Experience, aptitude and individual differences in linguistic attainment

Dabrowska, Ewa

DOI:
10.1111/lang.12323

Running head: EXPERIENCE, APTITUDE AND INDIVIDUAL DIFFERENCES

Experience, aptitude and individual differences in linguistic attainment: A comparison of L1 and L2 speakers

Ewa Dąbrowska

University of Birmingham

Author note

Ewa Dąbrowska, Department of English Language and Linguistics, University of Birmingham, 3 Elms Road, Edgbaston, Birmingham, B15 2TT, United Kingdom

email: e.dabrowska@bham.ac.uk

Acknowledgments

I would like to thank Sarah Duffy, Steven McCarthy and Alex Nicholson for help with the data collection and Sible Andringa, 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.

Abstract

This paper compares the performance of native speakers and adult L2 learners on tasks tapping proficiency in three linguistic domains: grammar, vocabulary and collocations. In addition, data was collected on several predictors of individual differences in linguistic attainment, including some related to language experience (print exposure, education, and, for L2 speakers, length of residence and use of English) and some relating to an individual's aptitude to learn (language analytic ability and nonverbal IQ), as well as age and, for L2 speakers, age of arrival. As anticipated, the native group outperformed L2 speakers on all three language measures, although the effect sizes were much larger for collocations than for grammar or vocabulary. Crucially, there were vast individual differences in both groups, and considerable overlap between groups, particularly for grammar. Regression analyses revealed both similarities and differences between native and non-native speakers in which non-linguistic measures best predict performance on the language tasks.

Introduction

Most language acquisition researchers take it for granted that child language learning is successful: that is to say, except in impaired populations or in cases of severe deprivation all learners converge relatively rapidly on (more or less) the same grammar (see, for example Bley-Vroman, 2009: 179; Chomsky, 1975: 11; Lidz & Williams 2009: 177). This outcome contrasts sharply with the outcome of L2 acquisition, which is characterized by large individual differences, particularly in adult learners. Furthermore, adult learners rarely, if ever, attain native-like competence. Such differences between first and second language acquisition are often attributed to a biologically determined critical period, and are used to

support claims of a "fundamental difference" between the two processes (see, for example, Bley-Vroman 2009).

However, a number of recent studies have shown that native speaker convergence is a myth: there are, in fact, considerable individual differences in adult L1 speakers' linguistic competence (for recent reviews, see Dąbrowska, 2012, 2015; Farmer, Misyak & Christiansen, 2012; Hulstijn, 2015; Kidd, Donnelly and Christiansen 2017). Individual differences have been found in speakers' mastery of aspects of morphology and complex syntax, but also for relatively simple grammatical structures such as passives (Dąbrowska & Street, 2006; Street & Dąbrowska, 2010, 2014), object relatives (Street 2017), and quantifiers (Brooks & Sekerina 2006; Street & Dąbrowska, 2010). Many, though not all, of the observed differences are related to education; and when education-related differences are observed they show the same characteristic pattern, with highly educated participants performing at or near ceiling and less educated participants showing a wide distribution of scores. It is important to note that the differences reported in these studies cannot be attributed simply to some participants' failure to understand the task, uncooperativeness, or some other linguistically irrelevant performance factors, as all the studies mentioned included control conditions designed to rule out such interpretations (see Dąbrowska 2012 for further discussion).[1]

Commenting on the implications of the existence of individual differences in native speakers' grammatical knowledge for work on second language acquisition, DeKeyser (2012: 260) suggests that researchers interested in age effects should "avoid structures for which quite a bit of variability [in L1 speakers] has been documented; otherwise it is a foregone conclusion that the ranges of L1 and L2 variation are going to overlap". However, this recommendation seems unwarranted. First, overlap between L1 and L2 ranges for such structures is not a "foregone conclusion": if a structure is difficult for native speakers, we could reasonably expect it to be even more difficult, and perhaps unlearnable, for adult

second language learners. Secondly, while this may be a good strategy for researchers who want to demonstrate the existence of age effects, it is not advisable if our goal is to understand the nature of language acquisition in its entirety: to accomplish the latter, we need to study all types of structures. Finally, and most importantly, given that there is variability even in fairly basic structures, there may not be much left to study if we avoid structures for which there is variability in native speakers!

Thus, current work on second language attainment presents a distorted picture in that it typically focuses on structures which are known to be difficult for L2 learners (which is often explicitly acknowledged, cf. Granena & Long, 2013). Furthermore, the vast majority of studies use highly educated participants, which tends to exaggerate differences between native and non-native speakers (Andringa, 2014; Dąbrowska, 2012).

The main aim of the present study is to redress the balance by testing participants recruited from a variety of socioeconomic backgrounds on a range of grammatical structures which differ in difficulty, including some which are difficult for native speakers as well. This ensures that the results are more representative, as well as making it highly likely that there will be some variation in both language groups, which will make it possible to carry out meaningful analyses of possible reasons for individual differences.

In addition, the study also collected measures of vocabulary size and collocational knowledge. These are interesting for two main reasons. First, while age effects in grammar and phonology are often attributed to the existence of a critical period, vocabulary learning is generally assumed to rely on general learning mechanisms which remain active throughout adulthood. However, some researchers (e.g. Granena & Long, 2013; Long, 2013; Spadaro, 2013) have argued that critical period effects are also observable in the lexical domain; and there is considerable evidence that L2 learners have particular difficulty with acquiring a large store of collocations (Bolibaugh & Foster, 2013; Hoffman & Lehmann, 2000; Laufer &

Waldman, 2011). To my knowledge, however, no one has proposed that vocabulary size is subject to age effects. Secondly, it will help us understand the relationship between these three aspects of linguistic knowledge in both populations. It is well established that there is a strong relationship between vocabulary and grammar in early L1 acquisition (Bates, Bretherton & Snyder, 1988; Huttenlocher, 1998; Szagun et al., 2006). However, according to modular theories (e.g. Chomsky, 1981; Pinker, 1997, 1999; Ullman, 2006) these aspects of language rely on different mechanisms and become dissociated later in development. According to usage-based theories (e.g. Barlow & Kemmer, 2000; Bybee, 2010; Langacker, 1988), in contrast, all symbolic units rely on the same mental mechanisms (albeit possibly to different extents). Thus, usage-based theories predict that performance on measures of grammar, vocabulary and collocations should be correlated and subject to similar restrictions, while modular theories predict no relationships between grammar and lexis. Moreover, according to modular theories, lexis, or at least vocabulary size, should correlate with IQ and with measures of exposure, but grammar should not. Comparing the performance of L1 and L2 speakers in all three domains will shed light on all these issues.

The second aim of the study is to examine the relationship between linguistic abilities and four variables which are potentially relevant for both L1 and L2 speakers – print exposure, IQ, language aptitude, and education – and four variables which are relevant in an L2 context only -- age of first exposure, age of arrival, length of residence, and use of English.

