Experience, aptitude and individual differences in native language ultimate attainment

Dabrowska, Ewa

DOI:
10.1016/j.cognition.2018.05.018

Experience, aptitude and individual differences in native language ultimate attainment

Ewa Dąbrowska, University of Birmingham

Several recent studies have demonstrated that some native speakers do not fully master some fairly basic grammatical constructions of their language, thus challenging the widely-held assumption that all native speakers converge on the same grammar. This study investigates the extent of individual differences in adult native speakers' knowledge of a range of constructions as well as vocabulary size and collocational knowledge, and explores the relationship between these three aspects of linguistic knowledge and four nonlinguistic predictors: nonverbal IQ, language aptitude, print exposure and education. Individual differences in grammatical attainment were comparable to those observed for vocabulary and collocations; furthermore, performance on tests assessing speakers' knowledge of these three aspects of language was correlated (rs from 0.38 to 0.57). Two of the nonlinguistic measures, print exposure and education, were found to contribute to variance in all three language tests, albeit to different extents. In addition, nonverbal IQ was found to be relevant for grammar and vocabulary, and language aptitude for grammar. These findings are broadly compatible with usage-based models of language and problematic for modular theories.

Keywords: individual differences, grammar, vocabulary, collocations, language aptitude, print exposure, education, IQ

1. INTRODUCTION

It is a widely held assumption in linguistics – almost an article of faith – that all native speakers converge on (more or less) the same grammar (see, for example, Bley-Vroman, 2009: 179; Chomsky, 1965, p. 11, 1975, p. 11; Crain & Lillo-Martin, 1999, p. 9; Crain, Thornton & Murasugi, 2009, p. 124; Herschensohn, 2009, p. 264; Lidz & Williams, 2009, p. 177; Montrul, 2008, p. 4; Nowak, Komarova & Niyogi, 2001, p. 114; Seidenberg,1997, p.1600; Smith, 1999, p. 41). Along with poverty of the stimulus, this presumed fact is one of the most powerful arguments for an innate language faculty (cf. Chomsky, 1975, p. 11, 1999, p. 47; Crain, Thornton & Murasugi, 2009, p. 124; Lidz & Williams, 2009, p. 177). There is growing evidence, however, that this assumption is unfounded: a number of studies have demonstrated considerable individual differences in adult native speakers' knowledge of various aspects of the grammar of their language, including complex syntactic structures involving subordination (Chipere, 2001, 2003; Dąbrowska, 1997, 2013; Street 2010, 2017), some simpler constructions such as passives and quantified noun phrases (Dąbrowska & Street, 2006; Street, 2010; Street & Dąbrowska, 2010, 2014), and some aspects of inflectional morphology (Dąbrowska, 2008; Indefrey & Goebel, 1993); for recent reviews, see Dąbrowska (2012, 2015), Hulstijn (2015), and Kidd, Donnelly & Christiansen (2017). The existence of individual differences in native speakers' grammatical competence has important implications for the language sciences in that it undermines the convergence argument for UG and provides indirect support for usage-based (UB) theories of language, which are fully compatible with such differences. Furthermore, the study of IDs in language acquisition and ultimate attainment offers opportunities for testing and refining UB models of acquisition, in that it allows us to examine the relationship between properties of the input and learner characteristics on the one hand, and linguistic outcomes on the other.

Many, although not all, of the individual differences in ultimate attainment observed in earlier studies are education-related. All of the studies that demonstrated the existence of education-related differences compared two groups of participants: a control group of high academic attainment (HAA) participants – often postgraduate students – and a low academic attainment group (LAA) – typically people who left school as soon as it was legally possible and who work in relatively low-skill jobs (cleaners, factory workers, supermarket shelf-stackers, etc.). Most of these studies found a characteristic pattern: while the HAA participants performed at or close to ceiling, the LAA group showed considerable variation, with some participants at ceiling, some at chance (or even below chance), and the majority somewhere in between. The studies employed a number of control conditions to ensure that the observed differences could not be explained by appealing to linguistically irrelevant factors such as attention, cooperativeness, or the ability to perform the experimental task: thus, the observed differences can be regarded as differences in competence rather than performance (for a discussion, see Dąbrowska, 2012a, 2012b). Furthermore, since the studies targeted constructions that do not differ across dialects, the low-performing participants' difficulties cannot be attributed to dialectal differences: for instance, there are no dialects in English in which *The boy was kissed by the girl* means 'The boy kissed the girl'.

The research described in this paper builds on this earlier work but goes beyond it in several ways. First, it strives to give a more rounded picture of adult native speakers' knowledge by targeting a much wider range of constructions than the earlier studies, and also measuring speakers' knowledge of vocabulary and collocations. Furthermore, rather than comparing two groups which are as different from each other as possible, it uses a sample which is representative of the UK population – participants with a range of educational and occupational backgrounds. This will make it possible to assess the extent of the incomplete acquisition identified in the earlier research.

The second goal of the study is to establish whether there is a relationship between different aspects of linguistic knowledge. It is well known that there are robust relationships between vocabulary size and grammar in language acquisition (Bates, Bretherton & Snyder, 1988; Dionne, Dale, Boivin & Plomin, 2003; Huttenlocher, 1998; Stolt, Haataja, Lapinleimu & Lehtonen, 2009; Szagun, Steinbrink, Franik & Stumper, 2006). It is unclear, however, whether this is merely a developmental phenomenon, or whether such correlations persist into adulthood. Cognitive and educational psychologists often refer to “verbal ability”, noting that performance on various verbal tasks tends to be more strongly correlated than performance on verbal and nonverbal tasks; however, such studies typically use broad measures of language proficiency such as reading comprehension, essay writing (and sometimes listening comprehension and speech production) or tasks involving verbal analogies or anagrams, which tap intelligence rather than linguistic knowledge *per se;* apart from vocabulary, they typically do not measure specific aspects of linguistic knowledge, such as the ability to produce or comprehend a particular grammatical construction.

Thus, the relationship between different aspects of linguistic knowledge in adults remains something of a mystery. This is unfortunate, as such knowledge could make an important contribution to our understanding of how linguistic knowledge is organized, and, since different theories make different predictions about these relationships, help us distinguish between alternative accounts of the same phenomena. According to usage-based models (Barlow & Kemmer, 2000; Bybee, 2006, 2010, 2013; Goldberg, 2003, 2006; Langacker, 1988, 2000), linguistic knowledge is represented by a network of constructions, i.e. form-meaning pairings varying in size (from single morphemes to sentence level constructions and beyond) and degree of specificity (from fully specified phonologically through partially abstract to fully abstract). Crucially, grammatical, lexical, and collocational knowledge are all represented in the same format and involve the same mental mechanisms (though possibly to varying extents). Hence, usage-based models assume a close relationship between these types of knowledge, and thus predict that measures of speakers' vocabulary size, grammatical ability, and knowledge of collocations should be correlated.

Modular models (e.g. Chomsky, 1981; Pinker, 1997, 1999; Ullman, 2006), in contrast, assume that lexical and grammatical knowledge rely on different types of representations and belong to different “components”; hence, modular theories do not predict a correlation between these two aspects of speakers' knowledge, although they do not explicitly rule it out. On the other hand, those modular theories which also assume that the development of grammar is underpinned by an innate domain-specific mechanism do predict the absence of a correlation. This is because, if we assume that all speakers are equipped with the same language acquisition device and that the acquisition of basic grammatical constructions is not sensitive to the learners' non-linguistic abilities or to properties of the input (as long as a minimal amount of input is available), variation in adult native speakers' knowledge of core grammar should be minimal, if it exists at all.

Linguistic theories also differ in the relationships they assume between speakers' knowledge of collocations and other aspects of linguistic knowledge. Since collocations are clearly learned from the input, nativist theories predict no relationship between collocational knowledge and core grammar, while allowing for a relationship between collocational knowledge and vocabulary, to the extent that the latter depends on the former, which is very plausible (see Dąbrowska, 2009 for evidence that this is the case). On the other hand, since grammar and collocations both arguably involve procedural memory, Ullman's Declarative-Procedural model (Ullman, 2001, 2004, 2006) would predict a relationship between these two aspects of linguistic knowledge and no link between these abilities and knowledge of word meanings (which relies on declarative memory).

The third goal of the present study is to explore some possible causes of individual differences in linguistic knowledge in adults. According to usage-based models, these should be related, on the one hand, to the quality and quantity of the input available to learners, and on the other, to individual differences in learning mechanisms supporting language. Thus, examining relationships between these factors and linguistic outcomes allows us to test predictions of usage-based models. The present study investigates the effects of two learner-internal factors, namely nonverbal IQ and language aptitude, and two factors related to the quality of the input: education and print exposure.

UB researchers often assert that language depends on "general cognition" but are rather vague when it comes to identifying the specific cognitive abilities that are supposed to be involved. One obvious candidate would be general intelligence, which is all the more interesting in that the development of language, and, in particular, grammatical knowledge, has often been claimed to be unrelated to intelligence (see, for example, Chomsky, 1965, pp. 57-58; Pinker, 1999; Reber, 2011, p. 30; Smith & Tsimpli 1995). Although robust correlations have been observed between grammatical abilities and IQ in development, these could be mediated by vocabulary: if higher IQ is associated with better word learning skills, but grammatical development cannot get off the ground until the learner has acquired a sizeable vocabulary, we can expect to find a significant relationship between IQ and grammar even if grammatical development does not depend on intelligence. However, such an explanation would be difficult to maintain if it turned out that the correlation between grammar and IQ is also found in adult speakers, who may be presumed to have long acquired the critical mass of vocabulary items necessary to trigger grammatical development.

The second learner-internal variable considered in this study is language analytic ability, that is to say, the capacity to (consciously) infer linguistic rules and generalizations. Language analytic ability is known to have an effect on second language learning, especially in classroom settings, and thus most tests of (foreign) language aptitude such as the Modern Language Aptitude Test (or MLAT: see Carrol & Sapon, 1959), the Pimsleur Language Aptitude Battery (PLAB; see Pimsleur, Reed & Stansfield, 2004) or the LLAMA Language Aptitude Test (Meara, 2005) include a component designed to test it. On the other hand, it is generally assumed to be irrelevant for first language development, which is supposed to rely almost entirely on implicit learning mechanisms (Bley-Vroman, 1990; DeKeyser, 2000; DeKeyser, Alfi-Shabtay & Ravid, 2010; Ellis, 2007; Reber, 2011).

There are some reasons to question this assumption. One is theoretical. Tests of language analytic ability such as the Language Analysis subtest of the PLAB, LLAMA F or the York Aptitude Test (Green, 1975) present participants with some vocabulary and sentences in an artificial language and ask them to use that information to infer the form of a novel sentence in the language (see section 2.2.6 below for an example). In order to succeed on such a task, an individual must be able to establish correspondences between chunks of form and chunks of semantic representation and use analogy to extend the pattern exemplified in the example sentences to novel utterances. These are operations which, according to usage-based accounts, underlie children's ability to acquire language (see Dąbrowska, 2004, 2009, 2014a; Tomasello, 2003). Thus, although such tests tap explicit analytical abilities which are usually assumed to be very different from the abilities that children bring to the language learning task, perhaps the difference is not as large as many researchers believe.

