Personality Traits Moderate the Relationship between Statistical Learning Ability and Second-Language Learners' Sentence Comprehension

Elma Kerz (elma.kerz@ifaar.rwt-aachen.de)
RWTH Aachen University, Department of English Linguistics,
Kármánstr 17/19, 52062 Aachen, Germany

Daniel Wiechmann (d.wiechmann@uva.nl)
University of Amsterdam, Institute for Logic, Language and Computation,
Science Park 107, 1098 XT Amsterdam, The Netherlands

Tandis Silkens (tandis.silkens@rwth-aachen.de)
RWTH Aachen University, Department of English Linguistics,
Kármánstr 17/19, 52062 Aachen, Germany

Abstract

An accumulating body of evidence has demonstrated a tight coupling between individual differences (ID) in statistical learning ability (SL) and variation in language performance in child and adult native speaker populations, with some initial evidence that this coupling extends to second language (L2) speakers. However, surprisingly little work has been conducted to assess potential interactions between SL and other experience-related and affective ID factors. Using a within-subjects design embedded in an ID framework, the present study attempts to fill this gap by investigating whether the impact of SL ability on language is moderated by individual differences in personality traits and the amount of experience an individual has had with the L2. The results of the study revealed a complex interplay between ID factors and variation in L2 comprehension of different types of complex sentences indicating that the effect of SL ability on language comprehension is moderated by personality traits.

Keywords: language comprehension; second language learning and processing; statistical learning; personality

Introduction

There is now a growing awareness that at every stage across the lifespan there are substantial individual differences (IDs) in language competencies. IDs are ubiquitous across the entire linguistic system, from the acquisition of properties of speech as well as the graphotactic regularities of written words to complex grammatical structures and discourse (see, e.g., Kidd et al., 2018, for a recent review). Understanding the determinants and consequences of individual variation in language outcomes is of paramount importance for informing theories of the acquisition, processing and use of language. Statistical learning (SL) – succinctly defined as the ability to automatically detect and extract distributional properties in the input – has been implicated as a strong candidate, since it systematically varies within the population and is found to correlate with language performance in both child and adult populations: An individual's capacity for SL is found to predict, among other things, the comprehension of various types of grammatically complex sentences (Kidd & Arciuli, 2016; Misyak & Christiansen, 2012), semantic and phonological lexical access (Mainela-Arnold & Evans, 2014), and speech perception (Conway et al., 2010) (for a detailed overview of studies, see, Siegelman, Bogaerts, Christiansen, & Frost, 2017). Most relevant to the current study, Misyak and Christiansen (2012) have shown that performance on artificial grammar learning (AGL) tasks that tested adult native speakers' ability to learn adjacent and nonadjacent dependencies successfully predicted the comprehension of three types of complex sentence structures involving these dependency types over and above the influence of other cognitive IDs factors, including verbal working memory and fluid intelligence. This study has in an exemplary way taken into account other cognitive IDs factors next to SL ability as well as experience-based ones. However, the study did not test for potential interaction effects between SL and other IDs factors investigated in the study. More importantly, like all other studies on the SL–language relationship, it did not consider another potentially significant IDs factor, namely personality traits. This affective IDs factor has received a considerable amount of attention in the language learning literature, in particular in the area of second language learning (Dewaele, 2012). First evidence that statistical learning ability may be impacted by personality comes from an extensive study by Kaufman et al. (2010). In this study participants were administered a battery of tasks assessing the association of implicit SL with a variety of cognitive and personality variables. Implicit SL was more weakly related to psychometric intelligence than was explicit associative learning, and was unrelated to working memory capacity. The findings of structural equation modeling reported in the study revealed that implicit SL was independently related to two components of psychometric intelligence (verbal analogical reasoning and processing speed) and was also independently related to academic performance on two second language exams. Importantly, implicit SL was significantly associated with aspects of self-reported personality, including intuition, Openness to Experience, and impulsivity. The results reported in Kaufman et al. (2010) call for further studies directed at elucidating the potential impact of personality traits on SL ability as well as their relative contribution to language abilities. Compared to the available research reviewed above on the relationship of SL and language in native speakers, attempts to investigate the role of SL in adult second language learning and processing have

been scarce ((but see Frost et al., 2013, for exceptions) and Ettlinger et al. (2016). The results of these studies provide some initial support suggesting that SL is also involved in second language learning in adults.