Print exposure is known to correlate strongly with vocabulary size (Cunningham & Stanovich, 1998; Stanovich & Cunningham, 1992), and it would be reasonable to expect it to correlate with collocational knowledge as well (Dąbrowska, 2014a). There is relatively little work on the relationship between print exposure and sensitivity to grammatical structure, although a few studies (Dąbrowska & Street, 2006; Montag & MacDonald, 2015; Street &

Dąbrowska, 2010, 2014a; Wells et al., 2009) suggest that it may be relevant, particularly for constructions such as passives and object relatives which occur more frequently in written language.

IQ and language aptitude are known to be relevant for foreign and second language learning, particularly in classroom settings (Ehrman & Oxford, 1995; Li, 2014; Sasaki, 1999; Sparks et al., 2011), but are supposedly irrelevant for child first language acquisition, which is thought to depend almost entirely on implicit learning; thus, evidence of a link between intelligence and/or language aptitude in child L1 acquisition would undermine the fundamental difference claim. Both IQ and language aptitude are multifaceted concepts and it is beyond the scope of this paper to measure all aspects of either. Therefore, the decision was taken to focus specifically on nonverbal or 'fluid' intelligence on the one hand, and language analytic ability on the other. Fluid intelligence, or the ability to solve novel problems, is a 'purer' measure of ability than crystallized, or verbal, intelligence, which depends strongly on knowledge and experience. Furthermore, the existence of a relationship between fluid intelligence and language would also be theoretically more interesting (crystallized or verbal intelligence is measured using vocabulary and comprehension tests, hence we can expect a correlation with linguistic knowledge). Language analytic ability, or the capacity to infer linguistic rules and generalizations, is the aspect of language aptitude which is arguably most relevant for grammar, which is the main focus of this paper (cf. Li 2014).

Education has been shown to predict linguistic abilities, including knowledge of grammar, in both L1 (Chipere, 2003; Dąbrowska & Street, 2006; Street & Dąbrowska, 2010) and L2 (Tarone, Bigelow & Hansen 2009; Hakuta et al., 2003). However, it is unclear whether these effects are attributable to education per se, or to other factors – such as print exposure, IQ and language aptitude – that correlate with it. Examining the role of all four variables in the same study will help to elucidate this issue.

Finally, examining factors which are unique to second language learners (age of first exposure, age of arrival, length of residence, and use of English) will add to a debate which has been raging in the second language literature for decades, namely, whether the failure of most second language learners to attain native-like competence should be attributed to maturational factors, amount of exposure to the L2, or some combination of the two (see, for example, Abrahamsson & Hyltenstam 2009; Birdsong, 2006; DeKeyser et al., 2012; Flege, 2009; Flege & Liu, 2001; Granena & Long, 2013; Hakuta et al., 2003; Johnson & Newport, 1989; Long, 2013).

Unlike earlier studies addressing similar issues, which tended to use grammaticality judgement tasks (GJTs), the present study used a picture selection task. Picture selection was chosen for two reasons. First, although – like nearly all experimental tasks – it is not very natural, it was thought to be more similar to ordinary language use than a GJT in that it involves a judgement based on meaning rather than a metalinguistic judgment about form. Secondly, and more importantly, it is a more suitable method of assessing linguistic knowledge in the target population, which included low-educated participants. Picture selection poses relatively small cognitive demands on the testee, and is often used in experiments with young children and for clinical assessment. For example, the Test for Reception of Grammar (Bishop, 2003), which is routinely used in clinical practice to diagnose language impairment in children, uses picture selection. Grammaticality judgments, in contrast, are notoriously difficult for children and illiterate speakers (Karanth & Suchitra, 1993), and hence could pose some problems for the low-educated participants as well.

Method

Participants

90 native speakers (42 male and 48 female) and 67 non-native speakers of English (21 male and 46 females) were recruited through personal contacts, church and social clubs, and advertisements in local papers. Participants were told that the purpose of the study was to examine individual differences in native and non-native speakers' knowledge of English, and whether these differences are related to their linguistic experience and abilities. All participants signed a written consent form before the research commenced.

The L1 participants were all born and raised in the UK and were selected to ensure a range of ages, occupations and educational backgrounds. The age range was from 17 to 65 (mean 38, SD 16). 22% of the participants held manual jobs, 24% held clerical positions, and 28% had professional-level jobs or were studying for a degree; the remaining 26% were occupationally inactive, i.e., unemployed, retired or housewives. In terms of education, participants' backgrounds ranged from no formal qualifications to PhD, with corresponding differences in the number of years spent in full-time education (from 10 to 21, mean 14, SD 2). Six participants reported a working knowledge of another language; the rest described themselves as monolinguals.

The non-native participants ranged in age from 20 to 62 (mean 33, SD 9). Ages of arrival ranged from 16 to 49 (mean 25, SD 8) and length of residence from 3 to 42 years (mean 7, SD 6). They came from a variety of language backgrounds, including Polish, Russian, Lithuanian, Mandarin/Cantonese, German, Greek, French, Italian, Spanish, and Malaysian. The participants varied widely in terms of educational attainment (from 8 to 24 years spent in full time education), although as a group they had more schooling than the

native participants (mean 16 years, SD 3). 39% of the participants obtained all of their educational qualifications in their home country; 45% had up to three years education in the UK or another English-speaking country; and 16% had studied in an English-speaking country for 4 or more years. 31% of the sample various manual jobs, 49% had professional level jobs or well full-time students, and 15% held clerical posts; the remaining 4% were retired, unemployed or housewives.

Materials

The testing materials included three language tests designed to tap participants' knowledge of grammar, vocabulary and collocations, and measures of nonverbal IQ, language analytic ability, and print exposure. In addition, participants completed a background questionnaire which included questions about education, use of English, and reading habits.

Grammatical comprehension

Grammatical comprehension was tested using the Pictures and Sentences test (available from https://www.irisdatabase.org/iris/app/home/detail?id=york:935511; for a detailed description, see Dąbrowska, 2018). Pictures and Sentences assesses comprehension of 10 grammatical constructions of varying degrees of difficulty (see Table 1) using a picture selection task. Participants are presented with a sentence and two pictures and asked to choose the picture that matches the sentence. There are 8 items for each structure, giving a total of 80 items.

Table 1

Constructions tested by Pictures and Sentences

Construction	**Example**
Active	The boy scratched the dancer.
Passive	The dancer was scratched by the boy.
Subject cleft	It was the boy that scratched the dancer.
Object cleft	It was the dancer that the boy scratched.
Subject relative	The boy was the one who scratched the dancer.
Object relative	The dancer was the one that the boy scratched.
Simple locative	The lamp is on the table.
Locative w/ quantifier	Every lamp is on a table.
Possessive locative w/ quantifier	Every table has a lamp on it.
Postmodifying PP	The lamp on the table is white.