The second reason to question the assumption that the ability to reason explicitly about language is irrelevant to first language acquisition is empirical. Skehan and Ducroquet (1988) administered a battery of (foreign) language aptitude and foreign language achievement tests to 103 fourteen-year-olds who had participated in the Bristol Language Project, a longitudinal study of early language development (Wells, 1985), and compared these data with the measures taken during the original study. Table 1 shows the correlations observed for two of the aptitude measures used in the study, the Matching Words subtest of the EMLAT and the York Aptitude Test (Green, 1975)[1], and selected measures of L1 development.

1 The EMLAT (Modern Language Aptitude Test: Elementary, Carroll and Sapon 1967) is a version of MLAT developed for use with children. In the Matching Words subtest, participants are presented with a model sentence containing an underlined word and asked to identify a word in another sentence which fulfils the same grammatical function. The York Language Aptitude Test (Green 1975) measures inductive language learning ability by asking participants to produce foreign language forms on analogy to forms presented; it is thus similar to the Language Analysis subtest of the PLAB which will be discussed later.

Table 1

Correlations between measures of L1 development and language aptitude (based on Skehan & Ducroquet, 1988, older cohort)

Developmental measure	Matching Words	York
MLU at 42 months	0.52***	0.35**
Comprehension at 39 months	0.45***	0.35**
Vocabulary at 39 months	0.48***	0.52***
Comprehension at 57 months	0.51***	0.43**
Vocabulary at 66 months	0.21	0.37**
Range of synt. complexity	0.28*	0.04
Range of NP complexity	0.43**	0.26*

*** Correlation is significant at 0.001 level (two-tailed).
** Correlation is significant at 0.01 level (two-tailed).
* Correlation is significant at 0.05 level (two-tailed).

The Bristol study involved two cohorts of children. The data in the table are for the older cohort, who were followed from the age of 39 months to 60 months. All correlations with coefficients of 0.24 or higher are statistically significant. As we can see, the data reveal a number of robust correlations between first language development in early childhood and measures of language aptitude. The younger cohort, who were followed from 15 to 42 months, also showed a number of significant correlations between the two sets of measures although they tended to be lower than those of the older cohort. Given the size of these correlations, and the fact that the aptitude tests were administered more than 10 years later, these results are quite striking, and suggest that the relationship between language aptitude and first language achievement is worth exploring further.[2] Accordingly, a measure of language analytic ability was incorporated into this study.

In addition, this study investigates the relationship between linguistic abilities and two input-related factors, namely print exposure and education. With regard to the former, it is well established that written texts are lexically richer and contain more difficult constructions that spoken texts (Biber, 1995; Cameron-Faulkner & Noble, 2013; Cunningham & Stanovich, 1998; Roland, Dick & Elman, 2007). Therefore, to the extent that mastery of a linguistic unit, be it lexical or grammatical, depends on the amount of experience with it (cf. Diessel, 2007; Divjak & Cardwell-Harris, 2015; Ellis, Römer & O'Donnell, 2016), we would expect speakers with more print exposure to perform better on tasks tapping mastery of less frequent words and constructions. Exposure to a particular linguistic unit in writing may be more conducive to learning in other ways as well. In ordinary face-to-face communication, listeners can rely on a variety of sources of information in order to infer the speakers' meaning, including lexical and grammatical cues, prosody, gesture and the situational context; and, when all else fails, they can ask for clarification. In most written communication, in contrast, the information to be conveyed must be expressed through lexical or grammatical means (although punctuation, capitalization and font size may provide some additional cues). Thus, processing a written text requires more focus on the linguistic form itself, which may benefit learning. There is another sense, however, in which processing written text is less demanding: a reader – in contrast to a listener – can process the text at their own pace, and can go back and reread passages which they find difficult. This means that once they have mastered the basic mechanics of reading, language learners can use the written form as a "processing crutch" which enables them to comprehend and produce – and hence practice – more complex sentences.

2 Correlations between language aptitude and foreign language achievement were even higher – from 0.66 to 0.69 for Matching Words, and from 0.65 to 0.74 for the York test in the older cohort; the figures for the younger cohort were also statistically significant, but somewhat lower.

It is well known that reading leads to vocabulary growth. There is very little research about the effects of reading on knowledge of specific constructions. However, two earlier studies which examined the issue are encouraging. Street and Dąbrowska (2010) found correlations between print exposure in adults matched for education and comprehension of passives and quantifiers. More conclusively, Wells, Christiansen, Race, Acheson and MacDonald (2009) manipulated the reading experience of adults over several weeks and found that the experimental group, which was exposed to a large number of relative clauses, increased reading speeds for relatives, in particular object relatives, while the control group did not.

The last possible determinant of individual differences that will be examined here is education. As mentioned above, several earlier studies have demonstrated a strong relationship between educational attainment and grammatical abilities. However, it is unclear whether this correlation should be attributed to the specific properties of language encountered in educational settings – decontextualised expository prose, which has many characteristics of written language (Ravid, 2004; Nippold, Hesketh, Duthie & Mansfield, 2005; Verhoeven et al., 2002) and poses similar demands on the learner, or whether the relationship is attributable to other factors that correlate with educational attainment such as exposure to written language, IQ and language aptitude. Examining all these factors in the same study will enable to unpick the relationships between them.

2. METHOD

2.1 Participants

90 adult native speakers of English (42 males and 48 females) aged from 17 to 65 (mean 38, median 32) were recruited through advertisements in local press, church groups, schools and personal contacts. The distribution of qualifications roughly reflected the demographics of the UK

population, according to data from the 2011 Skills for Life Survey (Department of Business Innovation and Skills, 2011), although the proportion of participants with postgraduate degrees was slightly higher and the proportion of participants with no formal qualifications slightly lower. Three participants had no formal qualifications, 60 held a secondary school or vocational certificate, 16 held an undergraduate degree or were studying for one, and 11 held postgraduate degrees. Years spent in full-time education ranged from 10 to 21 (mean 13.7, median 13). Fifteen of the participants were in full-time education, 6 were unemployed, 4 were housewives and 14 were retired. Of the remaining 41, approximately one-third held manual jobs such as cleaner, shop assistant, waitress, or roofer; another third held clerical jobs (office worker, IT support) and the last third held graduate-level jobs such as web designer, teacher, or quality surveyor. All were born and raised in English-speaking families in the UK. 84 described themselves as monolingual; 6 reported a reasonable working knowledge of another language (French, Spanish or German).

2.2 Materials

2.2.1 Grammatical comprehension

Grammatical comprehension was measured using the “Pictures and Sentences” test developed especially for this project (see Supplementary Materials, Appendix A). The test targets ten constructions (see Table 2), with eight items for each construction. Four of the constructions (actives, simple locatives, subject relatives, object relatives) are very basic structures which are acquired early by children and which were not expected to pose difficulties for any participants. They were included in the test to ensure that there were some items that all participants could respond confidently to. They also function as control conditions: if participants perform at ceiling on these sentences, then difficulties with other sentence types cannot be explained away simply as failure to understand the instructions or unwillingness to engage with the task. The other constructions are also relatively basic in that they involve structures that one would expect to find in ordinary conversation; however, there were expected to be more challenging. Object relatives, object clefts and passives all involve noncanonical word order, are acquired late and are difficult even for adult speakers (Ferreira, 2003; Hachmann, Müller & Konieczny, 2009; Street & Dąbrowska, 2010, Wells et al. 2009). The contrast between the two sentence types containing the universal quantifier *every* (Q-is and Q-has) requires establishing the scope of the quantifier, which has also been found difficult for some adult native speakers (Brooks & Sekerina, 2006; Street & Dąbrowska, 2010). Finally, processing sentences with complex subjects involves attending to the structural relations between the head of the phrase and the noun in the postmodifying prepositional phrase rather than purely linear relations, i.e. assuming that the noun immediately preceding the verb is the subject. Such structures are characteristic of written language and relatively rare in speech, and hence could be difficult for speakers with low print exposure.

Each item in the test consisted of two pictures and a sentence; the participants' task was to choose the picture that goes with the sentence. There were four versions of the test which differed in the ordering of items and the assignment of sentences to pictures, so that for sentences describing transitive events, each person mentioned in the sentence occurred equally often in the agent and the patient role. Participants were randomly assigned to one of the four versions, approximately a quarter to each.

Table 2

Structures assessed by “Pictures and Sentences”

Construction	Example	Pictures
Active	The man fed the girl.	Target: man feeding girl Distractor: girl feeding man
Passive	The girl was fed by the man.	
Subject Cleft	It was the man who fed the girl.	
Object Cleft	It was the girl that the man fed.	
Subject Relative	The man was the one who fed the girl.	
Object Relative	The girl was the one that the man fed.	
Locative w/ quantifier (Q-is)	Every pencil is in a box.	Target: three boxes, each containing a pencil plus an extra box Distractor: three boxes, each containing a pencil plus an extra pencil
Possessive w/ quantifier (Q-has)	Every box has a pencil in it.	Target: three boxes, each containing a pencil plus an extra pencil Distractor: three boxes, each containing a pencil plus an extra box
Simple locative	The spoon is in the cup.	Target: spoon in cup Distractor: spoon next to cup
Postmodifying PP	The spoon in the cup is red.	Target: red spoon in green cup Distractor: yellow spoon in red cup

2.2.2 Receptive Vocabulary

Receptive vocabulary was tested using a shortened version of the Vocabulary Size Test developed by Nation and Beglar (2007; see also Beglar, 2009). The original test systematically samples 14 frequency levels established using data from the spoken part of the BNC. Frequency level 1 comprises the 1000 most frequent “word families”, where a family is a root plus all words derived from it by regular word formation processes and their inflectional forms; frequency level 2 comprises the second 1000 word families, and so on. Each test word is presented in a short non-defining context and followed by 4 simple definitions (see examples (1) and (2) below); the participant's task is to choose the best definition for the target word. The full test contains 10 items for each vocabulary level, and hence 140 items in total. Since participants in this study had to complete five other tasks, the test was abridged to avoid fatigue. Levels 1 and 2, which contain the most frequent words, were omitted, and only the odd-numbered items were selected from the remaining levels. Thus, the abridged version contained 60 items.

