Differences in Implicit vs. Explicit Grammar Processing as Revealed by Drift-Diffusion Modeling of Reaction Times

David Abugaber (dabuga2@uic.edu)
Department of Hispanic and Italian Studies,
University of Illinois at Chicago, 606 S. Morgan Street, Chicago, IL 53706 USA

Kara Morgan-Short (karams@uic.edu)
Department of Hispanic and Italian Studies; Department of Psychology,
University of Illinois at Chicago, 606 S. Morgan Street, Chicago, IL 53706 USA

Abstract

Learning new languages is a complex cognitive task involving both implicit and explicit processes. Batterink, Oudiette, Reber, and Paller (2014) report that participants with vs. without conscious awareness of a hidden semi-artificial language regularity showed no significant differences in behavioral measures of grammar learning, suggesting that implicit/explicit routes may be functionally equivalent. However, their operationalization of learning via median reaction times might not capture underlying differences in cognitive processes. In a conceptual replication, we compared rule-aware (n=14) and rule-unaware (n=21) participants via drift-diffusion modeling, which can quantify distinct subcomponents of evidence-accumulation processes (Ratcliff & Rouder, 1998). For both groups, grammar learning was manifested in non-decision parameters, suggesting anticipation of motor responses. For rule-aware participants only, learning also affected bias in evidence accumulation during word reading. These results suggest that implicit grammar learning may be manifested through low-level mechanisms whereas explicit grammar learning may involve more direct engagement with encoded target meanings.

Keywords: artificial language; drift-diffusion; evidence accumulation; second language; grammar; implicit; explicit

Background

Learning a new language is a complex cognitive task involving both explicit and implicit processes (i.e., that do/do not involve conscious awareness). Understanding how these processes interact is essential to a full account of second language (L2) learning (for a review, see Leow, 2015). One way to study implicit and explicit language processing comes from semi-artificial language paradigms involving covert regularities in pseudoword articles encoding word meaning (e.g., a word's living/non-living status). Such studies indicate that learning an untaught rule can proceed in the absence of conscious awareness in rule-unaware participants, as indexed by above-chance accuracy on forced choice tasks (e.g., Williams, 2005) and by reaction time slowdowns to rule-violating exemplars (e.g., Batterink, Oudiette, Reber, & Paller, 2014; Leung & Williams, 2011; but cf. Faretta-Stutenberg & Morgan-Short, 2011), suggesting that explicit knowledge is not strictly necessary for learning to occur. At the same time, other studies with untaught rules demonstrate that learners are likely to acquire both implicit and explicit knowledge (e.g., Grey, Williams, & Rebuschat, 2014). A core question that underlies the interpretation of such studies is whether rule-aware and rule-unaware participants in such semi-artificial language paradigms take qualitatively different routes to grammar processing, or if their performance is underlyingly based on the same mechanisms. One possibility is that participants with rule awareness consciously and willfully apply strategies such as rule searches, generation of explicit predictions, and systematic hypothesis testing (Leow, 2015). Indirect evidence for such an account comes from word sequence learning paradigms that suggest that experiment-initial instruction can improve performance on rule-adhering trials but worsen detection of rule-violating trials, suggesting functional differences between rule-aware/rule-unaware processing (Batterink, Reber, & Paller, 2015). However, the presence of explicit knowledge does not preclude the possibility that other, more implicit kinds of learning can drive task performance (Rebuschat et al., 2013). In support of such an interpretation, Rose, Haider, and Büchel (2010) found that neural and behavioral markers of learning can emerge at time points in the experiment *before* participant reports of the emergence of rule awareness, suggesting that conscious rule awareness might emerge as a consequence of implicit learning.