Receptive vocabulary

Receptive vocabulary was measured using a shortened version of the Vocabulary Size Test (Nation & Beglar, 2007). The original test systematically samples 14 frequency levels, with 10 items for each level, for a total of 140 items. Since participants in this study had to complete five other tasks, the test was shortened in the following way: 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 shorter version contained 60 items.

Each item of the test consists of a target word presented in a non-defining context followed by four simple definitions, as in example (1) below; the participant's task is to choose the correct definition.

(1) remedy: We found a good **remedy**.

a. way to fix a problem

b. place to eat in public

c. way to prepare food

d. rule about numbers

Collocations

Collocational knowledge was assessed using the Words that Go Together test (Dąbrowska, 2014a). This instrument consists of 40 sets of five short phrases (see (2) and (3) below for examples; target answers are *blank expression* and *achieve one's objectives*) of which one is an established collocation and the other four plausible alternatives. Participants are instructed to select from each set the phrase which "sounds the most natural or familiar". The target items vary in frequency (from 0.06 to 6.19 per million words in the BNC) and in collocation strength (mutual information from 4.4 to 15.6). The distractors all have a mutual information score of less than 2.0.

(2) blank expression

frightful expression

plain expression

sinister expression

terrible expression

(3) achieve one's objectives

complete one's objectives

finish one's objectives

follow one’s objectives

tackle one’s objectives

Nonverbal IQ

Nonverbal IQ was assessed using the Shipley-2 Block Patterns test (Shipley et al., 2009), which is a pen-and-paper version of Kohs' Block Design test in which participants are required to replicate patterns of black-and-white squares. There are 12 items in total, some with several subparts, so the maximum possible score is 27.

Language analytic ability

Language analytic ability was measured using the Language Analysis subtest of the Pimsleur Language Aptitude Battery (PLAB, Pimsleur et al., 2004). The PLAB was developed to predict achievement in foreign language learning in classroom settings, although it is also known to correlate with learning in more naturalistic settings. In the Language Analysis subtest, participants are presented with some vocabulary and sentences in an unknown language and asked to predict the form of a novel sentence (which they choose from an array of four alternatives). To be able to do this, the participant must be able to determine which chunks of form in the model sentence correspond to which chunks of meaning and then construct a new form by performing an analogical mapping. Thus, although the test does not require knowledge of grammatical terms such as 'subject', 'agreement' or 'case marking', it is strongly metalinguistic in the sense that it involves explicit reasoning about language. The test contains 15 items; thus, scores can vary from 0 to 15.

Print exposure

Finally, print exposure was measured using the Author Recognition Test (ART, Acheson et al., 2008). The test consists of a list of 130 names, half of which are names of real authors. The participants' task is to mark the names that they know to be those of published authors. To penalize guessing, the score is computed by subtracting the number of foils from the number of real authors selected. Thus, the maximum possible score is 65, and the minimum score could be negative if a participant selects more foils than real authors. When this happened, the negative number was replaced with 0.

The Author Recognition Test has been shown to a valid and reliable measure of print exposure, which, unlike questionnaire-based measures is not contaminated by socially desirable responses and assesses lifetime reading experience as opposed to current reading (see Acheson et al., 2008; Stanovich and Cunningham 1992).

Background questionnaire

The background questionnaire for native speakers included questions about age, gender, education (highest qualification and the number of years spent in full-time education), occupation, linguistic background, and reading habits. In the reading habits part of the questionnaire, participants were asked to estimate the number of hours they spent reading in a typical week and in the preceding week. The estimates were to include reading any type of written or printed material, including novels, newspapers, magazines, course books, poetry, blogs, e-mails, instruction manuals, etc. Participants were asked to circle one of the following options: less than 1 hour, 1-5 hours, 5-10 hours, 10-15 hours, 15-20 hours and more than 20 hours. These choices were assigned scores of 0, 1, 2, 3, 4 and 5 respectively. The third question asked how much participants read compared to their friends: much less (0 points), a little less (1 point), about the same (2 points), a little more (3 points) and much more (4

points). The reading score was computed by summing the points for all three questions.

The L2 speakers answered the same questions, but the questions about education and reading habits were separated into two parts, one for education and reading in English, and the other for their native language. In addition, they were asked to provide information about their native language, the age at which they started learning English, the age at which they started living in an English-speaking country, the total number of years spent living in an English-speaking country, and use of English at work/college/university and in their private lives, i.e. with friends and family. For the latter two questions, the options were: less than 20% of the time, 20-40%, 40-60%, 60-80%, and over 80% of the time. These were scored 0, 1, 2, 3, 4 and 5 respectively.

Procedure

The participants were tested individually in a quiet room by a research assistant. The participants completed the background questionnaire first, followed by the language tasks and the three cognitive tasks in the same order as described in the preceding section. The language tasks were administered both orally and in writing: the RA read the questions out loud while the participant followed on their copy of the test and responded orally, and the RA recorded their responses. This was done in order to avoid complications due to poor literacy skills, and to make sure that participants answered all questions, even if they said they did not know the answers. Ensuring that participants supplied answers to all questions was important, as earlier work (see for example e.g. Dąbrowska, 2014a, 2014b) had shown that people often guess at well above chance levels (particularly on the collocations task) even when they believe they do not know the answer; thus encouraging them to guess provides a more accurate measure of their performance. The non-linguistic tasks were administered in writing,

following the instructions given in the test manuals. The entire testing session lasted 1.5 to 2 hours, with short rest periods between tasks.

Results and discussion

This section is organized as follows. First, I compare the performance of native and non-native speakers on the three linguistic tasks. This will be followed by analyses of the correlations between all the variables, focusing in particular on relationships between the three linguistic measures in natives and non-natives (which is relevant to the modularity debate) and on correlations between the predictor variables in non-natives (which prepares the ground for the analyses of the relationships between these variables and the linguistic outcomes). Finally, I conduct two sets of regression analyses. The first set examines the role of factors which were previously shown to predict performance on the grammar, vocabulary and collocations tasks in native speakers, namely education, print exposure, nonverbal IQ and language aptitude (see Dąbrowska, 2018), as well as age, which, as we will see, shows a different pattern of relationships with linguistic abilities in the two language groups. The second set of regressions examines the relationship between linguistic abilities and some additional predictors which are relevant only for non-native speakers, namely age of first exposure, age of arrival, education in native language, education in English, use of English, and length of residence.