(1) ___ fracture: They found a **fracture**.

a. break

b. small piece

c. short coat

d. rare jewel

(2) ____ marsupial: It is a **marsupial**.

a. an animal with hard feet

b. a plant that grows for several years

c. a plant with flowers that turn to face the sun

d. an animal with a pocket for babies

2.2.3 Collocations

Knowledge of collocations was assessed using the Words that Go Together test (Dąbrowska, 2014b). The test contains 40 multiple choice items, each of which consists of five short phrases (examples 3-4; the target answers are *blatant lie* and *raise prices*); the participants' task is to select one phrase from each set which "sounds the most natural or familiar". The target phrases are adjective-noun and verb-noun collocations which vary in frequency (from 0.06 to 6.19 per million) and in collocation strength (mutual information from 4.4 to 15.6). The distractors are grammatically correct phrases containing the same noun and adjectives or verbs which could plausibly combine with the noun but with much lower mutual information scores; thus, in order to identify the target participants must know which combinations of words tend to occur together. Dąbrowska (2014b) reports that the test shows robust correlations with various measures of language exposure but not with IQ, which suggest that the test is a valid measure of collocations knowledge.

(3) blatant lie
clear lie
conspicuous lie
distinct lie
recognizable lie

(4) elevate prices
grow prices
lift prices
raise prices
stimulate prices

2.2.4 Nonverbal IQ

Nonverbal IQ was measured using the Shipley-2 Block Patterns test (Shipley, Gruber, Martin & Klein, 2009). Block Patterns is a pen-and-paper adaptation of Kohs' Block Design Test. Participants are presented with a pattern made up of black and white squares and an incomplete version of the same pattern which is larger and sometimes rotated, and are asked to choose the squares that would complete it. The test consists of 12 multiple-choice problems arranged in order of difficulty. The difficulty gradient is quite steep, so the test can be regarded as measuring not only the construct that it was designed to measure, i.e. fluid intelligence, but also, indirectly, persistence and cooperativeness.

2.2.5 Print exposure

Print exposure was measured using the Author Recognition Test (ART) developed by Acheson, Wells and MacDonald (2008). The ART consists of 130 names, of which 65 are the names of real authors and the remaining 65 foils. Participants are asked to mark the names that they believe to be those of real authors, and instructed not to guess. The score is the total number of correctly identified authors minus the number of foils marked. Thus the highest possible score is 65. The ART measures the amount of print exposure over a life-time (rather than just current reading habits); moreover, it is thought be more accurate than measures based on self-report, which tend to be contaminated by the tendency to give socially desirable answers (Stanovich & West, 1989). Numerous studies have established that the ART correlates highly with measures of orthographic processing, spelling, vocabulary size, reading comprehension and general knowledge, demonstrating that it is a valid measure of print exposure (Cunningham & Stanovich, 1998; Stanovich & Cunningham, 1992; Stanovich & West, 1989).

2.2.6 Language analytic ability

Language analytic ability was measured using the Language Analysis subtest of the Pimsleur Language Aptitude Battery (Pimsleur et al., 2004). The PLAB was developed for use in schools (grades 7-12, i.e. with 13- to 19-year-olds) to predict foreign language success. The first two parts elicit information about the participants' grade point average and interest in foreign language learning. The remaining four parts measure English receptive vocabulary, language analytic ability, the ability to discriminate non-native sounds, and the ability to learn sound-symbol associations. In the Language Analysis subtest, participants are presented with some vocabulary and sample sentences in an unknown language, e.g.

jiban '(a) boy'

jojo '(a) dog'

jiban njojo za 'A boy likes a dog.'

and asked to predict a novel form (e.g. how would you say 'A dog likes a boy?'). They have to select the correct answer (*jojo njiban za*) from an array of four.[3] The test thus measures the ability to map form onto meaning, reasoning by analogy and, to some extent, working memory capacity.

2.3 Procedure

Participants were tested individually in a quiet room. After the initial briefing and giving consent, they completed a background profile which elicited information about their age, sex, education, occupation, knowledge of languages other than English, and reading habits. This was followed by the three language tasks (Pictures and Sentences, Vocabulary Size Test, Words that Go Together) administered orally. The experimenter read the printed questions while participants

3 This example is a practice item available from the web (see (http://lltf.net/aptitude-tests/language-aptitude-tests/pimsleur-language-aptitude-battery/plab-sample-items/plab-part-4/) rather than an item from the actual test.

followed in their own version of the test and responded orally. Participants were given as much time as they needed, but most completed each test in 10-12 minutes. Finally, participants completed the three cognitive tests (Block Patterns, Language Analysis, and Author Recognition). These were administered in writing following the standard procedure as outlined in the test manuals. There were short breaks between the tests. The entire testing session lasted from 1.5 to 2 hours.

3. RESULTS AND DISCUSSION

3.1 Performance on the grammar test

Table 3 shows the means, interquartile ranges and ranges for each of the 10 constructions tested; the entire dataset is available in Appendix B in the Supplementary Materials. As expected, performance on the four control conditions (Active, Locative, Subject Relative, and Subject Cleft) was at ceiling, with mean performance at 7.9 out of 8, i.e. 99% correct, for all four. This shows that participants had understood the task and were cooperative; thus, problems with the experimental sentences cannot be attributed to linguistically irrelevant performance factors.

Table 3

Performance by construction on the grammatical comprehension test

Construction	Mean	IQR	Range
Active	7.9	8-8	5-8
Locative	7.9	8-8	6-8
SubRel	7.9	8-8	6-8
SubCl	7.9	8-8	5-8
Passive	7.8	8-8	2-8
PostMod	7.6	7-8	4-8
ObjCl	7.4	7-8	4-8
ObjRel	7.4	7-8	1-8
Q-is	6.5	6-8	0-8
Q-has	6.2	5-8	0-8

Performance on passive sentences was also very good (mean 7.8); only one participant performed at chance, scoring 2 out of 8. This is much better than the performance on the passive observed by Street and Dąbrowska (2010), where about one-third of the low academic attainment group performed at chance. However, only three of the participants in the current study had educational backgrounds comparable to that of the LAA group in the former study (i.e., no formal qualifications), so the proportion of participants who had not mastered the passive was about the same in both studies. Thus, the present findings are not incompatible with the earlier research; they just show that problems with passives are quite rare in a sample that is more representative of the general population.

The remaining five constructions were harder. While the mean score for sentences with subjects containing a postmodifying PP is quite high (7.6), 9 out of 90 participants, i.e. 10%, performed at chance, with scores ranging from 4 to 6. The mean scores for the next two constructions, object relatives and object clefts were identical (7.4). For object relatives, 13 participants performed at chance (3-6 correct) and one participant below chance (1 correct). For object clefts, 13 participants were at chance (4-6 correct); of these, 7 also had problems with object relatives. Thus, performance on these two constructions was very similar. This is not surprising, given the structural similarities between them.

The two quantifier constructions were the most difficult, possibly because quantifier scope is conceptually more complex than identifying semantic roles. The mean score on Q-is was 6.5, with 29 participants performing at chance and another 8 below chance. For Q-has, the mean was even lower (6.2), and 32 participants performed at chance and 4 below chance. Of these 36, 13 were also at chance on Q-is. Surprisingly, the correlation between scores on these two sentence types was close to zero ($r = -0.11$, $p = 0.32$). This is probably due to some participants adopting a particular response strategy, i.e. having a strong preference for the picture with the extra object or the extra container in both conditions. As a result, five participants were at ceiling (8 out of 8) in the Q-has condition and below chance (0 or 1 out of 8) on Q-is, and another three showed the opposite pattern.

Thus, while average performance on the grammar task was good for all constructions, a sizeable minority of participants performed at chance, and in some cases even below chance, on some constructions. This shows that, even for the relatively simple structures examined here, incomplete acquisition is far from being a marginal phenomenon.

Table 4 shows correlations in scores on individual constructions. It is quite striking that most of the correlations are relatively weak, with 30 out of 45, i.e. 67%, having coefficients of ±.2 or below. Doubtless this is partly due to the fact that there were only 8 items per construction, and relatively little variation in scores, and hence the measurements of individual scores are not very precise. However, correlations between the three constructions with canonical SVO order (subject cleft, subject relative, and active) are actually relatively high (between .44 and .53), and these were the very constructions where performance was closest to ceiling. Likewise, there are moderately strong correlations between two pairs of constructions with non-canonical word order, namely between object relatives and object clefts and between object relatives and passives (0.49 and 0.38 respectively), but, interestingly, not between object clefts and passives (0.08). Thus, while caution must be used when interpreting these results, the overall low correlations between scores in different conditions suggests that knowledge of different constructions is relatively independent.

Table 4

Correlations between scores on individual constructions

	Active	Locative	ObjCl	ObjRel	Passive	Postmod	Q.has	Q.is	SubCl	SubRel
Active	1.00	0.00	-0.03	0.17	0.01	0.17	0.16	0.05	0.53***	0.44***
Locative	0.00	1.00	0.12	0.23	-0.05	-0.02	0.02	-0.13	0.04	0.06
ObjCl	-0.03	0.12	1.00	0.49***	0.08	0.39**	0.35*	0.11	-0.06	-0.08
ObjRel	0.17	0.23	0.49***	1.00	0.38**	0.32†	0.22	0.11	0.22	0.11
Passive	0.01	-0.05	0.08	0.38**	1.00	0.24	0.13	0.01	0.12	0.20
Postmod	0.17	-0.02	0.39**	0.32†	0.24	1.00	0.09	0.41**	0.25	0.10
Q.has	0.16	0.02	0.35*	0.22	0.13	0.09	1.00	-0.11	0.17	0.18
Q.is	0.05	-0.13	0.11	0.11	0.01	0.41**	-0.11	1.00	0.10	-0.02
SubCl	0.53***	0.04	-0.06	0.22	0.12	0.25	0.17	0.10	1.00	0.47***
SubRel	0.44***	0.06	-0.08	0.11	0.20	0.10	0.18	-0.02	0.47***	1.00

*** Correlation is significant at 0.001 level (two-tailed).
** Correlation is significant at 0.01 level (two-tailed).
* Correlation is significant at 0.05 level (two-tailed).
† Correlation approaches significance ($p<0.10$).
The reported *p* values have been adjusted using Holm's method.

3.2 Test-retest reliability

To establish test-retest reliability, 25 of the original participants were invited to come to a second testing session administered 3 to 6 months after the main experiment. In this session, participants completed only the three language tasks. The correlations between scores on the two tests were 0.81 for grammar, 0.94 for vocabulary and 0.83 for collocations, with $p < 0.001$ for all three.

3.3 Overall results

In the remainder of this paper, I will consider aggregate scores for grammar obtained by summing the scores for the six experimental structures (i.e. the Passive, Postmod, ObjCl, ObjRel, Q-is and Q-has conditions). As explained earlier, although all three of the language tests used a multiple-choice format, they differed in the number of choices: two for the grammar test, four for vocabulary, and five for collocations. To allow for meaningful comparisons, the raw scores were corrected for guessing and transformed into percentages using the formula in (5). One participant had marked more non-authors than authors on the ART and thus obtained a negative score; this was changed to zero. There were two missing values, one for ART and one for Blocks; these were imputed using the `mice` package in R.

(5) corrected score = 100*(score – chance)/(perfect score – chance)

The descriptive statistics for performance on the six experimental tasks are presented in Table 5. As we can see, there was considerable individual variation on all three linguistic tasks, with overall performance being highest for grammar and lowest for collocations.