The present study seeks to contribute to this body of research and advance our understanding of the role of statistical learning in language and its interactions with other cognitive, experience-related and affective IDs factors. More specifically, through a within-subjects design embedded in an IDs framework, the study is aimed at investigating whether the impact of SL ability on comprehension of different types of complex sentence structures is moderated by individual differences in personality traits and the amount of language experience. In pursuit of this aim, the study sets out (1) to replicate the effects of SL abilities on language comprehension in adult native speakers reported in Misyak & Christiansen (2012) in a sample of adult second language learners and (2) to assess potential interactions between the SL abilities, the amount of experience with the L2 and five personality traits. To this end, we analyzed data from a sample of adult German speaking L2 learners of English on a total of thirteen measures assessing individual differences in the four constructs of interest. We first present the results of a correlational analysis investigating the bivariate relationships between all individual differences measures. We then report on the results from mixed-effects regression models that investigated the joint effects of statistical learning ability, personality traits and the amount of language experience on L2 comprehension accuracy of three types of sentences structures.

Methods

Participants

Fifty native German-speaking L2 learners of English (34 female and 16 male, $M = 22.36$ years, $SD = 2.86$) participated in this study. All participants were students at RWTH Aachen University and classified as having a Common European Framework (CEFR) English proficiency level of at least B2 (upper intermediate) or C1 (lower advanced) based on their institutional status. In addition, participants were asked to fill out the Language Experience and Proficiency Questionnaire to obtain general demographic information and more specific information on self-rated proficiency and current knowledge of L2 English and exposure to the L2 (Kaushanskaya et al., 2019). Regarding their English acquisition, the L2 speakers started learning English around the age of 9 and reported to have acquired fluency at around 15 years of age. On average, their current experience with English comes mainly from listening to music (mean score of 7.08 out of 10), social media (mean score of 6.84 out of 10) and reading (mean score of 5.49 out of 10). Self-ratings of their English language proficiency based on a 10-point scale were relatively high (all mean scores greater 7.16).

Materials

Statistical Learning Ability Individual differences in statistical learning ability were assessed using the two auditory-verbal artificial grammar learning (AGL) tasks used in Siegelman & Frost (2015). These tasks were designed to probe into participants' ability to extract statistical regularities between adjacent and non-adjacent elements. Both tasks involved two phases: In an initial 'familiarization' phase, participants were auditorily presented with a 9-10 minute stream of trisyllabic 'words' from an artificial language. In a subsequent testing phase, participants were presented with 36 pairs of trisyllabic sequences each consisting of one item from the familiarization stream (a 'word') and one previously unseen variation of a word (a 'part-word') and had to decide which of the two items sounded more familiar to them. All stimuli were synthesized in PRAAT software (Boersma, 2001), at a fundamental frequency of 76 Hz, with a mean syllable duration of 290 ms. The stimulus material in the adjacent AGL task was based on 12 'words', built from 18 CV syllables. Items in the familiarisation stream were designed such that the transitional probabilities (TPs) between every two adjacent syllables within a word was 0.5 whereas the TP across words boundaries was only 0.187 on average. The detection of word boundaries in the familiarization stream was possible only in terms of this TP contrast. Performance on the task was assessed in terms of the total number of correct word identifications in the testing phase (0 - 36).

