Title
Grammatical Generalisation in Statistical Learning: Is it Implicit and Invariant Across development?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Hickey, Amanda J.
Hayiou-Thomas, Marianna E.
Mirković, Jelena

Publication Date
2019

Peer reviewed

Grammatical Generalisation in Statistical Learning: Is it Implicit and Invariant Across development?

Amanda J. Hickey (ajh650@york.ac.uk)
Department of Psychology, University of York, York, YO10 5DD

Marianna E. Hayiou-Thomas (emma.hayiou-thomas@york.ac.uk)
Department of Psychology, University of York , York, YO10 5DD

Jelena Mirković (jelena.mirkovic@york.ac.uk)
School of Psychological and Social Sciences, York St John University, Lord Mayor's Walk, York, YO31 7EX
Department of Psychology, University of York , York, YO10 5DD

Abstract

The learning and generalisation of grammatical regularities is fundamental to successful language acquisition and use. Research into statistical learning has started to consider how this process occurs through the implicit detection and assimilation of grammatical regularities. This study focuses on how adults and children generalise regularities and explores the role of explicit knowledge in this process. Across three experiments, adults and children learnt an artificial language containing two semantic categories denoted by a co-occurring determiner and suffix. Explicit knowledge of the regularities was associated with generalisation performance in adults but not children, even when adult word level knowledge was similar to children's. The implications of these results for developmental theories of grammatical generalisation are discussed.

Keywords: statistical learning, explicit knowledge, grammatical categories, artificial language, learning, generalisation.

Introduction

A key aspect of language acquisition and use is the ability to learn grammatical regularities that are present in the input, and generalise them to novel situations. Statistical learning (SL) has been suggested as one of the key mechanisms for the acquisition of grammatical regularities (Gómez & Gerken, 2000). For example, corpus studies of child directed speech have shown the presence of two statistical cues, which reliably indicate grammatical categories: phonological and distributional cues (e.g. Monaghan et al., 2005). Phonological cues consist of speech sounds which are associated with word class (e.g. in English, phoneme length can indicate a noun or a verb), and can be at the word, syllable or phoneme level. Distributional cues relate to the linguistic context within which a word usually sits, for instance where two co-occurring words (or morphemes) frame an interleaved word stem (e.g. '…is walking; Monaghan et al., 2005). Both adults and children use these cues in natural language learning and processing (e.g. Farmer et al., 2006; Lew-William & Fernald 2007; 2010). Research using artificial languages incorporating these types of cues has also shown that adults and children are able to utilise them when learning grammatical categories (e.g. Lany & Saffran, 2010; 2011; Mirković, Forest & Gaskell, 2011; Mirković & Gaskell, 2016; Frost, Monaghan & Christiansen, 2019). For example, Gómez (2002) found that both young infants (17-19 months old) and adults were able to detect and learn distributional cues and Hall, Horne and Farmer (2018) demonstrated the use of distributional cues in the learning and generalisation of grammatical categories in older children (6-9 years old).

The role of semantic cues has also been assessed. For example, in 20-month-old infants, the learning of semantic cues was supported by deterministic phonological and distributional cues for grammatical categories (Lany and Saffran, 2011), and to a lesser extent by probabilistic mappings between semantics and distributional cues (Lany, 2014). More recently, distributional cues have also been shown to enhance the learning of word-referent mappings in adults (Frost, Monaghan & Christiansen, 2019). Adults have also been shown to use semantic cues to generalise grammatical gender-like classes to previously unseen items, in a probabilistic artificial language (Mirkovic et al., 2011; Mirkovic & Gaskell, 2016).

In sum, the research on both adults and children and infants demonstrates that they can use SL to learn grammatical categories from statistical cues. Although not all studies assess generalisation of newly formed grammatical knowledge to previously unseen items, those that do show that both adults (e.g. Mirković et al., 2011) and children and infants (e.g. Lany & Saffran, 2010; 2011; Wonnacott et al., 2012) are able to do so. However, an important open question concerns the processes that support successful generalisation, and whether these processes differ in adults and children.