The available evidence on grammar processing in semi-artificial language experiments (e.g., Leung & Williams, 2011; Rebuschat, et al., 2013; Batterink et al., 2014) is not sufficient to determine whether rule-aware and rule-unaware participants use different grammar-processing mechanisms. Such studies have typically used linear analyses of measures of central tendency (e.g., means or medians) to compare reaction times, such that grammar learning can be measured as general slowdowns to rule-violating trials relative to rule-adhering trials. However, a finding that rule-aware and rule-unaware participants do not differ in these measures (as in Batterink et al., 2014) does not allow one to infer that the underlying processes were not different in subtler ways that do not affect means or medians directly in ways that can be detected through a traditional linear analysis (Balota & Spieler, 1999; Whelan, 2008). Although electroencephalography (EEG) data from one study using a semi-artificial language paradigm showed different brain responses for rule-aware vs. rule-unaware participants (Batterink et al., 2014), as the authors note, interpretation of these components is problematic because overlap in component timing may cause EEG signals to essentially cancel each other out. As such, in addition to the processing signal that rule-aware participants evidenced (claimed to reflect explicit processing), they may have also engaged the same grammar processing activity as rule-unaware participants. This processing signal (claimed to reflect implicit processing) may have been obscured by overlap with processing signal that was detected in these learners.[1] Due to limitations of previous behavioral and EEG data analysis approaches, it is not yet established whether grammar processing in rule-aware/rule-unaware participants involves distinct mechanisms.

This study aims to address this gap in the extant research through drift-diffusion modeling (Ratcliff & Rouder, 1998), which belongs to a family of evidence-accumulation models that allow one to determine precisely how different participant groups might vary in their response time characteristics even when the central tendencies of their reaction time distributions are the same. The drift-diffusion model is based on the idea that each decision in a two-choice context is made in a continuous fashion by sampling noisy evidence that accumulates until a decision boundary threshold has been crossed in favor of one response or the other for each trial. Because such models simultaneously fit response times and accuracy/choice direction, drift-diffusion modeling can also account for speed-accuracy tradeoffs.

Drift-diffusion modeling allows us to determine whether participants differ in terms of model parameter estimates that capture certain constructs from cognitive psychology, i.e.:

- ***v***: speed of evidence accumulation towards the response in each experimental trial;
- ***z:*** bias in evidence accumulation towards one response or another, at the start of each trial;
- ***t0***: time spent in non-decision-related processes, e.g., tied to factors like speed of motor responses or of low-level perception;
- ***a*****:** threshold of accumulated evidence before a response is provided in each experimental trial

Trial-to-trial within-person variance in any of the parameters listed above can be formally included in the model, e.g., as s_v, the standard deviation of v; as s_z, the standard deviation of z, etc. Finally, testing for significant differences in these drift-diffusion parameters allows one to determine how experimental manipulations can affect manifestation of the constructs from cognitive psychology mentioned above.

How can the drift-diffusion modeling approach be leveraged to determine whether and how conscious rule knowledge affects grammar learning? In the original experiment design from Batterink et al. (2014; based on the semi-artificial language from prior studies, e.g., Williams, 2005; Leung & Williams, 2011; Faretta-Stutenberg & Morgan-Short, 2011), participants are shown four novel articles (*gi, ul, ro,* and *ne*) and told that these encoded the distance of a co-occurring English noun, such that two of the articles are used with distant referents and two are used with nearby referents. However, there was also an underlying, untaught regularity in the semantic features encoded by these articles: namely, two of the articles were usually used with living things (e.g., *horse*,), and two articles usually used with inanimate nouns (e.g., *stereo*,). As such, learning of the underlying rule can be captured via response times/accuracies to a "living/non-living" response task across trial conditions that adhere to vs. violate the underlying rule. Critically, in the experiment presented in Batterink et al. (2014), the trial design is such that the living/non-living-encoding article is presented with an English noun together on a screen simultaneously. However, if the pseudoword is shown *before* the English noun, then it is possible to disentangle different cognitive processes as described below.