The stimulus material of the non-adjacent AGL included 12 trisyllabic artificial 'words' that were built from three groups of consonantal patterns ($p\_v\_g\_$, $d\_k\_b\_$ and $m\_t\_s\_$) into which four different vowel combinations were inserted. In the testing phase, participant had to decide between 'legal words', i.e. items that were constructed from the consonantal patterns from which the words in the familiarization stream with novel vowels, and 'non-legal words', i.e. items that contained only one or two consonants from a given consonantal pattern and one or two from another pattern. To ensure that the only cue for extracting the underlying patterns of words in the familiarization stream was the non-adjacent TPs, the TPs between adjacent syllables was held constant, both between and within words. Scores on the task ranged from 0 to 36, based on the number of correct identifications of legal words.

Personality Personality traits were assessed through the Big Five Inventory (BFI) questionnaire that measures an individual on the Big Five Factors of personality: extraversion, agreeableness, conscientiousness, neuroticism and openness (John et al., 2008). The BFI consists of 44 self-rating statements, such as "I see myself as someone who generates a lot of enthusiasm". Participants were asked to rate each statement on a 5-point Likert scale ranging from 1 ('disagree strongly') to 5 ('strongly agree'). Participants' scale scores for each of the five dimensions were expressed as person-centered z-scores that were adjusted for differences in acquiescent response styles. Individuals with a high extraversion factor are expected to have a dynamic and active manner to-

ward the social and material world, subsuming traits such as sociability, activity, assertiveness, and positive emotionality. The agreeableness factor contrasts a prosocial and communal orientation towards others with antagonism and covers traits such as altruism, tender-mindedness, trust, and modesty. Individuals with a high conscientiousness factor are commonly described as having a socially prescribed impulse control facilitating task- and goal-directed behavior. Individuals with a high neuroticism factor are prone to psychological distress and more likely than average to be moody and to experience such feelings as anxiety, worry, and frustration. Individuals with high scores on Openness-to-Experience are more curious, creative, original, imaginative, and untraditional, and have broad interests.

Amount of L2 experience Following previous work, we quantified individual participants' linguistic experience using three measures that provide proxy estimates of an individual's exposure to the L2: [1] the Author Recognition Test (ART), [2] the Need for Cognition (NFC) and [3] the Lexical Test for Advanced Learners of English (LexTALE). An updated version of the ART from West et al. (1993) was used a measure of print exposure. Numerous studies have established that the ART correlates significantly with measures of orthographic processing, spelling, vocabulary size, reading comprehension and general knowledge, indicating that it is a valid measure of print exposure (Moore & Gordon, 2015). Participants were presented with a list of 93 names and asked to mark those that they believe are names of real authors. The list comprises 46 names of real authors and 47 foils, i.e. names not associated with a published author. The ART was scored as the number of real author names identified by a participant minus the number of foil names selected. The LexTALE test is an unspeeded visual lexical decision task designed to assesses the receptive vocabulary knowledge, which is often argued to be a strong indicator of the amount of time an individual spends reading. Participants performed the online version of the task consisting of 60 vocabulary trials on which they have to decide whether the presented item exists. The test consists of 40 words and 20 nonwords. Performance in the task was assessed in terms of the percentage of correct responses, corrected for the unequal proportion of words and nonwords in the test. The NFC scale (Cacioppo et al., 1984) measures a personality-based variable that concerns the degree to which an individual prefers cognitively engaging activities including such reading over activities that require less cognitive engagement. NFC has been used as a motivational proxy to linguistic experience based on the assumption that individuals with higher NFC will be more likely to engage with printed materials and thus to possess a higher degree of print exposure (Farmer et al., 2017). The NFC scale includes items such as "I would prefer complex to simple problems". Each item was answered on a 5-point Likert-type scale ranging from 1 (extremely uncharacteristic of me) to 5 (extremely characteristic of me). An 18-item version of the Need for Cognition (NCS) scale was used and the NFC scores was determined by summing responses to all items.