It is typically assumed that SL is an implicit (unconscious) process that is invariant across different ages (e.g. Aslin & Newport, 2012). However, more recent studies (Batterink et al., 2015; Franco et al, 2011; Conway & Christiansen, 2005) suggest that both implicit and explicit processes play a role in adult statistical learning. By drawing parallels between the implicit learning literature (e.g. Reber, 1967) and SL, these authors consider how 'implicit' implicit learning tasks really are in the context of SL.

To test the relative roles of implicit and explicit processes in SL, Batterink et al., (2015) incorporated on-line measures of implicit learning (reaction times and ERPs) into a word boundary SL task with adults (based on the paradigm used by Saffran et al., 1996). The results suggested that both implicit and explicit processes were involved in the detection of word boundaries. In a similar vein, Smalle et al. (2017) used a Hebb sequence-learning paradigm to examine the relative role of explicit/implicit processing in children as compared to adults. Although both adults and children showed evidence of explicit awareness of the learned sequences, there were some notable developmental differences: adults' explicit awareness emerged at an earlier point during learning than that of children. Furthermore, while explicit awareness was significantly associated with Hebb learning performance in adults, this association was not present in children. This suggests that adults were drawing on both explicit and implicit learning mechanisms during this task, while children relied on implicit learning (Smalle et al., 2017). These studies suggest that both explicit and implicit processes are involved in SL and that the relative contributions of these processes may differ between adults and children.

Current Study

The key aim of the current study was to examine the role of explicit knowledge in grammatical category generalisation, and whether this differs in children and adults. Across three experiments, participants were trained on an artificial language using phonological, distributional and semantic cues to create a grammatical gender-like noun class system (Mirkovic et al., 2011). We tested adult and school-aged child participants and examined the role of explicit knowledge in the generalisation of grammatical regularities to previously unseen items. We manipulated the type of training and the level of initial learning of the novel nouns.

The artificial language consisted of the noun "classes" based on semantic, phonological, and distributional cues. To create the semantic cues, two semantic categories were used: animals and artefacts. The phonological cues were incorporated using a "suffix" (e.g. mof*eem*). The distributional cues were incorporated as a co-occurrence of a "determiner" and a "suffix" (see Table 1 for examples). Each determiner and each suffix was paired with a semantic category (animals or artefacts). This provided an *a*X*b* structure for animals and *c*X*d* structure for artefacts, with X denoting the interleaving arbitrary stem, *a* and *c* the determiner and the *b* and *d* the suffix.

Across all studies participants were trained using a word learning task (with no reference to underpinning 'grammatical' regularities. After training, they were tested on three generalisation tasks focusing on the three different cues (explained below). Levels of emergent explicit knowledge were assessed at the end of the experiment. Experiment 1 included adults and children, while Experiments 2 and 3 included adults only. Across the three experiments, we manipulated two factors that we hypothesised would contribute to generalisation and the emergence of explicit

Table 1: Design of the noun classes

	Determiner	Suffix	Examples
animal	tib	eem	*tib* mof*eem* = dog *tib* zeap*eem* = duck
artefact	ked	ool	*ked* larsh*ool* = table *ked* snar*ool*=TV

knowledge: i) initial levels of word learning, and ii) type of training. In Experiment 1, participants were exposed to a fixed number of repetitions of the novel words at training using a word-picture matching task (WPM; Breitenstein et al., 2007), and a word repetition task. Experiments 2 and 3 used criterion learning, with adult participants matched in the level of initial word learning to the children in Experiment 1. We hypothesized that levels of initial word learning may influence generalisation performance and levels of explicit knowledge of the 'grammatical' regularities. In addition, we removed word-picture matching from the training procedure in Experiment 3, to test the hypothesis that explicit selection may contribute to the emergence of explicit knowledge of the 'grammatical' regularities. In all three experiments, we examined the extent to which generalisation performance was associated with the emergent explicit knowledge of the phonological, distributional, and semantic cues.