Table 5

Descriptive statistics for the six experimental tasks

	Mean	SD	Median	Range	IQR
Grammar	79	19	83	17-100	67-96
Vocabulary	69	20	74	9-96	56-86
Collocations	66	20	69	9-97	54-84
PrintExp	13	11	11	0-46	6-17
IQ	99	14	101	68-126	87-110
LgAnalysis	8	4	7	0-15	5-10

Note: The scores for Grammar, Vocabulary and Collocations are percentages corrected for guessing. The figures for given IQ (i.e., Block Patterns) are standard scores. The scores for print exposure (ART) and Language Analysis are raw scores; the maximum possible scores for these two tests are 65 and 15, respectively.

Table 6 shows correlation coefficients between the six experimental tasks and a seventh variable, education, operationalized as the number of years spent in full-time education. As we can see, there is a strong correlation between vocabulary size and collocational knowledge ($r=0.57$, $p<0.001$), and weaker, but nevertheless significant relationships between grammar and vocabulary ($r=0.40$, $p = 0.001$) and grammar and collocations ($r=0.38$, $p = 0.002$). However, two of the language measures are related to nonverbal IQ (as measured by the Block Design test) and all three are related to the number of years spent in formal education and to language analytic ability. This raises the possibility that the observed correlations are driven by general ability, familiarity with tests, or even cooperativeness with the researcher. To determine whether this was the case, partial correlations were computed between pairs of all three linguistic variables while partialling out the effect Block Design – a test which was designed to measure nonverbal IQ, but, because it is quite demanding, can also be assumed to measure test-taking skills and motivation (see Shipley et al., 2009). Partialling out the effect of IQ had very little effect on the results: although the correlation coefficient dropped slightly, all three relationships remained significant (collocations and vocabulary: $r = 0.53$, $p < 0.001$; grammar and vocabulary: $r = 0.25$, $p = 0.018$; collocations and grammar: $r = 0.32$, $p = 0.002$). Thus, as predicted by usage-based models, the three aspects of linguistic knowledge examined here – grammar, vocabulary, and collocations – are associated with each other, even in adults.

Table 6

Correlations between the six tasks and education

	Education	IQ	PrintExp	LgAnalysis	Grammar	Vocabulary	Collocations
Education	1.00	0.36**	0.47***	0.59***	0.36**	0.43***	0.37**
IQ	0.36**	1.00	0.09	0.50***	0.46***	0.43***	0.22
PrintExp	0.47***	0.09	1.00	0.45***	0.27*	0.60***	0.50***
LgAnalysis	0.59***	0.50***	0.45***	1.00	0.46***	0.47***	0.35**
Grammar	0.36**	0.46***	0.27*	0.46***	1.00	0.40***	0.38**
Vocabulary	0.43***	0.43***	0.60***	0.47***	0.40***	1.00	0.57***
Collocations	0.37**	0.22	0.50***	0.35**	0.38**	0.57***	1.00

*** Correlation is significant at 0.001 level (two-tailed).
** Correlation is significant at 0.01 level (two-tailed).
* Correlation is significant at 0.05 level (two-tailed).
† Correlation approaches significance ($p<0.10$).
The reported *p* values have been adjusted using Holm's method.

3.4 Regression analyses

The effect of the four predictors (education, print exposure, nonverbal IQ and language analytic ability) on speakers' knowledge of grammar, vocabulary and collocations was examined using standard multiple regression modelling. The initial models contained the four predictors and all their interactions; these were simplified following the procedure recommended by Crawley (2007), i.e., nonsignificant predictors were removed one at a time, beginning with the highest order interactions. In order to facilitate interpretation of model coefficients, all predictors were centred before modelling. The relative importance of each predictor was assessed by using the `lmg` metric, which was computed using the `relaimpo` package in R (see Grömping, 2006, 2007). This measure, named after Lindeman, Merenda and Gold (1980), is obtained by averaging the sequential sum-of-squares obtained from all possible orderings of predictors. It can be thought of as analogous to a squared semi-partial correlation (sr^2) and allows the researcher to quantify each individual regressor's contribution to the variance explained by the model (see Larson-Hall, 2010); the values of individual regressors add up to model R^2.

3.4.1 Grammar

The results of the regression analysis for grammar are presented in Table 7. Since the model violates the homoscedasticity assumption, the reported standard errors, *t* values and *p* values have been adjusted using the Eicker-Huber-White method. As we can see, the most important predictors for grammar are nonverbal IQ (as measured by Block Design) which accounts for over 12% of the variance in scores, and language analytic ability (as measured by the Language Analysis subtest of the PLAB), which accounts for almost 10% of the variance. Print exposure (measured by the ART) and years spent in full-time education also make a small but significant contribution. There are also significant interactions between education and language analytic ability and between nonverbal IQ and print exposure.

Table 7

Regression results for grammar

Variable	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	lmg
Intercept	81.791	1.761	46.445	<0.001***	
Education	1.541	0.740	2.082	0.040*	0.046
PrintExp	0.365	0.128	2.863	0.005**	0.034
IQ	0.302	0.123	2.463	0.016*	0.122
LgAnalysis	1.127	0.534	2.110	0.038*	0.097
Education × LgAnalysis	-0.520	0.171	-3.032	0.003*	0.053
IQ × PrintExp	-0.025	0.010	-2.514	0.014*	0.025
R^2					0.377

*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$ † $p < 0.10$

The relationship between language analytic ability, education and grammar is depicted in Figure 1. Language analytic ability and number of years in full time education are depicted on the x and y axes respectively, and the size of the dots represents performance on the grammar test. (Please note that in this and the following four graphs performance on the linguistic test is plotted as a continuous variable; thus the dots in the legend are for guidance only.) Two things are immediately obvious on inspecting the figure. First, nearly all highly educated participants have good language analytic abilities, while there is much more variation among the less educated participants. Secondly, low grammar scores are found only in the bottom left quadrant, i.e., among participants who have poor language analytic ability *and* below-average education level. In other words, it seems that good analytic ability can compensate for low education, and, conversely, education can compensate for low language-analytic skills.

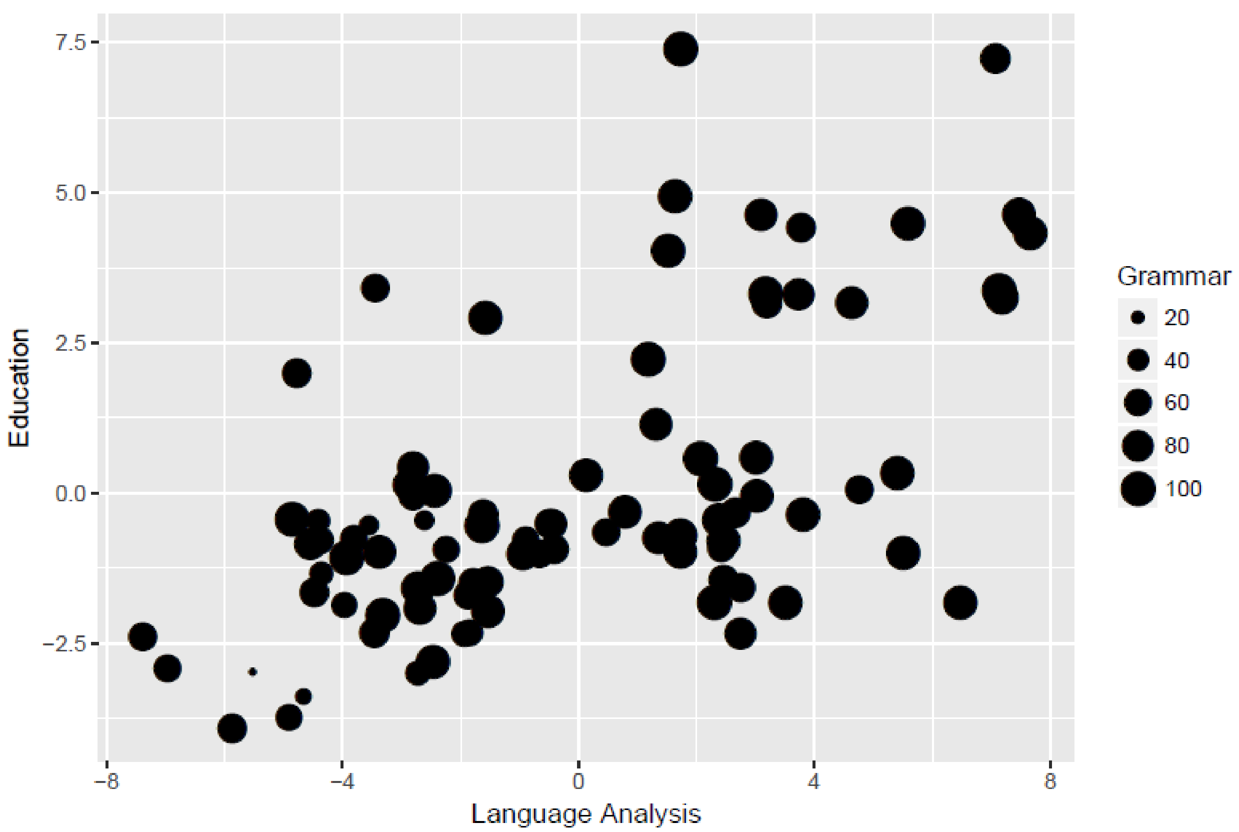

Figure 1. The relationship between language analytic ability, education and grammar scores

As shown in Figure 2, the relationship between nonverbal IQ, print exposure and grammar shows a similar pattern. Again, very low scores are seen primarily in the bottom left quadrant, i.e. when the two disadvantaging factors, low IQ *and* low print exposure, conspire. Thus, it appears that above average IQ can compensate for low exposure to print and vice versa.