L2 Sentence Comprehension Abilities Individual variation in L2 sentence comprehension abilities was assessed using the same task as in Misyak & Christiansen's (2010) L1 study. The stimulus material was carefully chosen so as to involve the tracking of different types of both adjacent (local) and non-adjacent (long-distance) relationships: It consisted of structurally complex sentences drawn from three prior studies. The S/OR sentence set contained subject relative clauses and object relative clauses. Processing of such structures requires tracking both non-adjacent (nonlocal, long-distance) and adjacent relationships between language units. The A/IN sentence set comprised animate and inanimate head nouns modified by a reduced or non-reduced relative clause. This set was designed to investigate whether and to what extent the resolution of local syntactic ambiguities was affected by semantic constraints, specifically the animacy of the head noun referent. The processing of such sentence types also involves keeping track of both local (adjacent) dependencies - i.e. between the head noun and its modifying RC - as well as and long-distance (non-adjacent) dependencies holding across the RC between the head noun and the main verb. The 'PT' set involved noun/verb homonyms with phonologically typical or atypical noun/verb resolutions. Their processing required tracking of local dependencies between the sentence's ambiguous homonym and the subsequent words that resolve the ambiguity. Four sentence lists were created, each incorporating 12 initial practice items followed by 88 test items (40 sentences with subject-object relative clauses (S/OR), 28 sentences involving clauses with animate/inanimate noun constructions (A/IN) and 20 sentences involving noun/verb homonyms with phonological typical or atypical noun/verb resolutions (PT). Versions of a given target sentence were counterbalanced across the four lists and presented in random order. A YES/NO comprehension question followed every sentence. Following Misyak & Christiansen (2012), performance on the task was assessed in terms of comprehension accuracy scores.

Procedure

The tests and measures were administered in the same order to all participants to avoid potential effects of administration ordering (LexTALE, BFI, NFC, ART, Adjacent SL, Sentence Comprehension, Non-adjacent SL). All experimental tasks were implemented and run using PsychoPy v3.0 (Peirce, 2007).

Results

The means, standard deviations, and range for all individual differences (IDs) measures are provided in Table 1 and the correlations among the measures are presented in Table 2. Accuracy scores on the L2 comprehension task were subject to considerable individual differences ranging between 59.04% and 97.59%. Like the results reported for native speakers in Misyak and Christiansen (2012), the performance

Table 1: Descriptive statistics for the individual-differences tasks and measures assessing (1) L2 sentence comprehension, (2) statistical learning, (3) personality and (4) amount of L2 experience.

	Dependent measure	M	SD	Observed range
Animate/inanimate (A/IA)	Percent correct (28 Y/N questions)	79.25	10.84	58.82–96.30
Phonological Typicality (PT)	Percent correct (20 Y/N questions)	87.12	10.72	60.00–100.00
Subject/object Relatives (S/OR)	Percent correct (40 Y/N questions)	75.2	12.93	48.28–97.44
Adjacent	Percent correct (36 2AFC items)	52.6	12.73	30.55–80.55
Nonadjacent	Percent correct (36 2AFC items)	56.97	9.48	36.11–83.33
Openness	Person-centered z-scores adjusted	0.12	0.46	-1.50–1.08
Conscientiousness	for differences in acquiescent	0.17	0.47	-0.78–1.39
Extraversion	response styles	0.56	0.49	-0.53–1.44
Agreeableness	('yea-saying' vs. 'nay-saying')	0.59	0.43	-0.88–1.37
Neuroticism		-0.35	0.59	-1.55–1.33
LexTALE	LexTALE score (averaged % correct)	71.607	12.489	47.50–98.75
ART	Sum of correct targets - foils	61.265	12.3	20.00–86.00
NFC	Sum of scaled responses	6.5714	4.8088	0.00–21.00

Table 2: Intercorrelations between tasks and measures.