Method

Participants

Experiment 1. Sixty-one participants took part: 31 adults with a mean age of 19.70 years (19.08-20.67 years; 1 male) and 30 children with a mean age of 10.21 years (9.67-10.82 years; 13 males). The adult sample was drawn from the undergraduate population at the University of York and received course credits for their participation. The child sample was drawn from primary schools in North Yorkshire.

Experiments 2 & 3. Thirty participants took part in Experiment 2 with a mean age of 20.77 years (18.17-32.58 years; 4 male), and thirty in Experiment 3 with a mean age of 21.09 years (18.25-31.58 years; 5 males). These two samples were drawn from the undergraduate population at the University of York and received course credits or payment for their participation.

Stimuli The training and testing tasks in all experiments used pictures drawn from Rossion and Purtois (2001) object database (281x173ppi) and artificial words created from the English database of pronounceable nonwords (Rastle et al., 2002). The artificial 'words' were constructed using the three elements described earlier (e.g. *a*X*b*) and were digitally recorded (produced by a native speaker of English). This process was based on the stimuli created by Mirković et al., (2011).

All arbitrary stem (X) elements consisted of one syllable with a CVC, CCVC or CVCC (C= Consonant, V = Vowel; 'CAT' = CVC) structure. An overall balance of CVC, CCVC and CVCC words between the animal and artefact training words was controlled for. The stem onset phoneme did not match the onset phoneme of the English word for the animal/artefact it was paired with. The same training and generalisation sets were used in all three experiments.

Training Set: Thirty-two word-picture pairs were created (16 in each semantic category). Each word was paired with a picture, which denoted the assigned meaning of the word, providing the non-arbitrary semantic cue (see Table 1 for examples).

Generalisation Sets: Three different sets of 8 generalisation items were designed to test post-training performance on previously unseen items. Each set consisted of 4 items that were consistent with the trained regularities, and 4 items that were inconsistent. Higher endorsement rates for consistent vs inconsistent items was taken to indicate learning of the regularities.

Determiner and Suffix Generalisation. This task was designed to test learning of the mapping between the determiner and suffix, and the associated semantic category. Eight novel words were presented with novel picture pairings. The four consistent items conformed entirely to the regularities present in the training set. In the four inconsistent items, the structure of the word conformed to the training set (e.g. *tib darleem*), but it was presented with a picture from the 'wrong' semantic category (e.g. *tib darleem* was paired with an artefact, instead of an animal).

Suffix-Only Generalisation. This task tested learning of the co-occurrence between the semantic category and the suffix specifically; that is, the 'phonological' cue. The 8 novel words were presented with novel picture pairings; as before, the 4 consistent items conformed to the regularities in the training set. In the inconsistent items, the determiner 'matched' the picture, but the suffix did not match either the determiner or the picture (e.g.*tib* sen*ool* was paired with a picture of a goat; where the co-occurrence of 'tib' with the picture of an animal conformed to the training set, but the suffix 'ool' was inconsistent with both the determiner 'tib' and the semantic category of animal).

Phonological Form Generalisation. This task specifically tested learning of the co-occurrence of the determiner and suffix; that is, the 'distributional' cue. Eight novel words were presented without pictures. The 4 consistent items conformed to the regularities used in the training set. The 4 inconsistent items had a mismatch between the determiner and the suffix (e.g. *tib* jit*ool* and *ked* narp*eem*).

Procedure Participants completed all tasks in one session of approximately 40-60 minutes. Responses were recorded by the 'DMDX' programme (Forster & Forster, 2003) on a PC laptop computer. Participants were introduced to experimental tasks as a series of games involving 'alien' words introduced by a visiting extra-terrestrial. The training procedure varied across the three experiments, but they all used the same testing protocol.