Overall performance

The descriptive statistics for the three language tasks are given in Table 2; see also Table S1 for raw scores on all measures. As explained earlier, participants were asked to

choose from an array of 2 in the grammar task, an array of 4 in the vocabulary task, and an array of 5 in the collocations task. Thus, chance performance was 40/80, or 50%, for grammar; 15/60, or 25%, for vocabulary and 8/40, or 20%, for collocations. To facilitate comparisons across tasks, all scores were corrected for guessing and converted into percentages using the following formula:

Corrected score = 100*(raw score – chance)/(perfect-chance)

where *chance* is the predicted number of items obtained simply by guessing (40 for grammar, 15 for vocabulary and 8 for collocations), and *perfect* represents the maximum possible score (80 for grammar, 60 for vocabulary and 40 for collocations). Thus, a corrected score of 0 means that the participant was at chance; a corrected score of 100 means that a participant gave target responses for all items; and a negative score means that a participant performed below chance level. Table 2 also provides information about the reliability of the language tests in both populations. All tests are highly reliable (KR-20 > .80).

As expected, L1 speakers performed better than L2 speakers on all three tasks. The differences were moderately large for grammar ($t = 3.9$, $df = 79.6$, $p < .001$, $d = 0.70$) and vocabulary ($t = 4.7$, $df = 133.8$, $p < .001$, $d = 0.77$) and very large for collocations ($t = 9.6$, $df = 117.2$, $p < .001$, $d = 1.61$). However, as shown in Figure 1, there were large individual differences in performance in L1 as well as L2 speakers, and considerable overlap between the groups. As a matter of fact, 75% of the L2 speakers performed within the native speaker range on the grammar task, and 51% performed above the native mean. For vocabulary, the corresponding figures were 94% and 28% respectively, and for collocations, 69% and 6%.

Table 2

Scores (% corrected for guessing) on the three language tasks in L1 and L2 speakers

	Mean	SD	Median	Range	IQR	KR-20
Grammar-L1	86	13	90	45-100	78-98	.82
Grammar-L2	69	34	88	0-100	46-95	.96
Vocab-L1	69	20	74	9-96	56-86	.96
Vocab-L2	53	23	53	-13-91	38-71	.97
Colloc-L1	66	20	69	9-97	54-84	.81
Colloc-L2	29	26	19	-16-88	6-55	.89

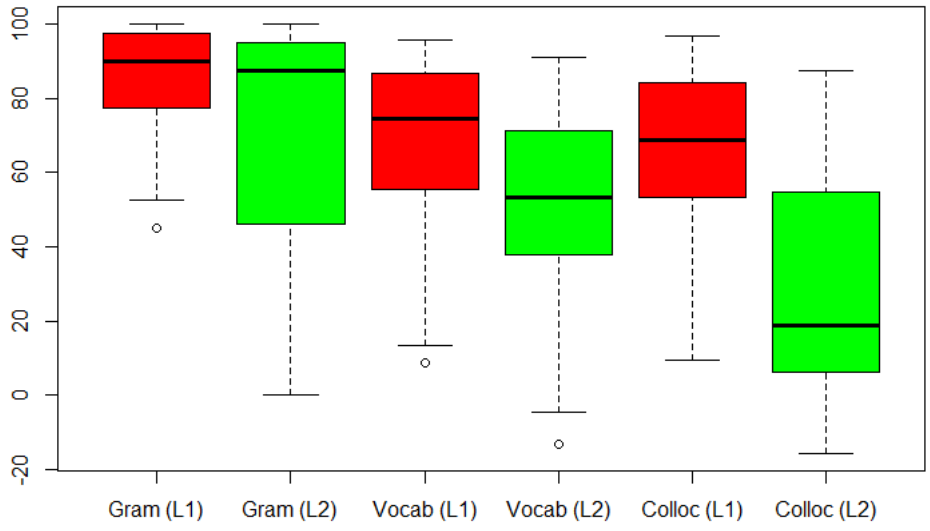

Figure 1. Distribution of scores on the language tests (N=90 for the L1 group and 67 for L2 group)

Correlational analyses

Correlation matrices showing the relationships between the linguistic measures and background measures are provided in Table S2 (Natives) and Table S3 (Non-natives). In native speakers, there are significant positive relationships between all three language measures, with correlation coefficients ranging from .36 for grammar and collocations to .57 for vocabulary and collocations. Likewise, there are positive correlations between all pairs of predictor variables except Blocks (i.e., nonverbal IQ) and the two measures of print exposure. Of the two measures of print exposure, ART shows stronger correlations with variables that are known to correlate with reading, viz. vocabulary and education, which suggests that this is a more accurate measure for native speakers than the self-report questionnaire, and accordingly was used in the regression analyses reported below.

For L2 speakers, the situation is somewhat different. There are significant correlations between scores on grammar and collocations (r=.50) and vocabulary and collocations (r=.41), but not between grammar and vocabulary. This may be due to the fact that scores for grammar and vocabulary depend most strongly on the amount and type of language instruction the learner has received, as language teaching often targets specific components (i.e., either grammar or vocabulary) more than others. As in native speakers, there was no correlation between nonverbal IQ and measures of print exposure. In fact, ART did not correlate with anything except (negatively) age of first exposure – not even with self-reported reading in English. This suggests that it is not a suitable measure of reading exposure for L2 speakers, who may score well on the test because they have read many of the authors in translation; accordingly, all subsequent analyses of L2 performance will use the self-report measure (ReadEng) instead.

The demographic questionnaire contained four questions about the use of English,

namely use of English at work, use of English in private life, self-reported reading in English, and education in English. All four of these measures are correlated, with correlation coefficients ranging from .20 (education in English and reading in English) to .58 (use of English at work and reading in English). Therefore, a composite variable called Use of English was created by adding up the standard scores for these four variables.

Finally, we should note the correlations between age and some of the other variables. First, age at testing is very strongly correlated with age of arrival (r = .77), and moderately strongly correlated with length or residence (r = .50). Since age at testing is equal to age of arrival plus length of residence, such confounds are unavoidable and have also been found in other studies as well. In DeKeyser et al. 2010, for instance, the correlation between age at testing and age of arrival was .97 for study 1 and .98 for study 2. A second, and more interesting finding is that there is a different pattern of correlations between age and language scores in the two groups. In native speakers, both vocabulary and collocations scores are positively correlated with age (r=.37 and .27 respectively), and there is no significant relationship between age and grammar. This suggests that vocabulary and collocations, but not grammar (or at least not the relatively basic constructions studied here) continue to develop in adulthood. In L2 speakers, in contrast, we have marginally significant *negative* relationship between age and grammar (r=-.23, p =.06). These relationships will be explored further in the following section.

Regression analysis 1: The role of print exposure, language analytic ability, IQ, education and age

Ordinary least squares regression analyses were conducted in order to examine the possible effects print exposure, language analytic ability, nonverbal IQ, education and age on

the three linguistic measures (grammar, vocabulary and collocations). All predictors were centred before analysis. 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). The metric is obtained by averaging the sequential sum-of-squares obtained from all possible orderings of predictors, thus making it possible to estimate each regressor's unique contribution to the total variance in the dependent variable. Larson-Hall (2016: 255) argues that comparing `lmg` values for individual regressors is more meaningful than comparing standardized regression coefficients; however, I also provide standardized regression coefficients for the benefit or readers who are more familiar with this measure.