How might the drift-diffusion model parameters align conceptually with different hypothesized cognitive processes of grammar learning in our experiment design? We argue that the effects of reading the meaning-encoding article in isolation can be manifested in at least two (non-mutually-exclusive) ways: if the information provided by the pseudoword article regarding the correct button selection for the upcoming living/non-living response involves any degree of motor response anticipation (i.e., if participants become attuned to the button response assignments in the experiment and thus learn to predict which button is usually associated with the correct upcoming response, regardless of what the button "means" in terms of grammatically-encoded semantics), then differences between rule-adhering and rule-violating trials would be manifested to some degree via the t_0 parameter, which captures time spent in decision processes that are *not* tied to evidence accumulation from presentation of the stimulus that initiates the evidence accumulation process, i.e., the English noun in the case of our experiment). By contrast, if the information provided by the pseudoword article involves any degree of higher-level processes (e.g., mentally activating the concept of "living-ness" from the semantics grammatically encoded by the artificial language article), then effects would be manifested through one of the other drift-diffusion model parameters. More specifically, if participants start each trial with pre-activation of the semantics encoded by the pseudoword article such that evidence towards the correct living/non-living response is "pre-accumulated," then differences between rule-adhering and rule-adhering trials would be manifested via the z parameter, which captures biases in evidence accumulation at trial start. By contrast, if participants' rule-learning entails become faster at activating semantics when a noun appears in a rule-adhering (vs. rule-violating) context, then differences between these trials would be seen in the v parameter, which captures speed of evidence accumulation towards the correct response. Alternately, participants could react to rule-violating combinations by changing the threshold of overall

[1] In this same study, slow-wave and REM sleep showed similar benefits for rule-aware and rule-unaware participants, suggesting a similar neural mechanism underlying both kinds of processing (Batterink et al., 2014). However, this does not rule out the possibility that sleep benefits were due to factors that were not strictly cognitive (e.g., effects on mood, physical comfort, etc.).

accumulated evidence that they require before providing a living/nonliving response in each trial (parameter *a*). Finally, it is possible that rule-adhering and rule-violating trials could differ systematically in how much any of these parameters vary on a trial-to-trial basis, in which case we would expect significant differences (across rule-adhering/rule-violating trials) in the parameters related to variance (i.e., the standard deviations captured in parameters s_v, s_z, and s_{t0} for v, z, and t_0, respectively). These interpretations of the drift diffusion model parameters in the context of our experiment are visualized in Figure 1 below.

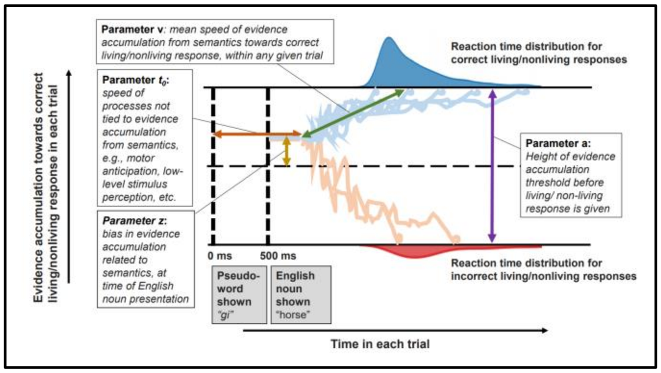

Figure 1: Visualization of drift-diffusion model in the context of our experiment paradigm. Not pictured: parameters capturing standard deviation of v, t_0, and z Visualization inspired by Fig. 1 in Vinding et al. (2018)

As discussed above, drift-diffusion modeling allows us to determine whether rule-aware vs. rule-unaware participants' response times differ in terms of model parameter estimates that capture constructs from cognitive psychology such as the speed of evidence accumulation; bias in evidence accumulation; the criterion threshold of evidence before response in each trial; and time spent in non-evidence-related processes, e.g., tied to factors like speed of motor responses or of low-level perception. Testing for significant differences in these drift-diffusion parameters would allow us to infer whether and how conscious rule knowledge affects grammar learning. We ask:

Research Question 1: Do learners in a semi-artificial language experiment show evidence of grammar learning without conscious awareness (conceptually replicating Batterink et al., 2014)?