Experiment 1 training:

Repetition: The thirty-two training stimuli were presented once within a block, for three blocks. Participants were instructed to look at the picture and listen to the 'alien' word and repeat the word aloud once. Participants completed this task twice.

WPM: Participants were presented with word-picture pairs and were instructed to judge if they thought the word and picture '*went well together'*. Participants were exposed to all 32 word-picture pairs once. In addition, 16 of the word items were presented again paired with a different picture from the same semantic category (mismatch trials) for the 'incorrect' response. The participant responded using keys on the computer keyboard: a "happy face" if they thought the picture and word went well together and a "sad face" if they did not. Participants completed this task twice.

Experiment 2 training:

Repetition and WPM: For this experiment, the repetition and WPM tasks were merged. In each block, participants were exposed to all 32 word-picture pairs once. In addition, 8 mismatch trials were included, in which the word items were paired with an incorrect picture from the same semantic category. Participants were instructed to look at the picture and listen to the 'alien' word and repeat it aloud once. They then pressed the space bar and then judged if the word and picture '*went well together'* using the same WPM response procedure from Experiment 1.

Each training block was followed by a word-learning test (described below). Training ended when the participant reached the same level of accuracy as that of the children in Experiment 1 (75%).

Experiment 3 training:

Repetition Only: The training set for Experiment 3 was the same as that for Experiment 2, including criterion learning. However, the training procedure was different in that it did not include WPM: participants had only to repeat the training items.

All Experiments: Testing

Word Learning –Two Alternative Forced Choice (2AFC): This task tested learning of the novel words. Each word was randomly presented once and was accompanied by the simultaneous presentation of two pictures (on either side of the screen), one of which was the correct trained picture. The 'foil' picture was drawn from the trained pictures and was from the same semantic category. Participants responded using keys on the computer keyboard which corresponded to the on-screen picture presentation position.

Generalisation: "Determiner and Suffix" and "Suffix Only" Generalisation. In both these tasks, participants were instructed to attend to 'alien*'* word and picture pairings (from the respective generalisation sets) and judge if they thought they '*went well together',* pressing the happy or sad face accordingly.

"Phonological Form": Participants were instructed to listen to the 'alien' words from the generalisation set and asked to judge if the words '*went well with'* the 'alien' language they had been listening to, pressing the happy or sad face accordingly.

Explicit Knowledge Questionnaire: Once all tasks were completed, participants were asked *'Did you notice anything about the alien language? Did you use any kind of strategies or clues to decide whether the word and the picture matched?'* Answers were recorded manually and a score from 0-3 was given separately for determiner and suffix knowledge: 0 for no reference to the morpheme or semantic dependency, 1 for knowledge of the morpheme but not the dependent semantic cue, 2 for partial knowledge of the morpheme and semantic dependency and 3 for full knowledge.

Results and Discussion

Word Learning: We examined the level of word learning in the three experiments by analysing performance on the 2AFC task at the end of training. One-sample t-tests against chance (.5) showed that all groups learned the novel words. (Experiment 1 adults, t(30)=28.93, *p*<.001; children, t(29)=9.10, *p*<.001; Experiment 2 adults, t(29)=21.17, *p*<.001; Experiment 3 adults, t(29)=21.83, *p*<.001; Figure 1). Adult participants in Experiments 2 and 3 were trained to the criterion matching the levels of child word learning in Experiment 1. To confirm that the word learning across the three studies matched as intended, we ran two multiple regressions with 2AFC performance as the outcome variable and group as the predictor variable coded using Helmert contrasts. The first set of contrasts showed that, as expected, adults learned more words in Experiment 1 (Adults1) than in Experiments 2 (Adults2) and 3 (Adults3; β=0.68, *p*<.001), and that there was no difference between the latter two (*p*=.293). The second set of contrasts showed that Adults2 (*p*=.341) and Adults3 (*p*=.757) learnt an equivalent number of words to the children in Experiment 1 (Children1). These findings show that all participants demonstrated word learning. Crucially, these results indicate that criterion learning method used in Experiments 2 and 3 was successful at reducing the level of adult participants' word-learning to that of the child participants in Experiment 1.