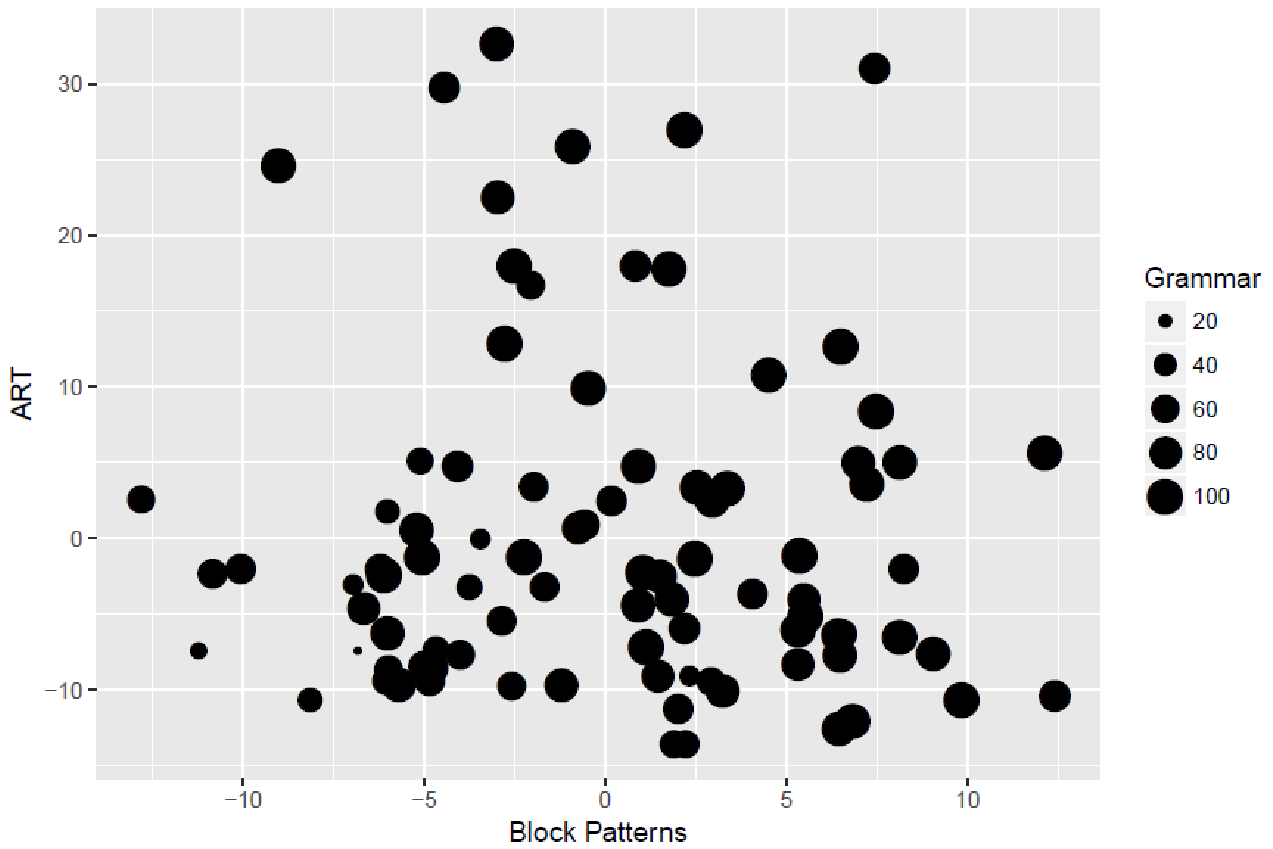

Figure 2. The relationship between nonverbal IQ (Block Patterns), print exposure (ART) and grammar scores

3.4.2 Vocabulary

Table 8 presents the results of the regression analysis for vocabulary. As we can see, print exposure is by far the most important predictor of vocabulary knowledge and accounts for almost 26% of the variance. Nonverbal IQ accounts for 11% of the variance and education for another 8%

(although the effect of education is only marginally significant). However, there is a significant interaction between education and print exposure which accounts for a further 5% of the variance. The interaction is represented graphically in Figure 3 and resembles the interactions we saw earlier when discussing grammar scores: again, very low scores are found only in the bottom left quadrant. In other words, the *combination* of low print exposure and low education is particularly detrimental to vocabulary scores. Language analytic ability is not significant on its own; however, there is a significant interaction between analytic ability and nonverbal IQ. As we can see from Figure 4, IQ doesn't seem to have an effect on performance on the vocabulary test in participants with language analysis scores of 4 or more points above average.

Table 8

Regression results for vocabulary

Variable	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	lmg
Intercept	73.401	1.733	42.338	<0.001***	
Education	1.545	0.866	1.784	0.078†	0.060
PrintExp	1.230	0.165	7.434	<0.001***	0.258
IQ	0.409	0.121	3.377	0.001***	0.113
LgAnalysis	0.017	0.533	0.032	0.975	0.079
Education × PrintExp	-0.160	0.048	-3.344	0.001***	0.051
IQ × LgAnalysis	-0.088	0.029	-3.053	0.003**	0.038
Model R^2					0.600

*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$ † $p < 0.10$.

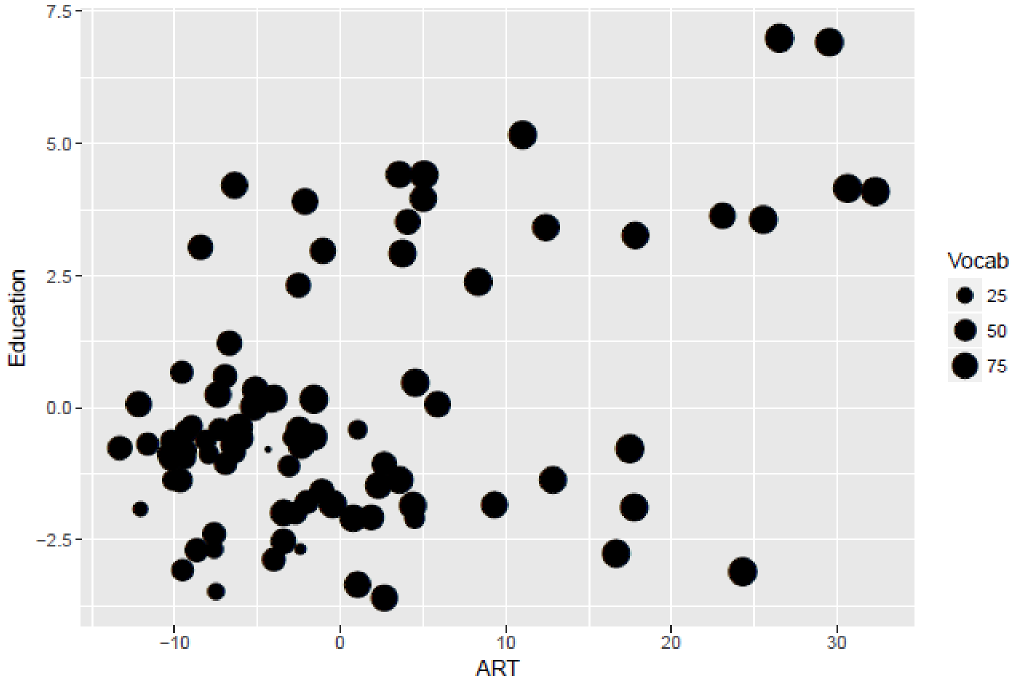

Figure 3. The relationship between print exposure (ART), education and vocabulary scores

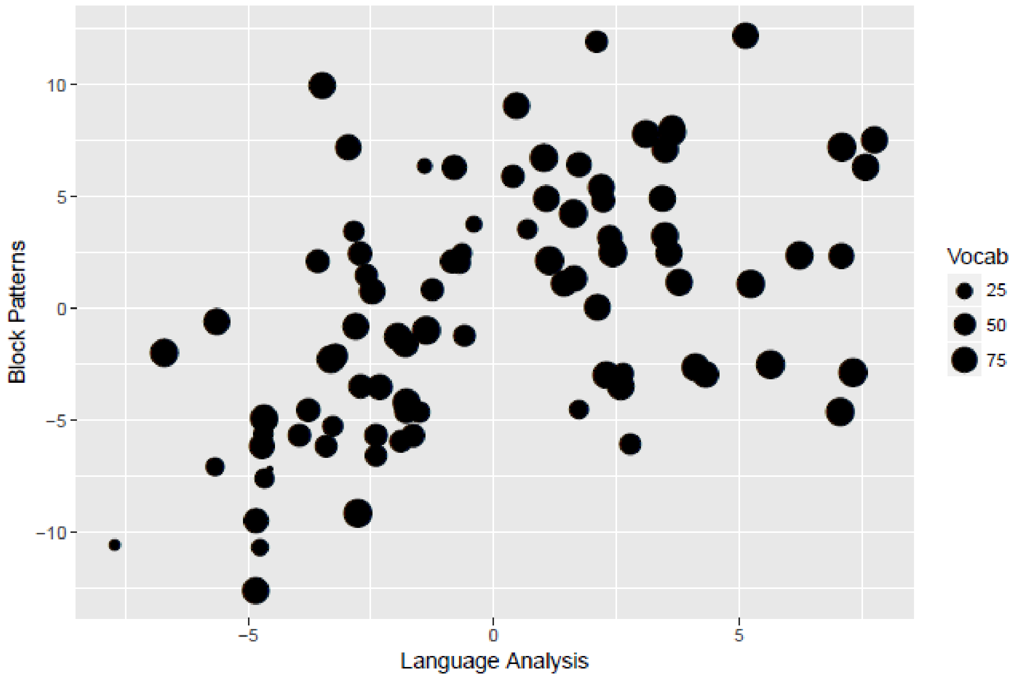

Figure 4. The relationship between language analytic ability, nonverbal IQ (Block Patterns) and vocabulary scores

3.4.3 Collocations

Finally, the results of the regression analysis for collocations are presented in Table 9. Since the model violates the homoscedasticity assumption, the reported standard errors, *t* values and *p* values have been adjusted using the Eicker-Huber-White method. As with vocabulary, the most important predictor of performance is print exposure, which accounts for over 20% of the variance. Education is also relevant, explaining an additional 8%. Last but not least, we have an interaction between these two variables which accounts for another 7% of the variance. The interaction, depicted graphically in Figure 5, shows the by-now familiar pattern: higher educational attainment can to a certain extent compensate for low print exposure and vice versa, while the combination of low education and low print exposure results in particularly low scores.

Table 9

Regression results for collocations

Variable	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	lmg
Intercept	68.518	1.693	40.468	<0.001***	
Education	2.269	0.908	2.951	0.004**	0.081
PrintExp	0.999	0.187	5.336	<0.001***	0.206
Education × PrintExp	-0.177	0.053	-3.331	0.001***	0.074
Model R^2					0.361

*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$ † $p < 0.10$.