Research Question 2: Do learners with vs. without conscious awareness of a covert grammar rule differ in grammar processing as revealed by drift-diffusion modeling, and if so, how?

Methods

Our study comprises a conceptual replication of a prior semi-artificial language learning experiment (Batterink et al., 2014) following a popular paradigm in the field of second language acquisition first introduced by Williams (2005).

Participants

Participants for this study were right-handed native speakers of English (N = 40, 27 female). All participants were undergraduate students at a large urban university who received psychology course credit for their participation. Table 1 shows participant attributes collected via a shortened version of the *Language Experience And Proficiency Questionnaire* (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007)

Table 1. Attributes of participants with reported languages of all participants.

Attribute	*Mean* (*SD*)
Gender	27 female, 13 male
Age	18.73 (0.91)
English reading proficiency[a]	4.86 (0.34)
English writing proficiency	4.84 (0.44)
English-speaking proficiency	4.84 (0.44)
% reporting additional lg.	90.45%
Add. lg. reading proficiency	3.54 (1.36)
Add. lg. writing proficiency	3.26 (1.45)
Add. lg. speaking proficiency	3.85 (1.14)

Note: [a]Self-report scale ranges from 1 to 5 with 1 labeled "low proficiency" and 5 labeled "high proficiency."

Procedure

Participants first provided informed consent and then completed a short language background questionnaire to confirm their native English proficiency. Then, they performed a vocabulary pre-training to become familiar with the four novel articles of the semi-artificial language (see Table 2). Subsequently, two blocks of the experimental reaction time task were performed. Finally, a debriefing was conducted to gauge participants' level of rule awareness.

Table 2: Living/non-living and distance assignment of the four semi-artificial language articles.

	Participants are not told…	
	Living	Non-living
Participants are told…		
Near	*gi*	*ro*
Far	*ul*	*ne*

Vocabulary Pretraining Participants were explicitly told that *gi* and *ro* denote nearby referents (e.g., "*gi* bear," "*ro* typewriter") whereas *ul* and *ne* denote distant referents (e.g., "*ul* snake," "*ne* teacher"). They then performed a written forward translation task and an audio-based backward translation task to criterion, just as in Batterink et al. (2014).

Reaction Time Task Each experimental trial (Figure 2) began with the presentation of a fixation cross for 1000 ms, followed by a pseudoword (*ul*, *gi*, *ro*, or *ne*) for 500 ms, and a noun (presented until a living/non-living response was provided until a maximum of 500 ms, after which point a blank screen replaced the noun on the display). After the living/non-living response, participants saw the cue "*Near/Far?*" until this second response was provided based on the pseudoword for that trial. Following Batterink et al. (2014), the four response options (living/nonliving/near/ far) were assigned unique buttons on a standard keyboard.

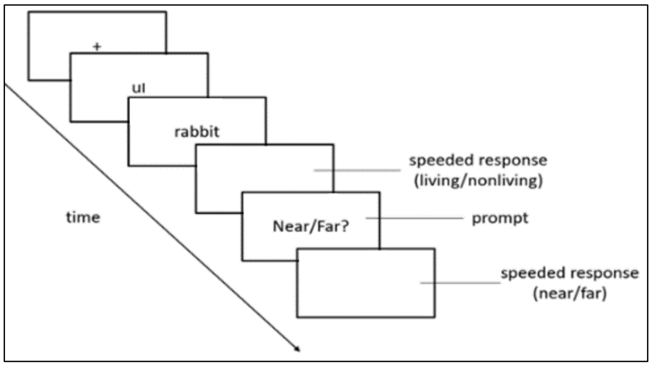

Figure 2: Trial structure for reaction time task.