In cases where any regression assumptions were violated, the analysis was followed up with robust regression using the lmRob function from the `robust` package in R. Since the results were very similar, only the OLS results are reported here. The focus of the discussion here is on L2 performance; native speaker results are reported for comparison only. For a more in-depth discussion of the L1 results, see Dąbrowska (2018).

Grammar

As shown in Table 3, the best predictor of performance on the grammar test in adult native speakers is IQ, which accounts for 21% of the variance, with print exposure accounting for an additional 8%. It should be noted, however, that nonverbal IQ and language analytic ability are correlated, and hence it is difficult to tease apart their contribution. When IQ is excluded from the model, the only significant predictor of performance is language analysis, which accounts for 21% of the variance. Moreover, when only performance on the six most difficult structures is considered, we do get a small but significant effect of language analytic ability in addition to an effect of nonverbal IQ (see Dąbrowska 2018).

Table 3

Predictors of grammar scores in L1 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	86.1	1.1	0.0	76.6	<.001	
Blocks	1.0	0.2	0.5	5.1	<.001	.21
ART	0.3	0.1	0.3	3.3	.002	.08
Model R^2						.29

Table 4

Predictors of grammar scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	
Intercept	63.6	4.6	-0.1	13.9	<.001	
Age	-1.4	0.5	-0.3	-2.5	.013	
Model R^2						.09

In contrast, the only factor which is significant for the L2 speakers is age at testing. The lack of effect of IQ and in particular language aptitude in the non-native group is surprising, since there is a large body of research showing that language aptitude in particular is a reliable predictor of both foreign and second language attainment (Carrol & Sapon, 2002; Li, 2014; Pimsleur et al., 2004; Sasaki, 1999). However, all of these studies tested learners with very similar patterns of exposure (typically students following the same course). The lack of relationship observed here is most likely due to the fact that the effects of language aptitude are masked by other factors such as differences in exposure and instruction.

The significant negative effect of age on grammar in L2 speakers is likewise surprising. Although age-related decline in grammatical processing has been observed in

native speakers as well, the effects do not become noticeable until after age 60 (Silagi et al. 2015) and even then may be apparent primarily in reaction time as opposed to accuracy (Kim et al. 2014). The results reported here suggest that this process may begin much earlier in L2 speakers, possibly because of the fact that the second language is less well entrenched. We will return to this issue in the concluding section.

Vocabulary

As shown in Table 5, the best predictors of native speakers' performance on the vocabulary test were print exposure, nonverbal IQ, and age which accounted for 28%, 14% and 13% of the variance respectively. The results of the regression analysis for L2 speakers are summarized in Table 6. As in native speakers, the single best predictor of performance was reading in English, which accounted for 18% of the variance. In addition, language analytic ability accounted for further 14% of the variance.

Thus, the L2 results differ from those for native speakers in several ways. First, the effects of print exposure are somewhat lower in the L2 group. This could be due to the fact that the self-report questionnaire is less accurate than the objective measure used with native speakers. Alternatively, vocabulary development in adult L2 learners may depend less on mere exposure and more on explicit teaching and learning. The second difference is that L2 vocabulary size is not related to nonverbal IQ. This could be because L2 learners are, to a large extent, learning labels for concepts they already have. On the other hand, in contrast to native speakers, L2 speakers' vocabulary size is related to their language analytic abilities. While this makes sense (the ability to learn vocabulary requires working out which chunk of form corresponds to which chunk of meaning), it is not clear why such a relationship is not observed in native speakers. One possibility is that the effects of language analytic ability are masked by the stronger effects of nonverbal IQ, which, as explained earlier, is moderately

strongly associated with it. Finally, age has a significant effect on native speakers: older L1 speakers have larger vocabularies than younger ones, presumably because they have had more time to acquire additional words. Interestingly, however, there is no corresponding effect in L2 speakers.

Table 5

Regression analysis 1: Predictors of vocabulary scores in L1 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	69.1	1.5	0.0	47.0	<.001	
ART	.9	0.2	0.5	5.8	<.001	.28
Blocks	1.7	0.3	0.5	5.6	<.001	.14
Age	0.5	0.1	0.4	4.0	<.001	.13
Model R^2						.55

Table 6

Regression analysis 1: Predictors of vocabulary scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	52.0	2.3	0.0	22.4	<.001	
ReadEng	2.3	0.6	0.4	3.8	<.001	.18
LgAnalysis	1.9	0.6	0.3	3.3	.001	.14
Model R^2						.32

Collocations

The regression results for collocations are presented in Tables 7 and 8. As we can see,

in both groups print exposure was the only significant predictor of performance, accounting for 26% of the variance in native speakers and 25% of the variance in non-native speakers. This is not surprising: the Words that Go Together test is strongly biased towards collocations that appear in writing, and hence speakers who read more have more opportunities to learn a larger number of collocations.

Table 7

Regression analysis 1: Predictors of collocations scores in L1 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	66.3	1.8	0.0	39.9	<.001	
ART	0.9	0.2	0.5	5.6	<.001	
Model R^2						.26

Table 8

Regression analysis 1: Predictors of collocations scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	29.1	2.8	0.0	10.4	<.001	
ReadEng	3.5	0.7	0.5	4.7	<.001	
Model R^2						.25

Regression analysis 2: Additional predictors

In this section, I examine the role of three of the predictors considered in the first analysis (Blocks, Language Analysis, and Education) as well as four additional predictors relevant for L2 speakers only: age of first exposure, age of arrival, length of residence, and use of English. As in the previous analyses, all predictors were centred before running the

regression, and when any assumptions were violated, a robust regression was also undertaken. Since the results were very similar, only the OLS results are reported here.[2]

Grammar

The results of the second regression analysis for grammar are presented in Table 9. As we can see, there is only one significant predictor, age of arrival, which accounts for 8% of the variance. Use of English, when entered into the model by itself, is also a significant predictor which accounts for a somewhat smaller amount of variance – 6%. However, when both predictors are entered into the model, neither is significant. This is because they are fairly strongly correlated ($r = -.52$, $p < .001$) – in other words, to a large extent, they measure the same thing. I will discuss to the implications of this finding in the concluding section.

Table 9

Regression analysis 2: Predictors of grammar scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	R^2
Intercept	68.8	4.0	0.0	17.2	<.001	
AoA	-1.3	0.5	-0.3	-2.4	.019	.08
Model R^2						.08

Vocabulary

The results of the second regression analysis of vocabulary scores, presented in Table 10, are similar to those of the first analysis reported in the preceding section. The most important predictor is language analysis, which accounts for 14% of the variance, with use of English accounting for an additional 10%. It is interesting to note that use of English accounts for less of the variance in scores than the ReadEng measure used in the first analysis (which,

as explained earlier, is incorporated into this composite measure). This suggests that L2 learners, like first language learners, acquire most of their vocabulary via reading.