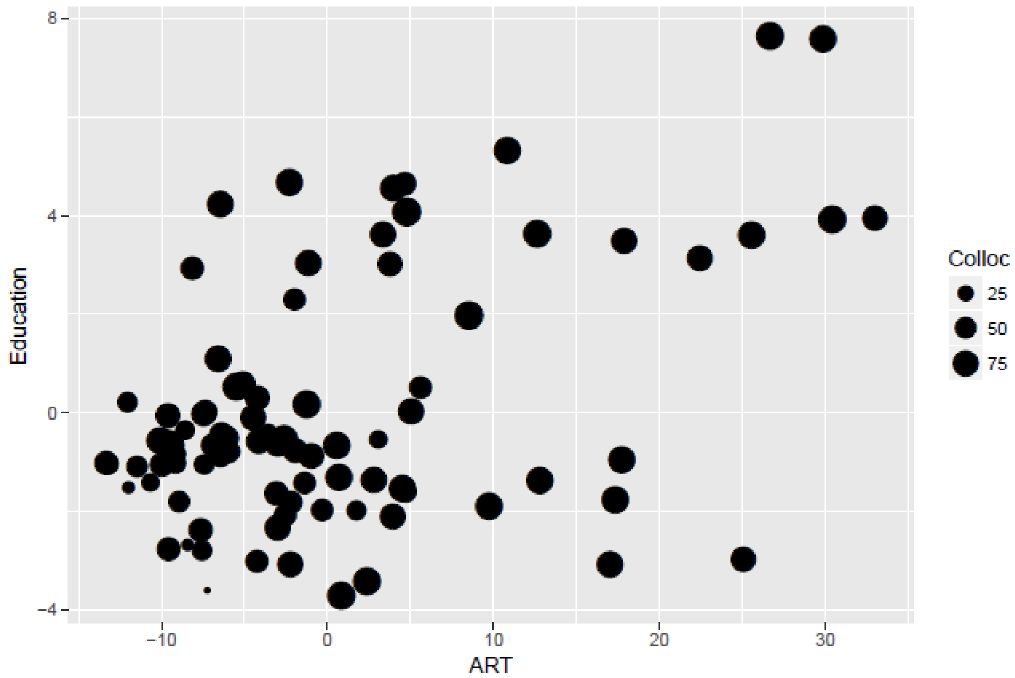

Figure 5. The relationship between print exposure (ART), education and collocation scores

3.4.4 Comparison of the effects of nonlinguistic predictors on performance on grammar, vocabulary and collocations tasks

So far, we have analysed each aspect of linguistic knowledge independently in an attempt to construct the best model for each – that is to say, a model which balances complexity and explanatory power. However, another goal of this paper was to compare the effects of education, print exposure, IQ and language analytic ability on grammar, vocabulary size and collocational knowledge. Since estimates of the contribution that each predictor makes depend on which other predictors are present in the model, in order to directly compare the different areas of linguistic knowledge we need to compare models which use the same explanatory variables. Accordingly, three new models were constructed to predict the three outcomes using the four predictors as well as all the interactions that were found to be significant for at least one of the outcomes, namely Education×PrintExp (relevant for vocabulary and collocations), Education×LgAnalysis and PrintExp×IQ (relevant for grammar), and IQ×LgAnalysis (relevant for vocabulary), and `lmg` values were computed for each model. The results are summarized in Table 10.

Table 10

The relative importance of all predictors for grammar, vocabulary and collocations

	Grammar	Vocabulary	Collocations
Education	0.047*	0.060†	0.054*
PrintExp	0.034*	0.259***	0.185***
IQ	0.121*	0.113**	0.020
LgAnalysis	0.096†	0.079	0.045
Education×PrintExp	0.032	0.046**	0.071**
Education×LgAnalysis	0.039†	0.005	0.003
PrintExp×IQ	0.018	0.004	0.008
IQ×LgAnalysis	0.004	0.035**	0.015*
total R^2	0.391	0.601	0.402

*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$ † $p < 0.10$.

As we can see, education seems to have a similar effect on all three areas of linguistic knowledge, accounting for 5-6% of the variance. We saw earlier that simple correlations between language attainment and education are higher, particularly for vocabulary (r= 0.43, i.e. 18% of the variance). This is due to the fact that education itself is correlated with the other three predictors (cf. Table 5): in fact, we can predict about 42% of the variance in educational attainment from the other three variables.[4] In other words, the observed correlation between education and linguistic attainment is largely attributable to other explanatory variables. Education itself does have a measurable effect, but it is relatively modest.

Print exposure is also relevant for all three language outcomes, but the effect on vocabulary and collocations (26% and 19% respectively) is much larger than for grammar (only 3%). This is probably due to the fact that the grammatical constructions tested are very frequent, and hence all

4 Interestingly, language analytic ability is the best predictor of educational attainment, accounting for about 21% of the variance. Print exposure is also important (14%), while nonverbal IQ makes a relatively modest contribution (7%).

participants had encountered numerous relevant exemplars regardless of how much they read, while the items on the vocabulary and collocations test sampled a wide range of frequencies.

The next predictor, nonverbal IQ, is highly relevant for grammar and vocabulary, but is not related to knowledge about collocations. The lack of effect for collocations is not surprising: one does not need highly developed reasoning skills in order to learn which words tend to co-occur in texts. The relationship between nonverbal IQ and vocabulary is well established (Jensen 1998, Raven 2000). A more surprising, and theoretically more interesting, finding is that nonverbal IQ was also related to grammatical comprehension. As a matter of fact, the relative importance of IQ was slightly larger for grammar than for vocabulary.

Finally, language analytic ability accounts for 10% of the variance in grammar scores and about 8% of the variance in vocabulary scores (although the latter effect is not statistically significant); it also enters into significant interactions with other variables. An examination of graphs plotting the relationship between language analytic ability and linguistic performance (see Figures 6 and 7) shows that above average language analytic ability is always associated with above average grammar and vocabulary scores; however, low language analytic ability is just as likely to be associated with high grammar scores as with low, and the same is true of vocabulary. Our tentative conclusion, then, is that good language analytic skills more or less guarantee highly successful grammatical and lexical development (given normal input, etc.), but other factors can compensate for poor language analytic skills. As shown earlier, these moderating factors include education (in the case of grammar) and nonverbal IQ (in the case of vocabulary), and we can anticipate that a number of other factors are also relevant.

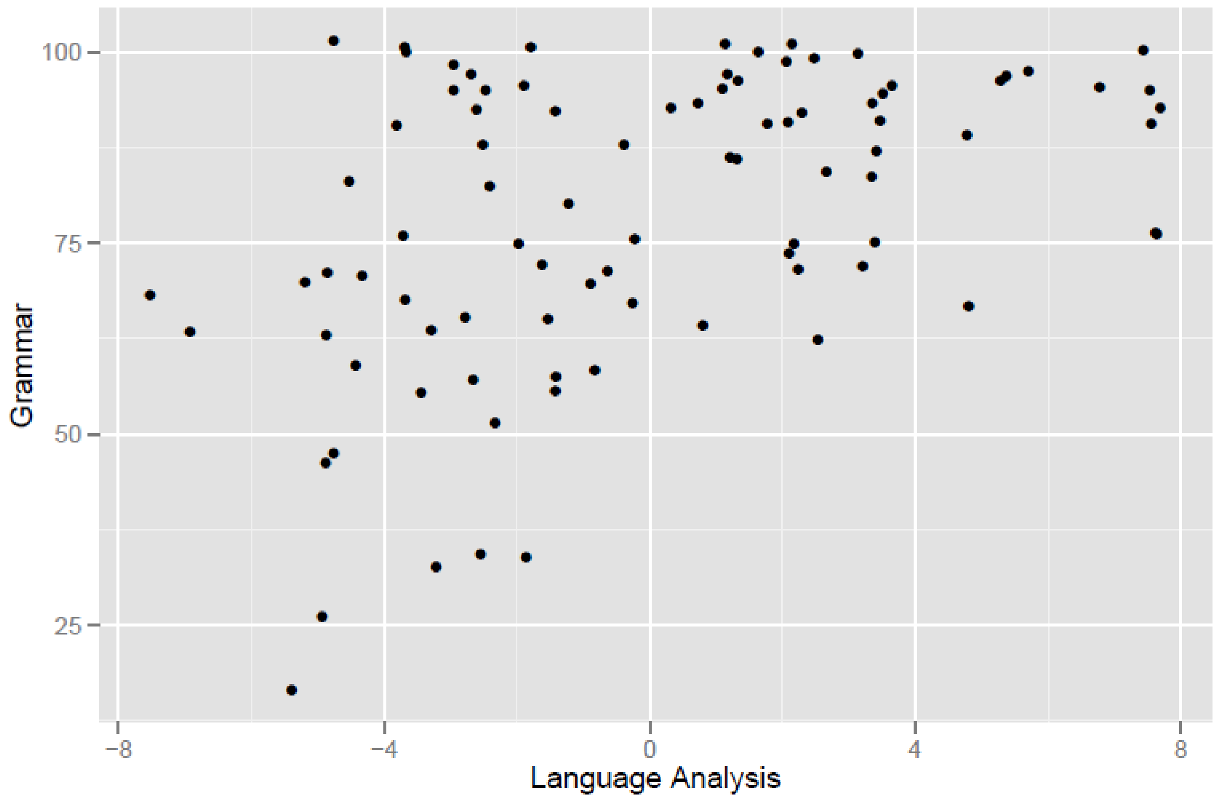

Figure 6. The relationship between language analytic ability and grammar scores

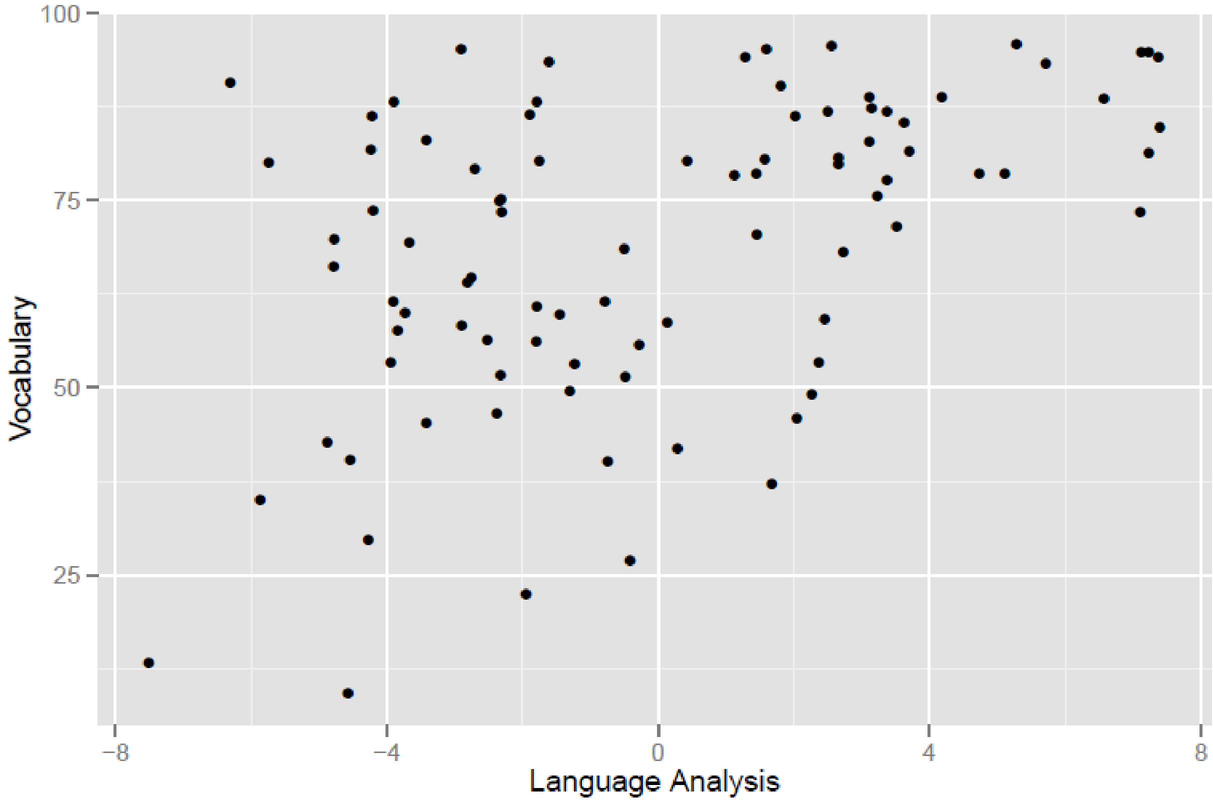

Figure 7. The relationship between language analytic ability and vocabulary scores

4. Conclusion

Several earlier studies have produced evidence that, contrary to the "conventional wisdom" in linguistics, some native speakers have not fully mastered the grammar of their language. As pointed out in the introductory section, incomplete acquisition is associated with low educational attainment, but, because of the sampling strategy employed in the earlier studies, it is not clear how widespread it is. The results of the present study confirm that it is a real phenomenon, and demonstrate that the extent of incomplete acquisition differs from construction to construction. While virtually all native speakers performed at ceiling on the most basic constructions of their language (simple locatives, and actives and other structures with SVO word order such as subject and object relatives), there were substantial individual differences in five of the six experimental conditions: sentences containing subjects with postmodifying PPs, where 9 participants – i.e. 10% of the sample – performed at chance, object clefts (14% at or below chance), object relatives (16%), simple locatives with *every* (41%) and possessive locatives with *every* (40%).