Half of the presented nouns were living (e.g., *horse*) and the other half were non-living (e.g., *stereo*). Six out of every seven trials were rule-adhering in that they conformed to the living/non-living assignment presented in Table 2, with *gi* and *ul* preceding living nouns and *ro* and *ne* preceding non-living nouns. One randomly selected (rule-violating) trial in each set of seven consecutive did not follow this pattern. To avoid confounds related to the specific nouns assigned to the rule-adhering/rule-violating conditions, stimuli were counterbalanced such that a given noun was presented in the context of a rule-adhering trial for six out of seven participants and in the context of a rule-violating trial for the seventh participant. Additionally, for each participant's stimulus list, nouns assigned to rule-adhering vs. rule-violating conditions did not differ on orthographic word length, frequency, concreteness, positive/negative valence, or arousal. Each noun's order of presentation was randomized within blocks, and assignment of nouns to either the first block or second learning block was counterbalanced across participants. Participants performed a short initial practice block of six (rule-adhering) trials followed by two learning blocks (each comprising 308 experimental trials) with a timed five-minute break in between.

Rule Awareness Debriefing Following the main experimental task, a systematic debriefing was administered to assess the extent of participants' rule awareness. Participants were first asked if they had noticed any pattern about when the different articles were used, beyond the overtly taught near/far rule. If at this point participants spontaneously reported that certain articles co-occurred with living/nonliving referents more often than others, participants were asked at what point they had noticed this pattern (i.e., during the first block, the second block, or only when directly asked during the debriefing). Following the procedure in Batterink et al. (2014), participants who produced the correct pattern and reported having noticed it during either the first or second experimental block were classified as rule-aware. Otherwise, they were classified as rule-unaware.

Analysis

Linear Analysis. To replicate Batterink et al.'s (2014) linear analysis procedure as closely as possible for Research Question 1, our initial measure of rule learning was the Rule Learning Index (RLI), which comprises response time slowdowns for the living/non-living response in rule-violating trials relative to rule-adhering trials. Data from each of the two experimental blocks were divided into four epochs of equal length, yielding eight total epochs. Participants' median RLIs were compared using a Greenhouse-Geisser-corrected mixed 2x2x8 ANOVA, with Awareness (rule-aware vs. rule-unaware) as a between-participants factor and Trial Condition (rule-adhering vs. rule-violating trial) and Epoch (for each of eight experimental epochs) as within-participant factors. Only trials with correct responses to the living/non-living judgment were included in this analysis.

Drift-Diffusion Analysis. To test for differences between rule-aware/rule-unaware participants as per Research Question 2, drift-diffusion modeling was performed on the living/non-living responses using the *rtdists* package (Singmann, Brown, Gretton, & Heathcote, 2020) for the R scripting language. Data were first cleaned by removing reaction times faster than 200 ms and slower than 3000 ms, and only correctly-responded trials were included. Only data from the second block were used, to ensure that enough time had elapsed for sufficient rule-learning to have occurred. Separate models were fit for each participant's rule-adhering and rule-violating trials following the model-fitting procedure used in Singmann (2020), with seven parameters: v (rate of evidence accumulation for the living/non-living response); z (bias in evidence accumulation at start of each trial); $t0$ (non-decision-related times); their standard deviations (s_v, s_z, and s_{t0}, respectively); and finally a (threshold of accumulated evidence before the living/non-living response was provided). For each model-fitting iteration, starting values for each of the parameters were drawn from a random distribution and fitting proceeded until relative convergence was achieved as per the *nlminb()* optimizing function. For each of the seven output parameters in the model, separate 2x2 mixed effects Analyses of Variance (ANOVA) were performed with the within-participant factor Trial Condition (rule-adhering vs. rule-violating trials) and the between-participant factor Awareness (for rule-aware vs. rule-unaware participants). Significant interactions were followed up via Bonferroni-