Generalisation Tasks To analyse performance in the generalisation tasks, we derived an A' metric based on the endorsement rates for consistent and inconsistent trials (Pallier, 2002). A' scores above 0.5 were taken as indication that a participant could reliably endorse consistent trials more often than inconsistent trials, demonstrating learning of the regularities.

"Determiner and Suffix": Figure 2 shows levels of generalisation performance for all groups on this task. One-sample t-tests showed that only Adults1 performed significantly above 0.5 (t(30)=6.13, *p*<.001). Thus, only this group demonstrated learning and generalisation of the mapping between the determiner and suffix, and the semantic category.

Using the same set of contrasts as in the analysis of word learning, with A' performance as the outcome variable and group contrasts as the predictor variables, group comparisons further confirmed that Adults1 were significantly better at generalising this regularity than Adults2 and Adults3 (β=0.30, *p*<.004). There was no difference between the Adults2 and Adults3 (*p*=.703), nor between Children1 and Adults2 (*p*=.451) or Adults3 (*p*=.916).

"Suffix Only": Figure 3 shows generalisation performance for all groups on this task. One-sample t-tests demonstrated that Adults1 (t(30)=3.45, *p*<.001) and Adults2 (t(29)=1.97, *p*=.029) performed significantly above an A' of 0.5. Children1 (*p*=.762) and Adults3 (*p*=.500) did not. Therefore, only Adults1 and Adults2 showed learning and generalisation of the mapping between the semantic category and the suffix (the phonological cue).

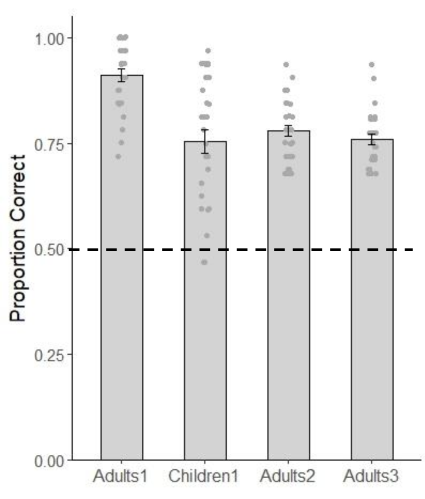

Figure 1: Word Learning: Accuracy on the 2AFC task at the end of training.

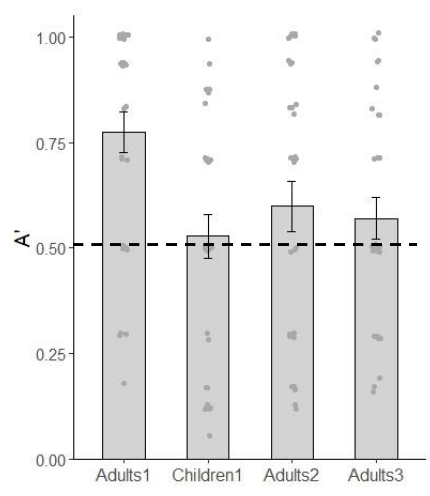

Figure 2: Performance on the "Determiner & Suffix" Generalisation Task

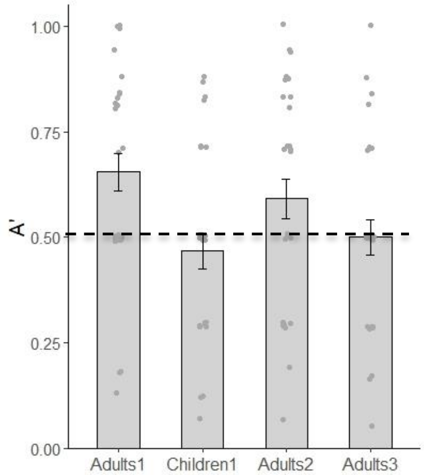

Figure 3: Performance on the Suffix Only Generalisation Task

Group comparisons showed differences between Adults 1 compared to Adults2 and Adults3 (β=0.25, *p*=.044) and between Children1 and adults 2 (β=0.35, *p*<.044). There was no difference between Adults 2 and 3 (*p*=.148) or between Children1 and Adults3 (*p*=.581).