	Statistical Learning		L2 Comprehension				L2 Experience				Personality		
	Adjacent	Non.-adj	A.IN	PT	S/OR	OVERALL	LexTALE	ART	NFC	OS	CS	ES	AS
Non.-adj	-0.21												
A.IN	-0.01	-0.08											
PT	0.09	-0.17	0.58***										
S/OR	0.01	-0.08	0.51***	0.54***									
OVERALL	0.02	-0.12	0.8***	0.76***	0.9***								
LEXTALE	-0.13	-0.02	0.62***	0.54***	0.38**	0.57***							
ART	0.04	-0.22	0.41**	0.37***	0.43**	0.49***	0.4**						
NFC	-0.02	-0.05	0.24	0.47***	0.3*	0.38**	0.45**	0.32*					
OS	0.13	-0.11	0.21	0.12	0.09	0.15	0.13	0	0.24				
CS	0.22	-0.11	0.08	0.28	0.1	0.14	0.01	0.08	0.23	0.08			
ES	0.06	-0.01	-0.27	-0.07	-0.14	-0.2	-0.29*	-0.39**	0.03	0.16	0.51***		
AS	-0.08	0.15	-0.31*	-0.32*	-0.41*	-0.43*	-0.4**	-0.47***	-0.08	0.24	0.27	0.41**	
NS	-0.03	0.13	-0.09	-0.13	0.06	-0.03	0.01	0.05	0.05	-0.01	-0.11	0.11	0.13

of the L2 group on sentence comprehension was highest on the PT set, followed by the A/IN set, and lowest on the S/OR set. However, the performance ranged between 7.28% and 10.85% below that of the native speaker group tested in the L1 study. Performance on the non-adjacent AGL was above chance ($t(49) = 5.4, p < 0.0001$) and similar in performance-level to the L2 group tested in Siegelmann and Frost (2015). On the adjacent AGL, overall group performance was not significantly different from chance ($t(49) = 1.4, p = 0.16$) and below the performance-level reported in Siegelmann and Frost (2015). L2 comprehension scores on the three sentence sets were significantly correlated with each other and - with one exception - also with all L2 experience proxy measures. Across sentence sets, there also were significant negative correlations between comprehension accuracy and scores on the agreeableness personality dimension. The three L2 experience proxies were all significantly correlated with one another. Furthermore, ART and LexTALE scores were negatively correlated with scores on the Extraversion and Agreeableness scales. Within the personality traits, there were positive correlations between Extraversion and Agreeableness and between Extraversion and Conscientiousness. Regarding statistical learning, there were no correlations neither between the two SL measures nor between a SL measure and any of the language- and personality-related IDs factors. Crucially, there were no significant correlations between performance on either of the two SL tasks and L2 comprehension scores on any of the sentence sets.

In a second step, we investigated the joint effects of statistical learning, personality and L2 experience on L2 comprehension accuracy using mixed-effects regression modeling. Separate models were fitted to the data from each of the three language comprehension sets. All models had the maximal random-effects structure justified by our design, including random intercepts for subjects and items. For each set, we first fitted a full model in which sentence comprehension accuracy scores were regressed onto the main effects of both statistical learning predictors, the five personality predictors, the three experience-related variables, as well as all two-way interactions between, on the one hand, the SL pre-