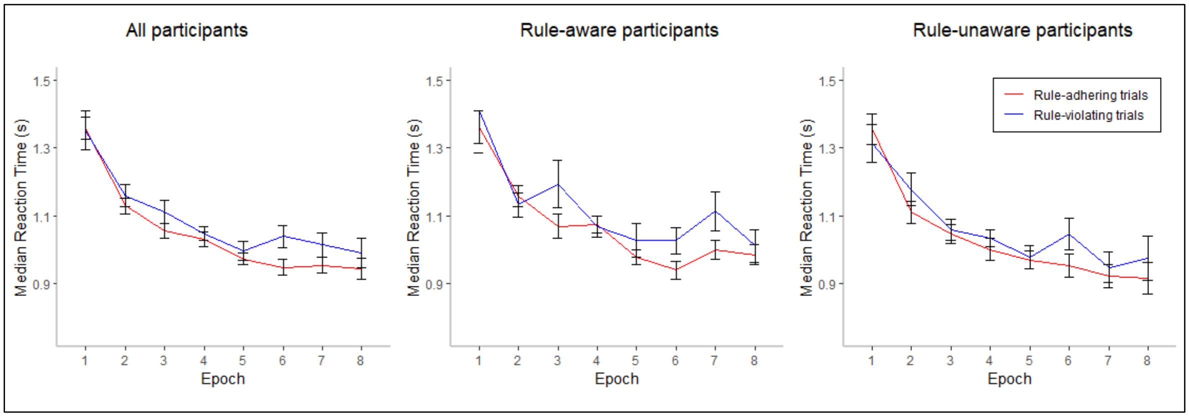

Figure 3: Epoch median response times to rule-adhering vs. rule-violating trials, calculated per participant. This is shown both overall and separately for rule-aware vs. rule-unaware participants.

corrected t-tests with degrees-of-freedom correction for unequal variances using the *emmeans* package for R (Lenth, 2020). Note that because these data comprise only one observation per participant per parameter, running mixed effects models to account for random effects is not possible.

Results

Of the 40 recruited participants, two were excluded due to technical issues and three were excluded due to excessively low accuracies that were not significantly different from chance levels (50%) as per a one-sample t-test on binarily-coded trial-level values (1=correct, 0=incorrect). Of the remaining 35 participants, 14 were coded as aware and 21 as unaware based on their debriefing questionnaire responses.

Linear Analysis Results. Figure 3 shows epoch median response times to rule-adhering vs. rule-violating trials overall as well as separately for rule-aware and rule-unaware participants. For both participant groups, we found decreasing reaction times over the course of the experiment and slower responses to rule-violating trials relative to rule-adhering trials, as confirmed by our three-way ANOVA which yielded a significant main effect of Trial Condition, $F(1, 33) = 7.36$, $p = .011$, $\eta_p^2 = .18$, and of Epoch, $F(3.63, 119.68) = 19.90$, $p < .001$, $\eta_p^2 = .38$. By contrast, there were no significant main effects or interactions with Awareness (all $p > .05$), reproducing Batterink et al. (2014) and suggesting that the learning effect was not different between rule-aware and rule-unaware participants, at least when measured in terms of median reaction times.

Drift-Diffusion Results. Figure 4 shows boxplots with drift-diffusion model parameters estimated separately for rule-adhering vs. rule-violating trials and for rule-aware vs. rule-unaware participants. Our mixed-effects ANOVAs showed no significant effects from either Trial Condition, Awareness, or their interaction on the parameters v, a, s_v, s_{z0}, or s_z. For the t_0 parameter, there was a main effect of Trial Condition, $F(1, 33) = 15.41$, $p < .001$, $\eta_p^2 = .32$ such that rule-violating trials showed higher t_0 values ($M = 0.46$, $SD = 0.19$) relative to rule-adhering trials ($M = 0.39$, $SD = 0.16$), $t(33) = 3.93$, $p < .001$.) Neither Awareness nor the interaction of Trial Condition and Awareness showed statistically significant effects on t_0 ($ps > .05$). The z parameter showed a significant interaction of Awareness by Trial Condition, $F(1, 36) = 6.14$, $p = .018$, such that rule-aware participants showed higher bias towards the correct response for rule-adhering ($M = 0.54$, $SD = 0.11$) relative to rule-violating trials ($M = 0.43$, $SD = 0.17$), $t(33) =$

