attention to specific aspects of form-meaning mappings, thus triggering the eureka experience earlier in life.

Finally, it is also possible that the participants in the less educated group were less efficient language learners, and hence needed more experience to reach the same level of attainment. We do not have data on the participants' verbal or nonverbal abilities; but since IQ is known to correlate strongly with educational achievement, it would be surprising if the HAA participants did not score higher on cognitive tests than the LAA participants, who all held

[6] Chipere (2001) observed something very similar in his training study.

relatively low-skill jobs. The results of Experiment 2 suggest a moderately strong association between performance on the comprehension task and the amount of time a person devotes to reading as well as willingness to engage in cognitively-demanding tasks. Since the results are purely correlational, we must be careful about drawing causal inferences; but both of these are plausible factors contributing to the group differences observed in Experiment 1. It is, of course, worth noting that the various explanations are not mutually exclusive. It may well be the case that the LAA participants need more experience with language, and get less.

We have argued that the differences between the highly educated and less educated participants cannot be attributed to non-linguistic factors such as working memory capacity, willingness to co-operate, and so on. Does this mean that they reflect differences in linguistic competence? It is difficult to give an unequivocal answer to this question, as it depends on some additional assumptions about the relationship between grammar and the processing system.

For theories which assume that the parser is essentially an implementation of the grammar – i.e., usage-based models such as those described in Barlow and Kemmer (2000), Bybee (2006), or Langacker (2000), but also some nativist approaches, e.g. Crain and Thornton (1998) – the answer must be affirmative. Since usage-based models predict frequency effects and individual differences attributable to experience, our results provide strong support for this approach to language. They are, however, problematic for Crain and Thornton, who maintain that all native speakers converge on the same mental grammar.

Generative theories which assume a more autonomous parser may be able to accommodate our results by attributing them to individual differences in parsing ability rather than grammar. Such an explanation would require an additional assumption, namely, that the parser is sensitive to frequency; this, however, does not appear to be particularly problematic. In fact, our results could be easily accommodated in two-stage processing theories such as LAST (Late Assignment of Syntax Theory; see Townsend & Bever 2001) and other "analysis-by-synthesis" models. Townsend and Bever suggest that sentence processing involved two distinct phases.[7] First, the processing system constructs a “pseudoparse”, a rough analysis based on superficial probabilistic cues and heuristics. The psuedoparse is then used to guide the true parse, an algorithmic process which accesses syntactic knowledge to construct a complete syntactic representation. Crucially, constructing the true parse is slower and computationally more demanding, and thus may not be carried out in certain circumstances (see also Ferreira 2003). Our results show that this is most likely to happen in less educated speakers, but also that changes to the processing system can occur relatively fast, even in adulthood.

5. Conclusion

Nativist theories of language acquisition are predicated on the claim that all speakers converge on essentially the same grammar, despite differences in experience. This is generally assumed to be self-evident, and therefore no empirical evidence is offered to support the claim. The results described here, and the research summarised in the introduction, show that convergence is not so self-evident and therefore is something that we cannot take for granted: important areas of ‘core’ grammar may not be fully mastered by some speakers, even by adulthood. The fact that our participants’ comprehension of the structures tested improved dramatically after training suggests that their earlier failure was due to lack of relevant experience, and that learners need more experience, or a different type of experience, than nativist theories usually assume is necessary.

The research reported in this paper does not invalidate nativist theories. As pointed out earlier, our results could be accommodated by theories which assume an autonomous parsing mechanism governed by its own principles. It is also possible, in principle, that individuals differ in their innate linguistic endowment – although it would make the term “universal grammar” a misnomer. However, the existence of substantial individual differences in native language attainment does raise doubts about one of the most widely accepted arguments for an innate

[7] We are grateful to an anonymous *Lingua* referee for this suggestion.

Universal Grammar: the “fact” that all native speakers of a language converge on the essentially same grammar. This may be true – but it needs to be demonstrated rather than simply assumed.

Acknowledgments: The authors would like to thank the staff and students at the Sheffield College (Norton), management and staff at Tesco supermarket, Sheffield (Abbeydale), and management and staff at Walton's, Chesterfield. James A. Street would like to thank the AHRC for funding (Award no. 06/125939) which supported this research.

Appendix: List of sentences used in one version of the comprehension task

Actives

The boy photographed the girl.

The soldier grabbed the sailor.

The man carried the woman.

The girl fed the boy.

The sailor hit the soldier.

The soldier pushed the sailor.

Passives

The girl was hugged by the boy.

The woman was chased by the man.

The woman was pulled by the man.

The soldier was frightened by the sailor.

The sailor was kicked by the soldier.

The man was kissed by the woman.

Q-*is*

Every umbrella is in a stand.

Every feather is in a vase.

Every toothbrush is in a mug.

Every ball is in a box.

Every pencil is in a jar.

Every cake is in a tin.

Q-*has*

Every shoe has a hamster in it.

Every bowl has a turtle in it.

Every cone has an ice cream in it.

Every pot has a windmill in it.

Every basket has a dog in it.

Every dish has an orange in it.

References

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

Baauw, S., Zuckerman, S., 2008. Principle B Delays as a processing problem: Evidence from task effects. Paper presented at Romance Turn 3, University of Southampton, 18-20 September 2008.

Barlow, M., Kemmer, S., 2000. Usage-Based Models of Language, Cambridge University Press, Cambridge.

Braine, M. D. S., O'Brien, D. P. (Eds.), 1998. Mental Logic. Erlbaum, Mahwah, NJ.