"Phonological Form": As illustrated in Figure 4, participants in all groups and experiments performed at a similar level. One-sample t-tests demonstrated that only Adults1 performed significantly above 0.5 (t(30)=2.93, *p*<.001). Thus, only this group showed evidence of learning and generalising the co-occurrence between the determiner and the suffix (the distributional cue).

The group comparisons showed there was no evidence of a differences in generalisation across the three experiments, or between adults and children (Adults1 vs. Adults2&3, *p*=.083; Adults2 vs. Adults3, *p*=.294; Children1 vs. Adults2, *p*=.513;

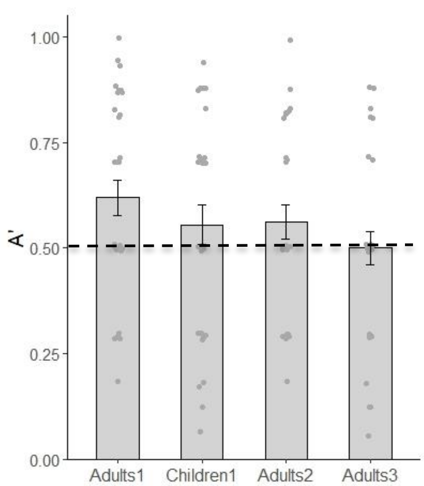

Figure 4: Performance on the "Phonological Form" Generalisation Task

Children1 vs. Adults3, *p*=.268). These results show only weak evidence for the learning and generalisation of the determiner + suffix co-occurrence in this paradigm.

In summary, only adults in Experiment 1 demonstrated the ability to utilise all three statistical cues to generalise newly formed grammatical knowledge. Their performance was reliably different from that of children, and of adults in Experiments 2 and 3, when generalising the trained cues to novel exemplars.

Given that Adults2 and Adults3 show the same level of word learning as Children1, this result suggests that the level of word learning may be a driver of grammatical generalisation and as such may explain some of the difference seen between adults and children in Experiment 1. This finding aligns with Bates and Goodman's (1997) lexicalist theory, which proposes that the emergence of grammar depends on lexical learning.

Although the lack of generalisation in children found here is in contrast to some previous studies in the literature (e.g. Hall et al., 2018; Lany & Saffran, 2010; 2011), this may be due to a number of factors, including the nature of the training and the structure and complexity of the regularities. In the current study, the training tasks always included simultaneous presentation of the referent with the novel words, unlike e.g. Lany and Saffran (2010; 2011) and Lany (2014), who trained participants on the phonological word form before introducing the referent, and Hall et al., (2018), who trained participants on a language that did not include a referent. Moreover, the simultaneous presentation of all noun class cues (phonological, distributional, and semantic) in the current study may have increased the complexity of the task, and affected the relative salience of the cues. Thus, further research exploring these methodological differences would help to clarify the role of semantic cues, and the effects of sequential vs simultaneous presentation of different type of cue.

Explicit Knowledge: Contributions to Generalisation

Table 2 shows the explicit knowledge scores for each group, presented separately for each morpheme. These scores suggest greater explicit knowledge for determiners than for suffixes across all groups/experiments.