It is important to note that the individual differences observed here involve fairly basic structures, in the sense that they can be found in normal everyday conversations, and that they involve "core" grammar (in the sense of Chomsky, 1981), i.e., deep regularities rather than idiosyncratic or "peripheral" constructions – that is to say, structures which, in the generative paradigm, are thought to depend on fundamental syntactic mechanisms such as movement (object relatives, object clefts and passives), recursion (complex NPs) and computation of quantifier scope (Q-is and Q-has), all of which have been claimed to be part of our innate endowment for language. Moreover, the interpretation of the structures studied here is not subject to dialectal variation, that is to say, it is not the case that *It was the girl that the man fed* is synonymous with *The man fed the girl* in some dialects of English, and with *The girl fed the man* in others, or that *The spoon in the cup is red* means that spoon is red in some dialects and the cup is red in others; therefore, some speakers' failure to correctly interpret some of the experimental sentences cannot be attributed to dialectal differences.

According to usage-based models, individual differences in grammatical knowledge are attributable to, on the one hand, differences in linguistic experience, and on the other, to differences in the cognitive abilities that language learners bring to the acquisition task. The research described here examined the role of two experience-related factors, print exposure and education, and two learner-internal factors, nonverbal IQ and language aptitude. Consistent with earlier research, print exposure was found to have a strong effect on vocabulary and collocations. Its effect on grammar was weaker, but nevertheless significant. The weaker effect on grammar is most likely due the fact that the constructions used in the grammar test were relatively frequent, and hence all participants had had considerable experience with them. Education also contributed to variance in scores on all three linguistic tasks, although its unique effect was relatively modest.

The results of this study also provide clear evidence for a relationship between nonverbal IQ and linguistic knowledge. While it is well-known that vocabulary and nonverbal IQ are correlated (see, for example, Jensen 1998, Raven 2000), it is often claimed that there is no relationship between intelligence and knowledge of basic grammatical constructions: the acquisition of L1 grammar is assumed to rely almost exclusively on implicit learning mechanisms, which are not related to IQ and which are believed not to differ much between individuals (Feldman, Kerr & Streissguth, 1995; Reber, Walkenfeld & Hernstadt 1991; Stanovich 2009). However, as we have seen, the effects of nonverbal IQ on grammar were of about the same magnitude as those observed for vocabulary (in fact, slightly larger). This suggests that general problem-solving abilities are also relevant for first language acquisition – as suggested by usage-based models.

Perhaps the most surprising finding is the robust relationship between language outcomes in native speakers and scores on the Language Analysis test – part of a battery developed to measure *foreign* language aptitude. Native language achievement is not supposed to be related to language aptitude (DeKeyser, 2000; DeKeyser, Alfi-Shabtay & Ravid 2010; Forsberg Lundell & Sandgren

2013; Granena, 2013); yet the correlations of language analytic abilities with grammar scores (0.46) and vocabulary size (0.47) are actually higher than those found in most studies that examined the effect of language analytic ability on foreign language achievement. For example, the correlations between Language Analysis and foreign language achievement (measured using the MLA Cooperative French Tests) ranged from 0.01 for listening to 0.41 for writing (Pimsleur et al., 2004), and the correlations between Language Analysis and foreign language grades ranged from 0.31 to 0.35 (Pimsleur et al., 2004). Sáfár and Kormos (2008) obtained virtually identical results (r = 0.34) when they examined the relationship between Language Analysis scores and EFL proficiency in a group of L1 Hungarian learners; and a recent metaanalysis (Li, 2014) of correlational studies investigating the relationship between language analytic skills and tests tapping foreign language grammar reports a mean correlation of 0.31.[5]

Why then has the relationship between language aptitude and native language attainment gone unnoticed in earlier research? As we saw earlier, it has not gone entirely unnoticed: Skehan and Ducroquet (1988) report a number of significant correlations between measures of early first language development and language aptitude taken more than 10 years later. Furthermore, Bylund, Abrahamsson and Hyltenstam (2010) found a moderately strong correlation (r = 0.52) between language aptitude and performance on a grammaticality judgment task in the heritage language (Spanish) in speakers who emigrated to Sweden before puberty (mean age of arrival 5.7), although a later study by Bylund and Ramírez-Galan (2016) which used a very similar design with participants who had arrived in the Sweden when they were older (24.3 on average) found no significant relationship between aptitude and GJT performance (r = .27, p = .10). Finally, a recent unpublished study by Abrahamsson, discussed in Skehan (2014), also found a significant correlation between grammaticality judgement scores and language aptitude in both native and nonnative speakers.

5 Note that correlations with whole PLAB or MLAT battery and foreign language achievement are usually higher.

However, it seems fair to say that most earlier research did not find a relationship in native speakers simply because it did not look for it. Language aptitude tests were developed to predict foreign language achievement, and have simply not been used by researchers interested in L1 abilities. Most studies that examine the relationship between aptitude and attainment are conducted in the context of work investigating age effects in second language acquisition and used tasks designed for L2 learners. Such tasks are usually very easy for native speakers, and as a result, there is simply not enough variation in their scores to detect a relationship, even if there is one. For example, a frequently cited study by DeKeyser (2000) found a significant correlation between aptitude and performance on a grammaticality judgement task in adult L2 learners but not in younger learners. DeKeyser uses this finding to argue that older learners of higher aptitude rely on explicit learning mechanisms to compensate for their impoverished implicit learning skills. However, the younger learners performed virtually at ceiling, with nearly all scores above 90%, and the majority above 95%. Thus, while DeKeyser's study provides strong evidence for the existence of age effects for some aspects of second language acquisition, it does not speak to the issue of the role of aptitude in first language ultimate attainment. Two other studies which examined the relationship between aptitude and language attainment (Abrahamsson & Hyltenstam, 2008; Granena & Long, 2013) found positive but statistically insignificant correlations in native speakers. Note, however, that the insignificant result is likely to be attributable simply to lack of power. In the Abrahamsson and Hyltenstam study, the correlation coefficient in native speakers was 0.47 (i.e., virtually identical to the one observed in the current study), but it was not significant. However, given that there were only 15 participants in the native control group, the probability of detecting a significant effect if there was one was only 50%. In the Granena and Long study, there were 12 controls and the correlation coefficient was 0.30; thus, power was only 15%. Bylund and Ramírez-Galan (2016) used a considerable larger group of participants (N=39), but with r = .27, power is still only 38%.

Our tentative conclusion, then, is that language aptitude *is* relevant also for first language acquisition, although its effect may not be the same as in second language learning; further research will be necessary to unpick the tangle of relationships between native language attainment, language aptitude, nonverbal ability, and other cognitive and environmental factors.

The research described here also demonstrated the existence of significant correlations between the three aspects of linguistic knowledge studied. Although the correlations were smaller than those observed in studies of early language development, they were still quite robust. Moreover, performance on the grammar and vocabulary tasks was related to the same nonlinguistic predictors, albeit to a different extent. These findings are consistent with usage-based models, according to which speakers' knowledge of grammatical constructions, words and multi-word units such as collocations relies on the same general-purpose learning mechanisms. They are, in contrast, problematic for modular theories, which assume that grammar depends on mental mechanisms which are largely separate from those subserving lexical learning and general problem solving.

It should be noted, however, that performance on the collocations test depended primarily on input measures: the relationships with language analytic ability and with IQ were relatively weak and not statistically significant. This is not surprising, since learning collocations such as *blatant lie* or *hazard a guess* does not require the learner to make inferences about linguistic form and function, but simply to track co-occurrence statistics of individual lexical items; hence, one would expect it to correlate with measures of statistical learning and phonological short-term memory rather than with inferential abilities (and indeed, such correlations have been found in L2 learners: see Forsberg-Lundel & Sandgren, 2013 and Bolibaugh & Foster, 2013). This does not mean that collocations depend on completely different mental mechanisms, since statistical learning is thought to play an important role in grammatical development (Ellis, Römer & O’Donnell, 2016; Kidd, 2012; Misyak, Christiansen & Tomblin, 2010). Further research is needed to determine exactly how inferential abilities on the one hand and statistical learning on another contribute to the acquisition of different aspects of language.

Finally, as we have seen, the correlations between performance on individual grammatical constructions are relatively low (in comparison with correlations between overall scores on the grammar, vocabulary and collocations tasks). Although these findings must be treated with caution due to the small number of items per construction, they suggest that knowledge or individual constructions may be relatively independent, and may depend differentially on different cognitive abilities. There is some evidence for this in the literature: Street and Dąbrowska (2010), for example, found that although scores on both passives and sentences with universal quantifiers correlated with both print exposure and need for cognition (how much people enjoy effortful cognitive activities) the former was a better predictor of performance on passives and the latter, on quantifiers. This makes sense: passives are more frequent in written discourse, so we would expect speakers who read more to perform better on this construction; and conversely, the computation of quantifier scope is arguably more difficult than establishing who did what to whom, so we would expect individuals who avoid effortful cognitive activities to be less likely to master the relevant routines. Thus, future research should examine the relationships between specific cognitive abilities and individual speakers' proficiency with specific constructions.

Acknowledgements

I would like to thank Sarah Duffy, Steven McCarthy and Alex Nicholson for help with the data collection and Dagmar Divjak, Nick Riches and three anonymous reviewers for their comments on earlier drafts of this paper. I am also grateful to Isla Jones for artwork in the "Pictures and Sentences" test.

REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30**,** 481-509.

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

Barlow, M., & Kemmer, S. (2000). *Usage-Based Models of Language.* Cambridge: Cambridge University Press.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From First Words to Grammar: Individual Differences and Dissociable Mechanisms.* Cambridge: Cambridge University Press.