dictors and, on the other hand, the personality and experience variables. To reduce multicollinearity, all variables were standardized prior to being entered into the model. Since our main goal was to determine whether the effects of SL ability are moderated by personality and/or L2 experience, we did not test for interactions within and across the personality- and experience-based IDs measures. We then employed a stepwise bidirectional variable selection procedure based on AIC to obtain the most parsimonious (minimal adequate) model. Only variables that decreased the AIC were retained. All models were fitted using the lme4 package (Bates et al., 2014) in R (R Core Team, 2019). The results of the final models for each language comprehension set are summarized in Table 3. For the PT set, the minimal adequate model accounted for 34% of the variance. The only two significant IDs variables retained in the model were the two L2 experience variables, LexTALE ($\hat{\beta} = 0.034, SE = 0.01, t = 3.1, p = 0.002$) and NFC ($\hat{\beta} = 0.021, SE = 0.01, t = 2.17, p = 0.03$). None of the other IDs measures had a significant effect on the response. For the A/IN set, the minimal adequate model accounted for 33% of the variance. It revealed a significant main effect of LexTALE ($\hat{\beta} = 0.02, SE < 0.01, t = 3.763, p < 0.001$). There were no main effects of any of the remaining IDs measures. However, there was a significant interaction between performance on the non-adjacent SL task and the Neuroticism personality trait ($\hat{\beta} = 0.08, SE = 0.03, t = 2.74, p = 0.01$), indicating a positive effect of SL ability in individuals scoring higher on that scale. Finally, for the S/OR set, the minimal adequate model accounted for 35% of the variance. In this set, there were negative main effects of Extraversion ($\hat{\beta} = -0.48, SE = 0.23, t = -1.98, p = 0.047$) and Agreeableness ($\hat{\beta} = -1.01, SE = 0.27, t = -3.77, p < 0.001$) as well as a positive main effect of Conscientiousness ($\hat{\beta} = 0.645, SE = 3.45, t = 2.53, p = 0.01$). In addition, a significant interaction between adjacent SL ability and Extraversion ($\hat{\beta} = 3.45, SE = 1.54, t = 2.23, p = 0.02$) indicated that SL abilities had differential effects in extroverts and introverts, with individuals on the higher end of the Extraversion scale benefitting more from higher SL abilities. The reverse effect was found for the Openness trait, where a stronger effect of SL ability was found in individuals with lower score on the personality scale ($\hat{\beta} = -6.145, SE = 2.15, t = -2.854, p = 0.004$).

Table 3: Regression coefficients (with 95% confidence intervals) from minimal adequate mixed-effects models fitted to the L2 comprehension accuracy data of the PT set (top), A/IA set (middle) and S/OR set (bottom).

	PT (adjacent)
Intercept	2.194^{***}
NFC	0.021^{*} (0.002, 0.040)
LexTALE	0.034^{**} (0.013, 0.056)
	A.IN (adj and non-adj)
Intercept	1.395^{***} (1.103, 1.688)
SL.non.adj	-1.197 (-2.886, 0.492)
ES	-0.322 (-0.667, 0.023)
NS	0.002 (-0.259, 0.263)
OS	0.323 (-0.051, 0.697)
LexTALE	0.026^{***} (0.012, 0.039)
SL.non.adj:NS	3.793^{*} (0.799, 6.787)
	S/OR (adj and non-adj)
Intercept	1.352^{***} (1.048, 1.655)
SL.adj	-0.166 (-1.720, 1.387)
ES	-0.476^{*} (-0.946, -0.006)
AS	-1.010^{***} (-1.537, -0.483)
CS	0.645^{*} (0.147, 1.144)
OS	0.175 (-0.290, 0.641)
SL.adj:ES	3.454^{*} (0.425, 6.483)
SL.adj:OS	-6.145^{**} (-10.365, -1.925)

Discussion

The present study aimed to determine whether and to what extent individual variability in language comprehension in adult second language learners can be accounted for by individual differences (IDs) in their statistical learning (SL) abilities, their personality traits and the amount of experience they have had with their L2. Through a within-subjects design embedded in an IDs framework, advanced L2 learners of English were administered two widely used tasks in the field of SL research that involved tracking of adjacent and non-adjacent statistical regularities in artificial grammar learning (AGL) tasks, along with three tasks designed to assess IDs in the amount of print exposure and the BFI questionnaire determining their personality scores. The participants' scores on IDs measures were then used as predictors of their performance on comprehension tasks across three sets of sentence materials entailing the tracking of adjacent and non-adjacent dependencies in natural languages originally used in the study on native language comprehension by Misyak and Christiansen (2012). As reviewed in the Introduction, a growing body of recent research demonstrates a tight coupling between SL ability and variability in native language learning and processing. There is also initial evidence that SL is implicated in second language learning and processing. Contrary to our expectations, there were no significant associations between performance in the two AGL tasks and L2 comprehension of relevant types of natural language sentences. However, several significant interactions indicated that the effect of statistical learning on language comprehension is moderated by personality traits, such that SL abilities had differential effects on introverts and extraverts, as well on individuals that are high or low on Openness and Neuroticism. These effects persisted even after controlling for the effects of L2 experience. Our findings also indicated a distinct role of language experience in language comprehension performance across sentence types, as evinced by several positive correlations between the three proxy measures of linguistic experience and L2 comprehension scores across the three sentence types.