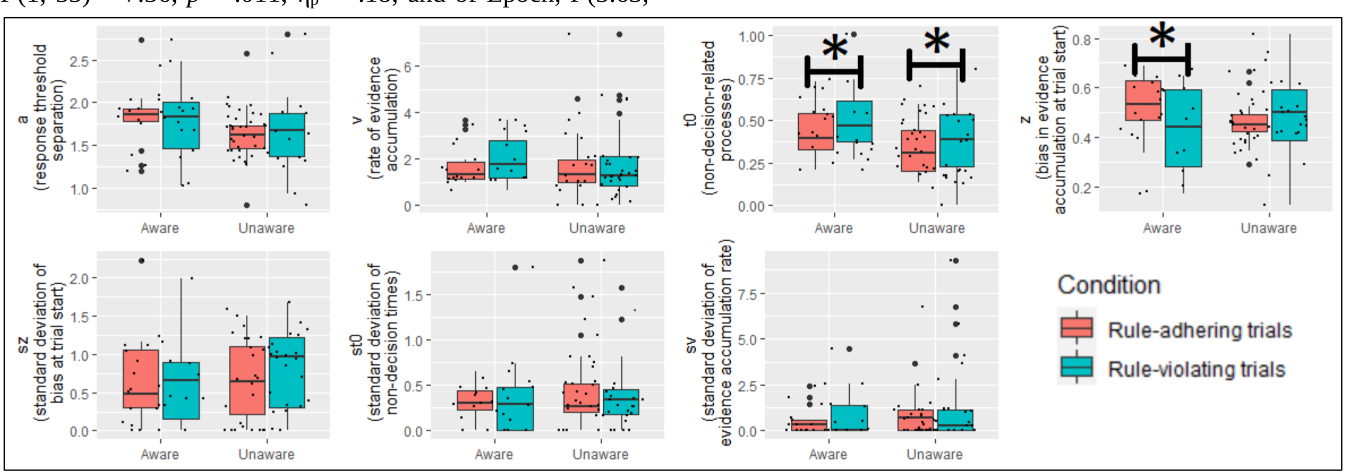

Figure 4. Drift-diffusion model parameters for rule-adhering vs. rule-violating trials, shown separately for rule-aware vs. rule-unaware participants.

2.51, p = .017. By contrast, for rule-unaware participants, bias at the start of the trial did not differ significantly between rule-adhering (M = 0.46, SD = 0.08) and rule-violating (M = 0.51, SD = 0.15) trials, $t(33)$ = 1.36, p = .183.

Discussion

We aimed to explore whether and how awareness of a covert grammatical rule would affect reaction times in a semi-artificial language learning task. For Research Question 1, our linear analysis reproduced prior findings from Batterink et al. (2014) in that both rule-aware and rule-unaware participants showed slow-downs to rule-violating trials, indicative of grammar learning. This aligns with other findings using a similar experimental paradigm (e.g., Williams, 2005; Leung & Williams, 2011) and contradicts the failure to replicate grammar learning effects in rule-unaware participants from Faretta-Stutenberg et al. (2011). More broadly, it suggests that overt instruction might not be necessary for learners to acquire L2 grammar regularities.

However, as discussed above, such a linear analysis that is based on measures of central tendency might not capture subtleties in *how* rule-aware and rule-unaware participants might perform differently in this task, even if the overall slowdown effect is similar (e.g., Balota & Spieler, 1999; Whelan, 2008). For this, we turn to our results from Research Question 2. Our drift-diffusion models suggest that rule learning (as captured by differences between rule-adhering vs. rule-violating trials) affected (a) non-decision-related response times (e.g., tied to factors like motor response speed that lie outside of the process of evidence accumulation) in all participants, and (b) bias in evidence accumulation (i.e., towards or against the correct response, at the beginning of each trial) in rule-aware participants only. The fact that significant differences were found between rule-aware vs. rule-unaware participants in the first place suggests that rule awareness is indeed tied to differences in task performance. This answers in the affirmative the question of *whether* rule-learning makes a difference for grammar learning. We turn now to a discussion of *how* rule-learning makes a difference.