Table 10

Regression analysis 2: Predictors of vocabulary scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	52.3	2.4	0.0	21.4	<.001	
LgAnalysis	1.9	0.6	0.4	3.1	.003	.14
UseEng	2.2	0.8	0.3	2.6	.013	.10
Model R^2						.24

Collocations

The results of the second regression analysis of collocation scores are presented in Table 11. Because the homoscedasticity assumption has been violated, the reported standard errors, *t* values and *p* values have been adjusted using the Eicker-Huber-White method.

Table 11

Regression analysis 2: Predictors of collocations scores in L2 speakers

Variable	*B*	*SE B*	*β*	*t* value	Pr(>|t|)	lmg
Intercept	30.2	2.4	0.0	12.6	<.001	
UseEng	5.3	0.7	0.6	7.1	<.001	.37
LoR	1.1	0.4	0.2	3.0	.004	.08
Model R^2						.45

By far the best predictor of performance on the collocations test is use of English, which accounts for over 37% of the variance in scores. In addition, there is a small but significant effect of length of residence, which accounts for an additional 8% of the variance. Clearly the ability to recognise collocations is an aspect of linguistic knowledge which is strongly dependent on experience. This is not surprising. Knowledge about which words collocate with what is something that can only be learned from observing usage. In contrast to vocabulary, use of English is a better predictor of performance on the collocations test than reading in English, possibly because use of English is strongly related to a desire to integrate with the target language community, which could result in learners' paying more attention to linguistic form, including aspects of form that have relatively little to do with communicative efficiency. This possibility will be discussed more fully in the concluding section.

Summary and conclusions

The effects of native speaker status

The main aim of this study was to compare the performance of native speakers and adult L2 learners on tasks tapping grammatical comprehension, vocabulary size and knowledge of collocations. As expected, native speakers as a group obtained higher scores on all three tasks. However, there were vast individual differences in performance in L1 speakers as well as L2 speakers, and considerable overlap between the two groups. These findings contrast with those reported in many earlier studies of adult L2 learners, which found little or no overlap between native and non-native groups, at least for grammar (Abrahamsson & Hyltenstam, 2009; Coppieters, 1987; Granena & Long, 2013; Johnson & Newport, 1989). The different outcome is most likely attributable to two factors.

First, as noted in the introduction, most earlier studies used highly educated

participants, which tends to exaggerate the difference between L1 and L2 speakers. This is evident when we compare Figure 1, which shows the distribution of language scores for the entire sample, with Figure 2, which shows the distribution of scores in participants with 14 or more years of formal schooling – in other words, the type of participants who are usually recruited for L2 research. It is clear from comparing the two figures that when we restrict the sample to highly educated participants, the differences between L1 and L2 speakers become much more pronounced, *even when the two groups are matched for educational attainment.* This confirms the suspicion expressed in the introduction that research which examines highly educated participants only presents a biased picture of L2 attainment because it underestimates the amount of variation found in native speakers (see also Andringa, 2014; Dąbrowska, 2012). It follows that if L2 researchers want to generalize to the entire population, rather than just the population of university graduates, they must test more varied samples of participants.

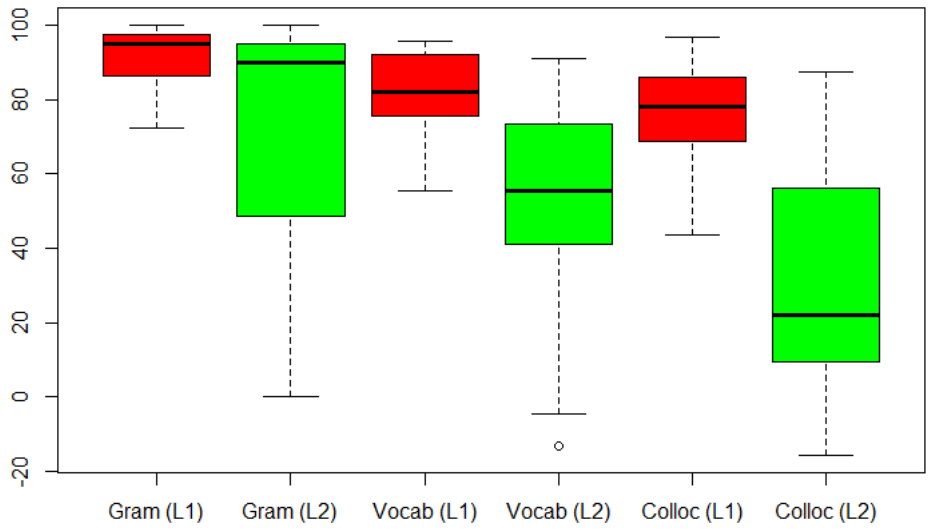

Figure 2: Distribution of scores on the language tests for speakers with 14 or more years of formal schooling (N=32 for the L1 group and 55 for L2 group)

There is another important difference between the present study and most earlier research comparing adult L2 learners with native speakers. As explained in the introduction, most earlier studies used a grammaticality judgment task rather than a comprehension task, and concentrated on certain structures, e.g. grammatical agreement, tense and plural marking, use of articles, and verb subcategorization patterns. It is possible – indeed, highly probable – that the differences between native speakers and adult L2 learners are particularly large in such structures, as they are selected precisely *because* they are known to be difficult for second language learners. This doesn't make the differences found in these studies any less real, of course: the point is that concentrating entirely on such structures gives us a distorted view of adult L2 learners' abilities.

More research is needed on which aspects of grammar are relatively easy or difficult for second language learners and why. One possibility consistent with the available data is that L2-difficult structures tend to be the more "decorative" as opposed to functional aspects of grammar, in the sense that their contribution to meaning is relatively small and often redundant: for instance, tense and number are often marked by adverbials and quantifiers as well as grammatical morphology. The acquisition of such aspects of grammar may depend strongly on the tendency to attend to and copy fine details of the behaviour of a model, including causally irrelevant aspects of behaviour. This tendency, which is referred to in the psychological literature as "overimitation", may be driven by a motivation to affiliate with social groups through participation in conventional activities (Legare & Nielsen, 2015), and is characteristic of human children (Lyons et al., 2007, 2011; Nielsen et al., 2014). Thus, children may be better at learning "decorative" grammar because they are more focussed on fitting in, while adult learners are more goal-directed, i.e., more focussed on getting the message across, and hence focus primarily on the more functional aspects of grammar.

Predictors of individual differences in language attainment: education, print exposure, language analytic ability, nonverbal IQ and age

The second aim of the study was to investigate possible reasons for differences in performance on language tasks in both native and non-native speakers. Five of the predictors examined here, viz., educational attainment, print exposure, language aptitude, nonverbal IQ and age are potentially relevant for both groups, while four additional predictors (age of first exposure, age of arrival, use of English, and length of residence) are relevant for L2 learners only.