Taken together, these findings are consistent with 'emer-

gentist' approaches to language, an umbrella term encompassing a broad class of approaches including usage-based approaches, constraint-based approaches, exemplar-based models and connectionist models that share the same core assumptions. According to these approaches, language (both L1 and L2) is learnt via analyses of, and generalization from, the input and IDs result from a complex interplay of endogenous cognitive systems and exogenous factors, in particular the quantity and quality of the language input (Kidd et al, 2017). These approaches call for an integrated approach to understanding the role of IDs in language learning and processing. The present study emphasize that such an integrated approach needs to take into account personality traits, one of the key affective factor in the educational psychology and second language learning literatures. To our knowledge, the present study is the first to show that personality traits impact the contribution of statistical learning to language abilities. At the methodological level, the results of the present study indicate that SL effects depends on the type of language structure to be comprehended: In this study, L2 comprehension of sentences involving clauses with animate/inanimate noun constructions was affected by the SL ability to detect non-adjacent dependencies, while L2 comprehension of sentences with subject-object relative clauses was impacted by the SL ability for adjacent dependencies. Such dependency of SL effect on sentence structure was also reported in Misyak and Christiansen (2012), albeit with a different pattern of associated tasks. This finding that the pattern of results varied across construction types draws attention to the fact that the study of individual difference requires not only the inclusion of multiple constructs and multiple measures per construct, but also needs to incorporate multiple measures to assess language performance in a particular domain (cf. James et al., 2018). The decision to assess SL abilities on the basis of AGL measures was motivated by our aim to examine whether the SL effects reported in Misyak and Christiansen (2012) can be replicated in a sample of adult L2 learners. The original study made use of carefully chosen stimulus materials to match the requirements of the artificial grammar learning and language comprehension tasks in terms of the types of dependencies between the elements that need to be tracked. However, the AGL tasks employed here have been designed to assesses group-level SL performance and have some psychometric shortcomings. Crucially, in AGL tasks a large proportion of the sample typically performs at chance level, as was the case in the adjacent SL task of the present study. To remedy these shortcomings, recent research has introduced novel measures to assess individual differences in SL abilities (Siegelman, Bogaerts, & Frost, 2017), including chunk-based measures (Isbilen et al., 2017). Such measures need to be integrated in future work.

Acknowledgments

This research was supported by a DFG research grant (KE 1476/2-2) to the first author.

References

Bates, D., Maechler, M., Bolker, B., Walker, S., et al. (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version*, *1*(7), 1–23.

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356–371.

Dewaele, J.-M. (2012). Personality in second language acquisition. *The Encyclopedia of Applied Linguistics*.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. (2016). The relationship between artificial and second language learning. *Cognitive Science*, *40*(4), 822–847.

Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading span task performance, linguistic experience, and the processing of unexpected syntactic events. *The Quarterly Journal of Experimental Psychology*, *70*(3), 413–433.

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*(7), 1243–1252.

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing statistical learning implicitly: A novel chunk-based measure of statistical learning. In *The 39th Annual Conference of the Cognitive Science Society (CogSci 2017)* (pp. 564–569).

James, A. N., Fraundorf, S. H., Lee, E.-K., & Watson, D. G. (2018). Individual differences in syntactic processing: Is there evidence for reader-text interactions? *Journal of Memory and Language*, *102*, 155–181.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, *3*(2), 114–158.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340.

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The language experience and proficiency questionnaire (leap-q): Ten years later. *Bilingualism: Language and Cognition*, 1–6.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, 154–169.

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without sli? *Journal of Child Language*, *41*(2), 327–351.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302–331.

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the author recognition test. *Behavior Research Methods*, *47*(4), 1095–1109.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.

R Core Team. (2019). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160059.

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, *49*(2), 418–432.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120.

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, 35–50.