Beglar, D. (2009). A Rasch-based validation of the Vocabulary Size Test. *Language Testing* 26**,** 1-22.

Biber, D. (1995). *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

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.

Bolibaugh, C., & Foster, P. (2013). Memory-based aptitude for nativelike selection. In G. Granena & M. Long (Eds.), *Sensitive Periods, Language Aptitude and Ultimate L2 attainment* (pp. 179–204). Amsterdam/Philadelphia: John Benjamins.

Brooks, P.J., & Sekerina, I.A. (2006). Shallow processing of universal quantification: A comparison of monolingual and bilingual adults. *Proceedings of the 28th Annual Conference of the Cognitive Science Society***,** 2450.

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language* 82**,** 711-733.

Bybee, J.L. (2010). *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Bybee, J.L. (2013). "Usage-based theory and exemplar representation," in *The Oxford Handbook of Construction Grammar,* eds. T. Hoffman & G. Trousdale. (Oxford: Oxford University Press), 49-69.

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2010). The role of language aptitude in first language attrition: The case of pre-pubescent attriters. *Applied Linguistics, 31*(3), 443–464.

Bylund, E., & Ramírez-Galan, P. (2016). Language Aptitude in First Language Attrition: A Study on Late Spanish-Swedish Bilinguals. *Applied Linguistics*, *37*(5), 621–638.

Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and Child Directed Speech. *First Language, 33,* 268–279.

Carroll, J.B., & Sapon, S.M. (1959). *Modern Language Aptitude Test.* Washington, DC: Second Language Testing Incorporated.

Carroll, J.B., & Sapon, S.M. (1967). *Modern Language Aptitude Test (EMLAT): Elementary: Form EA.* Psychological Corporation.

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 Variations in Syntactic Competence.* Basingstoke: Palgrave.

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

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

Chomsky, N. (1981/1993). *Lectures on Government and Binding: The Pisa Lectures.* Berlin: Walter de Gruyter.

Chomsky, N. (1999). On the nature, use, and acquisition of language. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Child Language Acquisition* (pp. 33–54). San Diego: Academic Press.

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

Crain, S., Thornton, R., & Murasugi, K. (2009). Capturing the evasive passive. *Language Acquisition*, *16*, 123–133.

Crain, S., Thornton, R., & Murasugi, K. (2009). Capturing the evasive passive. *Language Acquisition* 16**,** 123-133.

Crawley, M.J. (2007). *The R Book.* New York: Wiley.

Cunningham, A., & Stanovich, K. (1998). What reading does for the mind. *American Educator*, 1–8.

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

Dąbrowska, E. (2004). *Language, Mind and Brain. Some Psychological and Neurological Constraints on Theories of Grammar.* Edinburgh: Edinburgh University Press.

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

Dąbrowska, E. (2009). "Words as constructions," in *New Directions in Cognitive Linguistics,* eds. V. Evans & S. Pourcel. (Amsterdam: John Benjamins), 201-223.

Dąbrowska, E. (2012a). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2**,** 219-253.

Dąbrowska, E. (2012b). Explaining individual differences in linguistic proficiency. *Linguistic Approaches to Bilingualism* 2**,** 324-335.

Dąbrowska, E. (2013). Functional constraints, usage, and mental grammars: A study of speakers' intuitions about questions with long-distance dependencies. *Cognitive Linguistics* 24**,** 633 - 665.

Dabrowska, E. (2014a). Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25**,** 617-653.

Dąbrowska, E. (2014b). Words that go together: Measuring individual differences in native speakers' knowledge of collocations. *The Mental Lexicon*, *9*, 401–418.

Dąbrowska, E. (2015). "Individual differences in grammatical knowledge," in *Handbook of Cognitive Linguistics,* eds. E. Dąbrowska & D. Divjak. (Berlin De Gruyter Mouton), 649-667.

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.

DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31**,** 413-438.

DeKeyser, R.M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22**,** 499-533.

Department for Business Innovation and Skills (2011). "2011 Skills for Life Survey: Headline Findings". Department for Business Innovation and Skills.

Dionne, G., Philip S Dale, Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development* 74**,** 394-412.

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, *25*, 108–127.

Dionne, G., Philip S Dale, Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development*, *74*, 394–412.

Divjak, D., & Cardwell-Harris, C. (2015). Frequency and entrenchment. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 53–75). Berlin: De Gruyter Mouton.

Ellis, N.C. (2007). "Implicit and explicit knowledge about language," in *Encyclopedia of Language and Education. Volume 6: Knowledge about Language,* eds. J. Cenoz & N.H. Hornberger. Springer), 1-13.

Ellis, N.C., Römer, U., & O'Donnell, M. B. (2016). *Usage-based Approaches to Language Acquisition and Processing: Cognitive and Corpus Investigations of Construction Grammar*. Malden, MA: Wiley-Blackwell.

Feldman, J., Kerr, B., & Streissguth, A.P. (1995). Correlational analyses of procedural and declarative learning performance. *Intelligence* 20**,** 87-117.

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

Forsberg Lundell, F., & Sandgren, M. (2013). "High-level proficiency in late L2 acquisition: Relationship between collocational production, language aptitude and personality," in *Sensitive Periods, Language Aptitude and Ultimate L2 attainment* eds. G. Granena & M. Long. (Amsterdam/Philadelphia: John Benjamins), 231-258.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, *7*, 219–224.

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

Granena, G. (2013). "Reexamining the robustness of aptitude in second language acquisition," in *Sensitive Periods, Language Aptitude and Ultimate L2 attainment* eds. G. Granena & M. Long. (Amsterdam/Philadelphia: John Benjamins), 179-204.

Granena, G., & Long, M.H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29**,** 311-343.

Green, P. (1975). *The Language Laboratory in School: The York Study.* London: Oliver and Boyd.

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software* 17.

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician* 61**,** 139-147.

Hachmann, W., Müller, D., & Konieczny, L. (2009). Individual differences in the processing of complex sentences. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 309–314). Austin, TX.

Herschensohn, J. (2009). Fundamental and gradient differences in language development. *Studies in Second Language Acquisition* 31**,** 259-289.

Hulstijn, J. H. (2015). *Language Proficiency in Native and Non-native Speakers: Theory and Research*. Amsterdam: John Benjamins.

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

Indefrey, P., & Goebel, R. (1993). The learning of weak noun declension in German: Children vs. artificial network models. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society***,** 575-580.

Jensen, A.R. (1998). *The g Factor: The Science of Mental Ability.* Westport, CT: Praeger Publishing.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*, 171–184.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2017). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169.

Langacker, R.W. (1988). "A usage-based model," in *Topics in Cognitive Linguistics,* ed. B. Rudzka-Ostyn. (Amsterdam: John Benjamins), 127-161.

Langacker, R.W. (2000). "A dynamic usage-based model," in *Usage-Based Models of Language,* eds. M. Barlow & S. Kemmer. (Stanford, CA: CSLI Publications), 1-63.

Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using R.*

Li, S. (2014). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics* 36, 385-408.

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

Lindeman, R.H., Merenda, P.F., & Gold, R.Z. (1980). *Introduction to Bivariate and Multivariate Analysis.* Glenview, IL: Scott, Foresman.

Meara, P. (2005). *LLAMA Language Aptitude Tests: The Manual.* Swansea: Lognostics.

Misyak, J., Christiansen, M., & Tomblin, J. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1*.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor.* Amsterdam: John Benjamins.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher* 31**,** 9-13.

Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational v. expository discourse: A study of syntactic development in children, adolescents and adults. *Journal of Speech, Language and Hearing Research*, *48*, 1048–1064.

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

Pimsleur, P., Reed, D.J., & Stansfield, C.W. (2004). *Pimsleur Language Aptitude Battery: Manual 2004 Edition.* Bethesda, MD: Second Language Testing, Inc.

Pinker, S. (1997). Words and rules in the human brain. *Nature* 387**,** 547-548.

Pinker, S. (1999). *Words and Rules. The Ingredients of Language.* London: Weidenfeld and Nicolson.

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology* 41, 1-48.

Ravid, D. (2004). Emergence of linguistic complexity in later language development: Evidence from expository text construction. In D. Ravid & H. Bat-Zeev Shyldkrot (Eds.), *Perspectives on Language and Language Development* (pp. 337–356). Dordrecht: Kluwer.

Reber, A.S. (2011). "An epitaph for Grammar: An abridged history," in *Implicit and Explicit Language Learning: Conditions, Processes, and Knowledge in SLA and Bilingualism,* eds. C. Sanz & R. Leow. (Washington, DC: Georgetown University Press), 23-34.

Reber, A.S., Walkenfeld, E.F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning Memory and Cognition* 17**,** 888-896.

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.

Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics in Language Teaching* 46**,** 113-136.

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

Shipley, W.C., Gruber, C.P., Martin, T.A., & Klein, A.M. (2009). *Shipley-2 Manual.* Los Angeles: Western Psychological Services.

Skehan, P. (2014). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics* 36, 367-384.

Skehan, P., & Ducroquet, L. (1988). "A Comparison of First and Foreign Language Ability". (London: ESOL Department, Institute of Education, London University: Working Documents No. 8).

Smith, N. (1999). *Chomsky: Ideas and Ideals.* Cambridge: Cambridge University Press.

Smith, N., & Tsimpli, I. M. (1995). *The Mind of a Savant. Language Learning and Modularity*. Oxford: Blackwell.

Stanovich, K.E. (2009). "Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? ," in *In Two Minds: Dual Processes and Beyond,* eds. J.S.B.T. Evans & K. Frankish. (Oxford: Oxford University Press), 55–88.

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

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

Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2009). Associations between lexicon and grammar at the end of the second year in Finnish children. *Journal of Child Language* 36**,** 779-806.

Street, J. A. (2010). *Individual differences in comprehension of passives and universal quantifiers by adult native speakers of English.* University of Sheffield, Sheffield.

Street, J. A. (2017). This is the native speaker that the non-native speaker outperformed: Individual, education-related differences in the processing and interpretation of Object Relative Clauses by native and non-native speakers of English. *Language Sciences*, *59*, 192–203.

Street, J., & 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.

Street, J., & Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholinguistics* 35**,** 97-118.

Szagun, G., Steinbrink, C., Franik, M., & Stumper, B. (2006). Development of vocabulary and grammar in young German-speaking children assessed with a German language development inventory. *First Language* 26**,** 259-280.

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

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*, 231–270.

Ullman, M.T. (2006). The declarative/procedural model and the shallow structure hypothesis. *Applied Psycholinguistics* 27**,** 97-105.

Verhoeven, L., Aparici, M., Cahana-Amitay, D., Hell, J. V., Kriz, S., & Viguié-Simon, A. (2002). Clause packaging in writing and speech: A cross-linguistic developmental analysis. *Written Language & Literacy*, *5*, 135–161.

Wells, G. (1985). *Language Development in the Preschool Years.* Cambridge: Cambridge University Press.

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.