A key aim in the current study was to assess the extent to which explicit knowledge contributes to generalisation performance, and whether this contribution differs in children and adults. We were specifically interested not in the group differences between children and adults (as children may be less able to verbalise their knowledge), but in the extent to which individual variation in the levels of explicit knowledge

Table 2: Descriptive Statistics for Explicit Knowledge Scores

	Determiner		Suffix	
	M	*SD*	*M*	*SD*
Adults1	2.39	1.02	0.58	1.03
Children1	0.70	1.11	0.10	0.31
Adults2	1.20	1.19	0.33	0.76
Adults3	1.50	1.17	0.23	0.63

within each group contributes to generalisation performance. To address these questions, multiple regressions were carried out for each generalisation task. The outcome variable in each regression was the A' score, while explicit knowledge scores for the relevant morpheme(s) were the predictor(s). For example, in the "phonological form" and "suffix only" generalisation tasks, only knowledge of the suffix was necessary for successful performance, so only suffix knowledge was used as a predictor, while for the "determiner and suffix" task, knowledge of both determiner and suffix was relevant.

As illustrated in Table 3, for adults in Experiment 1 explicit knowledge of the regularities was a significant predictor of generalisation performance in the "determiner & suffix" and "suffix only" tasks, but not in the "phonological form" task. Explicit knowledge of the relevant morpheme facilitated performance in the generalisation tasks. The strongest effect was for knowledge of the determiner-semantic mapping, which accounted for 27% of the variance in the 'determiner & suffix' task. In contrast to the adults, explicit knowledge of the regularities in children did not significantly predict performance in any of the generalisation tasks.

The pattern of results in Experiment 2 provides an informative comparison because the adults in this group showed low levels of generalisation (comparable to children), and intermediate levels of explicit awareness. Nonetheless, variability in generalisation within this group was significantly predicted by explicit awareness of the relevant morphemes. As with adults in Experiment 1, the strongest effect was for knowledge of the determiner-semantic mapping, which in this case accounted for an even larger proportion (38%) of the variance in the 'determiner & suffix' task. Finally, the adults in Experiment 3, showed low levels of generalisation as well as low levels of explicit knowledge. In this case, and similarly to the children in Experiment 1, there was no clear evidence of facilitatory effect of explicit knowledge on generalisation performance. This may suggest that the use of the WPM training task could prompt the emergence and correct use of explicit knowledge, at least in adults.

Overall these results suggest a partial role for explicit knowledge in grammatical generalisation for adults but not children. This still seems to hold when adults demonstrate similar levels of word learning and generalisation to children, suggesting that there may be differences in the extent to which adults and children draw on explicit processes when generalising in a grammatical SL task.

Conclusion

The current set of experiments demonstrates that explicit knowledge plays a role in grammatical category generalisation in adults but not children. This may be partially due to children's lower level of word knowledge, given the lower level of generalisation performance in adults when levels of word knowledge were matched to those of children. However, adults with a lower level of word knowledge still demonstrated a partial involvement of explicit knowledge in the generalisation tasks. This suggests the possibility of developmental differences between adults and children in the role of explicit and implicit processes when generalising in SL tasks. In future studies, more sensitive measures that do not rely on verbal reports would provide further insights into the contributions of explicit knowledge in implicit learning tasks across development.

Table 3: Multiple Regressions for the Role of Explicit Morpheme Knowledge on Generalisation Performance.