Analyses of the relationship between the first set of predictors and performance on the language tasks revealed both similarities and differences between L1 and L2 speakers. As we have seen, education shows weak to moderately strong correlations with L1 performance (r = .36 for grammar, .37 for collocations, .43 for vocabulary), and a weak relationship with vocabulary in L2 speakers (r=.26). However, it did not emerge as a significant predictor of linguistic abilities in the regression analysis, suggesting that the correlation is attributable to other factors that are related to education, namely print exposure, reading, nonverbal IQ and language analytic ability.[3]

Reading in English contributes to variance in all three aspects of linguistic knowledge in native speakers, although the effect size for grammar is relatively small. For L2 speakers, reading predicted vocabulary size and collocational knowledge. The lack of effect on grammar, and somewhat smaller effect on vocabulary than in natives, may be due to the fact that these aspects of L2 knowledge are strongly dependent on the amount and type of language instruction; alternatively, it may simply reflect the fact that the self-report measure used for L2 speakers is less reliable than ART.

The findings reported here also revealed associations between nonverbal IQ and language aptitude (as measured by the Language Analysis subtest of the PLAB) on the one hand and grammar on the other in native speakers, but not in the L2 group. As pointed out earlier, the lack of effect for L2 speakers is most likely due to the fact that it was masked by other factors, in particular, language background and differences in amount and quality of instruction. Previous studies of the effects of aptitude on L2 learning tended to use much more homogenous populations of learners, often students from the same cohort who were of the same age and language background and had had similar amounts of language instruction. What is perhaps more surprising is the finding that language aptitude and nonverbal IQ were predictive of *native speakers'* performance on the grammar task (for a discussion of possible reasons for this, see Dąbrowska 2018).

Finally, an unexpected finding which emerged from the study was the different effect of age at testing in the two language groups. In L1 speakers, there was a positive relationship between age and two of the linguistic tasks, vocabulary and collocations, which suggests that lexical learning continues throughout adulthood. The lack of a relationship between age and grammar in native speakers could be due either to the fact that grammar reaches a steady state before adulthood or simply to the fact that the Pictures and Sentences test was not sensitive enough to pick up later changes (possibly because it targets fairly basic structures). In L2 speakers, however, there was no significant relationship between age and vocabulary size or age and knowledge of collocations, and a significant a *negative* relationship between age and grammar. Since age is also negatively correlated with use of English ($r = -.35$), this decline could be attributable to the different social context in which the older participants live. Alternatively, it is possible that age-related decline begins earlier in the second language, possibly due to its weaker entrenchment.

The role of age of acquisition and length and intensity of exposure

In addition to the factors discussed above, the study described here investigated the relationship between linguistic abilities and four factors which are relevant for L2 speakers only, namely age of first exposure, age of arrival, use of English and length of residence in an English-speaking country. Of the two age factors, only age of arrival was a significant predictor once other factors were controlled for. Furthermore, it was relevant only for grammar, and the effect size was very small (just 8% of the variance). Use of English was a strong predictor of performance on vocabulary and collocations. For grammar, it was not significant once age of arrival was controlled for; however, when it was entered into a regression model without age of arrival, its effects were almost as large as those of the latter factor. Interestingly, length of residence had no significant effect on grammar or vocabulary, and only a small effect on performance on the collocations task.

Several observations spring to mind in connection with these findings. In L2 research, age of arrival or age of first meaningful exposure are often used as a proxy for maturational changes in the brain. However, age is correlated with many other factors in addition to brain maturation – specifically, as we saw earlier, with age at testing and use of English – but also education in English and identification with the L2 community. Because of such correlations it is difficult to establish causation: it is possible that age is such a good predictor of L2 attainment simply because it correlates with so many different things which influence learning. (See Flege, 2009 for a more in-depth discussion of this issue.)

A second observation concerns the effects, or rather non-effects, of length of residence. LoR is the most widely used measure of experience in L2 ultimate attainment research. The results reported here show that, for this population at least – that is to say, speakers with an age of arrival of 16 or above, and length of residence of 3 years or more –

LoR is unrelated to use of English and a very poor predictor of L2 performance. This is in line with the results of several other studies. For example, Huang (2014) examined the effects of length of residence (as well as a number of other predictors) on L2 grammar and phonology and found that it had a *negative* effect on attainment. This was most likely due to confounding factors: LoR was positively associated with age, which was negatively associated with performance on language tasks. Flege and Liu (2001) compared L1 Chinese students and non-students learning English in the United States. There was a positive relationship between LoR and linguistic attainment in students, but no relationship (in fact, a trend in the opposite direction) in non-students. Interestingly, the two groups did not differ in amount of L2 use: the differences, the authors argue, are attributable to the fact that the students were exposed to more native input. All this suggests that length of residence is simply not a useful measure of exposure, and consequently, studies which pitch LoR against AoA are simply not very informative.

Limitations and suggestions for future research

The study described in this paper is relatively small-scale, and the results will need to be replicated using a larger sample of participants. It is also important to examine a wider range of predictors than it was possible to study here, the most obvious ones being working memory, phonological short-term memory and implicit learning abilities, all of which have been claimed to be associated with language learning. Finally, while the results reported here do not support the existence of critical period effects for second language learning, it does not follow that such effects do not exist. It is perfectly possible that CP effects exist for other aspects of language not studied here – notably phonology and “decorative” grammar. The important point is that, apart from phonology, the range of constructions that are subject to

strong age effects may be quite marginal – and yet the field of second language acquisition has been fixated on the CPH debate. It is time to overcome this fixation. There may well be structures for which there is little or no variation in natives and strong AoA effects in L2 speakers. If so, it is important to identify and study such structures. However, it is also important to study structures for which there is variability in both groups – and studying these structures, I submit, will yield more insights into the nature of human linguistic capacities.

Endnotes

[1] The only exception to this was Brooks & Sekerina (2006), who focused on comparing monolingual and bilingual speakers rather than demonstrating the existence of individual differences. However, the results of this study were replicated by Street & Dąbrowska (2010) who employed a variety of control measures.

[2] Given the fact that L2 speakers' language scores are negatively correlated with age, it would be interesting to explore this factor as well in the same model. However, this is not possible due to multicollinearity: as can be seen from Table S3, age at testing is strongly associated with both age or arrival and length of residence.