British National Corpus, The, version 2 (BNC World). 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Brooks, P. J., Sekerina, I., 2005/2006. Shortcuts to quantifier interpretation in children and adults. Language Acquisition 13, 177-206.

Bybee, J., 2006. From usage to grammar: the mind's response to repetition. Language 82, 529-551.

Cacioppo, J. T., Petty, R. E., Kao, C. F., 1984. The efficient assessment of need for cognition. Journal of Personality Assessment 48, 306-307.

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

Chomsky, N., 1975. Reflections on Language. Pantheon, New York.

Crain, S., Lillo-Martin, D., 1999. An Introduction to Linguistic Theory and Language Acquisition. Blackwell, Malden, MA.

Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., Woodams, E., 1996. Quantification without qualification. Language Acquisition 5, 83-153.

Crain, S., Thornton, R., 1998. Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics, The MIT Press, Cambridge, MA.

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.

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

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

Dąbrowska, E., 2008b. The effects of frequency and neighbourhood 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.

Drozd, K. F., 2004. Learnability and linguistic performance. Journal of Child Language 31, 431-457.

Ellis, N. C., 2005. At the interface: dynamic interactions of explicit and implicit language knowledge. Studies in Second Language Acquisition 27, 305-352.

Ferreira, F., 2003. The misinterpretation of noncanonical sentences. Cognitive Psychology 47, 164-203.

Gerken, L., Shady, M. E., 1998. The picture selection task. In: McDaniel, D. , McKee, C., Cairns, H. (Eds.), Methods for Assessing Children's syntax, MIT Press, Cambridge, MA, pp. 125-145.

Geurts, B., 2003. Quantifying kids. Language Acquisition 11, 197-218.

Ginsborg, J. 2006. The effects of socio-economic status on children's language acquisition and use. In: Clegg, J., Ginsborg, J. (Eds.), Language and Social Disadvantage, John Wiley & Sons, Chichester, pp. 9-27.

Gordon, P., 1998. The truth-value judgment task. In: McDaniel, D., McKee, C., Cairns, H. (Eds.), Methods for Assessing Children's syntax, MIT Press, Cambridge, MA, pp. 211-231.

Gordon, P., Chafetz, J., 1990. Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. Cognition, 36, 227-254.

Hart, B., Risley, T. R., 1995. Meaningful Differences in the Everyday Experience of Young American Children. Paul Brooks, Baltimore.

Hermon, G., 2002. General introduction: What can acquisition teach us about our language faculty? Available from http://www.ling.udel.edu/bruening/Courses/2001-2/496/hermonHO.pdf.

Hoff, E. 2006. How social contexts support and shape language development. Developmental Review 26, 55-88.

Huggins, A. W. F., 1964. Distortion of the temporal pattern of speech: Interruption and alternation. Journal of the Acoustical Society of America 36, 1055-1064.

Huttenlocher, J., 1998. Language input and language growth. Preventive Medicine 27, 195-199.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., Levine, S., 2002. Language input and child syntax. Cognitive Psychology 45, 337-374.

Kanji, G. K., 2000. 100 Statistical Tests, Third ed. Sage, London.

Kayne, R. S., 2000. Parameters and Universals. Oxford University Press, Oxford.

Langacker, R. W., 2000. A dynamic usage-based model. In: Barlow, M., Kemmer, S. (Eds.), Usage-Based Models of Language, CSLI Publications, Stanford, CA, pp. 1-63.

Locke, A., Ginsborg, J., 2003. Spoken language in the early years: the cognitive and linguistic development of three- to five-year-old children from socio-economically deprived backgrounds. Educational and Child Psychology 20, 68-79.

Maclay, H., Osgood, C. E., 1959. Hesitation phenomena in spontaneous English speech. Word 15, 19-44.

Maratsos, M., Fox, D. E. C., Becker, J. A., Chalkley, M. A., 1985. Some restrictions on children's passives. Cognition 19, 167-191.

Nowak, M. A., Komarova, N. L., Niyogi, P., 2001. Evolution of Universal Grammar. Science 291, 114 - 118.

Philip, W., 2004. Clarity of purpose in L1 acquisition research: a response to Ken Drozd's Learnability and linguistic performance. Journal of Child Language 31, 496-499.

Pinker, S., Lebeaux, D. S., Frost, L. A., 1987. Productivity and constraints in the acquisition of the passive. Cognition 26, 195-267.

Pressley, M., 2006. Reading instruction that works: The case for balanced teaching. Guilford Press, New York.

Rayner, K., Pollatsek, A., 1989. The Psychology of Reading. Prentice-Hall, Englewood Cliffs, NJ.

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.

Salis, C., Edwards, S., 2009. Tests of syntactic comprehension in aphasia: An investigation of task effects. Aphasiology 23, 1-16.

Seidenberg, M. S., 1997. Language acquisition and use: learning and applying probabilistic constraints. Science 275, 1603.

Street, J. A., Individual differences in comprehension of passives and universal quantifiers by adult native speakers of English. Unpublished PhD Thesis.

Tomasello, M. 2003. Constructing a Language: A Usage-Based Theory of Child Language Acquisition. Harvard University Press, Cambridge, MA.

Townsend, D. J., Bever, T. G., 2001. Sentence Comprehension: The Integration of Habits and Rules, MIT Press, Cambridge, MA.

Wolff, R. A., 1981. German past participles and the simplicity metric. Linguistics 19, 3-13.