	Experiment 1 Adults					Experiment 1 Children				
	R^2	B	SE B	ß	*p*	R^2	B	SE B	ß	*p*
Determiner & Suffix Generalisation:	0.27					-0.02				
Explicit Determiner Knowledge	**0.27**	**0.14**	**0.04**	**0.54**	**.002**	-0.02	0.03	0.05	0.11	.558
Explicit Suffix Knowledge	0.00	0.02	0.04	0.08	.599	-0.00	0.15	0.18	0.16	.403
Suffix Only Generalisation:	0.16					0.01				
Explicit Suffix Knowledge		**0.10**	**0.04**	**0.43**	**.015**		0.16	0.13	0.21	..259
Phonological Form Generalisation:	0.05					0.03				
Explicit Suffix Knowledge		0.07	0.04	0.29	.112		0.22	0.15	0.26	.171
	Experiment 2 Adults					Experiment 3 Adults				
Determiner & Suffix Generalisation:	0.46					-0.03				
Explicit Determiner Knowledge	**0.38**	**0.17**	**0.04**	**0.62**	**<.001**	-0.03	0.03	0.04	0.11	.553
Explicit Suffix Knowledge	**0.07**	**0.13**	**0.06**	**0.30**	**.037**	-0.00	0.08	0.08	0.19	.329
Suffix Only Generalisation:	0.10					0.11				
Explicit Suffix Knowledge		0.12	0.06	0.36	.052		**0.13**	**0.06**	**0.37**	**.044**
Phonological Form Generalisation:	0.18					0.12				
Explicit Suffix Knowledge		**0.14**	**0.05**	**0.46**	**.011**		**-0.13**	**0.06**	**-0.38**	**.036**

Significant results are highlighted in bold.

Only morpheme knowledge salient to the generalisation task were included.

References

Aslin, R. N., & Newport, E. L. (2012). Statistical Learning: From Acquiring Specific Items to Forming General Rules. *Current Directions in Psychological Science, 21,* 170-176.

Bates, E., & Goodman, J. C. (2001). On the Inseparability of Grammar and the Lexicon: Evidence from Acquistion, Aphasia and Real-time Processing. *Language and Cognitive Processes, 12,* 507-584.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78.

Breitenstein, C., Zwitserlood, P., de Vries, M. H., Feldhues, C., Knecht, S., & Dobel, C. (2007). Five days versus a lifetime: intense associative vocabulary training generates lexically integrated words. *Restorative Neurology and Neuroscience*, *25*(5-6), 493–500.

Conway, C. M. & Christiansen, M. H. (2005). Modality-Constrained Statistical Learning of Tactile, Visual and Auditory Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 24-39.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*(32), 12203-12208.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior research methods, instruments, & computers*, *35*(1), 116-124.

Franco, A., Cleeremans, A. & Destrebecqz, A. (2011). Statistical Learning of two artificial languages presented successively: how conscious? *Frontiers in Psychology, 2,* 229.

Frost, R. L. A., Monaghan, P., & Christiansen, M. H. (2019). Mark my words: High frequency marker words impact early stages of language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication*.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178–186.

Hall, J., Owen VAN Horne, A., & Farmer, T. (2018). Distributional learning aids linguistic category formation in school-age children. *Journal of Child Language*, *45*(3), 717–735.

Lany, J. (2014). Judging words by their covers and the company they keep: Probabilistic cues support word learning. *Child development*, *85*(4), 1727-1739.

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: infants' acquisition of lexical categories. *Psychological Science*, *21*(2), 284–291.

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science*, *14*(5), 1207–1219.

Mirkovic, J., Forrest, S., & Gaskell, G. (2011). Semantic Regularities in Grammatical Categories: Learning Grammatical Gender in an Artificial Language. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 33. Retrieved from https://escholarship.org/uc/item/54412022.

Mirković, J., & Gaskell, M. G. (2016). Does sleep improve your grammar? Preferential consolidation of arbitrary components of new linguistic knowledge. *PloS one*, *11*(4), e0152489.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*(2), 143–182.

.Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: the ARC Nonword Database. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *55*(4), 1339–1362.

Reber, A. S. (1967). Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behaviour*, *6*, 855-863.

Rossion, B., & Pourtois, G. (2001). Revisiting Snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, *1*(3), 413-413.

Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M. & Szmalec, A. (2017). Children retain implicitly learned phonological sequences better than adults: a longitudinal study. *Developmental Science, 21(5),* e12634.

Wonnacott, E., Boyd, J. K., Thomson, J., & Goldberg, A. E. (2012). Input effects on the acquisition of a novel phrasal construction in 5 year olds. *Journal of Memory and Language*, *66*(3), 458–478.