[3] Note, however, the study described in Dąbrowska (2018), which used a different measure of grammatical abilities, did find a small but significant effect of education in the same group of L1 speakers of English. In the earlier study, the dependent variable was the proportion of correct responses (adjusted for guessing) in the experimental conditions, i.e. Passives, Postmodified Subjects, Object Clefts, Object Relatives and the two quantifier constructions. The remaining four sentence types (i.e., Actives, Locatives, Subject Relatives and Subject Clefts) were treated as control conditions, where performance was predicted to be (and indeed was) at ceiling. However, since the L2 speakers in the current study made errors also

on these basic structure, the scores for all sentences were included in the analyses reported here.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*, 249–306. doi: 10.1111/j.1467-9922.2009.00507.x

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. doi: 10.3758/BRM.40.1.27

Andringa, S. (2014). The use of native speaker norms in critical period hypothesis research. *Studies in Second Language Acquisition, 36*, 565–596. doi: 10.3758/BRM.40.1.27

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.

Birdsong, David. (2006). Age and second language acquisition and processing. A selective overview. *Language Learning, 56*, 9–46. doi: 10.1111/j.1467-9922.2006.00353.x

Bishop, D. V. M. (2003). *Test for Reception of Grammar. Version 2. TROG-2 Manual.* London: The Psychological Corporation Ltd.

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition, 31*, 175–198. doi: 10.1017/S0272263109090275

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. doi: 10.1075/lllt.35.08bol

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

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

Carroll, J. B., & Sapon, S. M. (2002). *Modern Language Aptitude Test: Manual 2002 Edition*. N. Bethesda, MD: Second Language Testing, Inc.

Chipere, N. (2003). *Understanding Complex Sentences: Native Speaker Variations in Syntactic Competence*. Basingstoke: Palgrave.

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

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

Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language*, *63*, 544–573. doi: 10.2307/415005

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

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*, 219–253. doi: 10.1075/lab.2.3.01dab

Dąbrowska, E. (2014a). Words that go together: Measuring individual differences in native speakers' knowledge of collocations. *The Mental Lexicon, 9*, 401–418. doi: 10.1075/ml.9.3.02dab

Dąbrowska, E. (2014b). Implicit lexical knowledge. *Linguistics*, *52*(1), 205–223. doi: 10.1515/ling-2013-0060

Dąbrowska, E. (2015). Individual differences in grammatical knowledge. In E. Dąbrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 649–667). Berlin: De Gruyter Mouton.

Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition, 178*, 222–235. doi: 10.1016/j.cognition.2018.05.018

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. doi:10.1016/j.langsci.2005.11.014

DeKeyser, R. (2012). Individual differences in native language attainment and their implications for research on second language acquisition. *Linguistic Approaches to Bilingualism*, *2*, 260–263. doi: 10.1075/lab.2.3.03dek

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, *79*, 67–89. doi: 10.1111/j.1540-4781.1995.tb05417.x

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (2012). Individual differences in sentence processing. In M. Spivey, K. McRae, & M. Joannisse (Eds.), *Cambridge Handbook of Psycholinguistics*. Cambridge: Cambridge University Press.

Flege, J. E. (2009). Give input a chance. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 175–190). Bristol: Multilingual Matters.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*, 527–552.

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. doi: 10.1177/0267658312461497

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, *17*(1). Retrieved from http://www.jstatsoft.org/v17/i01

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, *61*, 139–147. doi: 10.1198/000313007x188252

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the Critical Period Hypothesis for Second Language Acquisition. *Psychological Science*, *14*, 31–38. doi: 10.1111/1467-9280.01415

Hoffmann, S., & Lehmann, H. M. (2000). Collocational evidence from the British National Corpus. In J. M. Kirk (Ed.), *Corpora Galore. Analyses and Techniques in Describing English* (pp. 17–32). Amsterdam: Rodopi.

Huang, B. H. (2014). The effects of age on second language grammar and speech production. *Journal of Psycholinguistic Research*, *43*, 397–420. doi: 10.1007/s10936-013-9261-7

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. doi: 10.1006/pmed.1998.0301

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Karanth, P., & Suchitra, M. G. (1993). Literacy acquisition and grammaticality judgments in children. In R. J. Scholes (Ed.), *Literacy and Language Analysis* (pp. 143–156). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Kim, H., Sung, J. E., Kim, J., Park, E. J., Yoon, J. H., Yoo, H. J., & Kim, S. R. (2014). Response time measurements in the iOS app -based Token Test in the elderly. *Geriatrics & Gerontology International*, *14*(4), 969–974. doi: 10.1111/ggi.12182

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

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R* (2nd ed.). Routledge.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672. doi: 10.1111/j.1467-9922.2010.00621.x

Legare, C., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. doi: 10.1016/j.tics.2015.08.005

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. doi: 10.1093/applin/amu054

Lidz, J., & Williams, A. (2009). Constructions on holiday. *Cognitive Linguistics*, *20*, 177–189. doi: 10.1515/COGL.2009.011

Long, M. (2013). Maturational constraints of child and adult SLA. In G. Granena & M. Long (Eds.), *Sensitive Periods, Language Aptitude and Ultimate L2 attainment* (pp. 3–41). Amsterdam/Philadelphia: John Benjamins.

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil., F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1158–1167. doi: 10.1098/rstb.2010.0335

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences, 104*(50), 19751–19756. doi: 10.1073/pnas.0704452104

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8- and 12-year-old children and adults. *Journal of Experimental Psychology: General, 144*, 447–468. doi: 10.1037/xge0000054

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.

Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2014). Where culture takes hold: "Overimitation" and its flexible deployment in Western, Aboriginal, and Bushmen children. *Child Development, 85*, 2169–2184. doi: 10.1111/cdev.12265

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. doi: 10.1038/42347

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

Sasaki, M. (1999). *Second Language Proficiency, Foreign Language Aptitude, and Intelligence*. New York: Peter Lang.

Silagi, M. L., Rabelo, C. M., Schochat, E., & Mansur, L. L. (2015). Healthy aging and compensation of sentence comprehension auditory deficits. *BioMed Research International*, Article ID 640657. doi: 10.1155/2015/640657

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

Spadaro, K. (2013). Maturational constraints on lexical acquisition in a second language. In G. Granena & M. Long (Eds.), *Sensitive Periods, Language Aptitude and Ultimate L2*

attainment (pp. 43–68). Amsterdam/Philadelphia: John Benjamins.

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2011). Subcomponents of second□ language aptitude and second□language proficiency. *The Modern Language Journal, 95*, 253–273. doi: 10.1111/j.1540-4781.2011.01176.x

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.

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. doi: 10.1016/j.langsci.2016.10.004

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. doi: 10.1016/j.lingua.2010.01.004

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. doi: 10.1017/S0142716412000367

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. doi: 10.1177/0142723706056475

Tarone, E., Bigelow, M., & Hansen, K. (2013). *Literacy and Second Language Oracy*. Oxford University Press.

Ullman, M. T. (2006). The declarative/procedural model and the shallow structure hypothesis. *Applied Psycholinguistics, 27*, 97–105. doi: 10.1017/S014271640606019X

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. doi: 10.1016/j.cogpsych.2008.08.002