We found that, for both rule-aware and rule-unaware participants, the rule-learning effect was manifested in the t_0 parameter, such that rule-violating trials had longer non-decision times than rule-adhering trials. Recall that the t_0 parameter captures processes that lie outside of evidence accumulation from the presented stimulus (in this case, the English noun). For instance, t_0 could be affected if participants anticipate the correct button press prior to the presentation of the noun. This seems plausible in the case of our experiment design, which (as mentioned above, and following prior work with this semi-artificial language paradigm, e.g., Leung & Williams, 2011; Batterink et al., 2014) assigns a unique button for each of the possible response options in the trial ("near," "far," "living," and "non-living"), making it possible for participants to prepare a living/non-living response immediately upon seeing the *ul*/*gi*/*ro*/*ne* pseudoword. Although this approach would not yield the correct response for the rule-violating trials, it would be sufficient for correctly responding to six-sevenths of the trials (i.e., the rule-adhering trials) and achieving 86% accuracy in the experiment overall. However—and most critically for the purposes of investigating how people learn grammatically-encoded *meanings*—this outcome would merely reflect arbitrary motor preparation responses from our specific idiosyncratic experiment context, devoid of the semantic meaning that is purportedly the target of learning.

For rule-aware participants only, rule learning was also manifested in the z parameter, such that rule-adhering trials showed significantly more bias towards the correct answer relative to rule-violating trials. Recall that the z parameter reflects bias in evidence accumulation at the start of each trial, i.e., if participants have acquired evidence for a living/non-living response before the noun is presented. This is distinct from other possible mechanisms of learning that could be detected by the drift-diffusion model, e.g., accumulating evidence from the noun more slowly or responding more cautiously to rule-violating trials relative to rule-adhering trials. Importantly, the z parameter is distinct from the t_0 parameter in that bias from the z parameter interacts with other decision-related components like the evidence accumulation rate (parameter v) and the response boundary threshold (parameter a), in affecting the reaction time that is ultimately measured for each trial. By contrast, t_0 is "agnostic" to these other components and instead shifts the entire evidence accumulation process to an earlier/later ultimate response time, regardless of the relative timing of its subcomponents. Seen in this way, our findings seem to distinguish between learning that involves higher-order cognitive processes (e.g., pre-activation of the semantics of a noun based on grammatically-encoded meaning) vs. learning that involves lower-level mechanisms (e.g., motor anticipation based on recurring patterns particular to a task context).

Our findings provide evidence that rule-aware and rule-unaware grammar learners engage different mechanisms. However, at this stage, our evidence cannot speak to the exact relationship between implicit and explicit learning. In what has been referred to as the "interface debate" (for a review see Leow, 2015), prior competing models in the field of second language acquisition have argued as to whether explicit L2 learning helps, has no direct relationship with, or (as in Ellis & Sagarra, 2010) can even hinder L2 implicit learning. Hopefully future studies can leverage the power of drift-diffusion modeling to expand on this line of inquiry, e.g., by determining whether the higher-level learning associated with conscious rule-awareness is predicated on lower-level learning tied to motor response prediction in this paradigm. As Rebuschat et al. (2013) write, "one needs to ask what processes contributed to participants suddenly becoming aware of a feature in the first place."

We have identified several future directions for this line of research. First, we have adapted this experiment so that motor response preparation from the hidden semi-artificial language grammar rule is not possible, e.g., by randomizing button assignment on each trial so that the correct response cannot be anticipated prior to noun presentation. This would allow us to test whether implicit grammar learning can occur in regard to word meaning, vs. in regard to lower-level processes related to motor anticipation of idiosyncratic button-pressing sequences in a particular task design. We are currently undertaking data collection for precisely such a study. Because this round of data collection also involves counter-balancing this new, randomized-button trial design with the non-randomized trial design presented in these data, we aim for a controlled comparison across the two conditions as well as a larger dataset to validate the findings presented in the current study. This would also allow us to investigate how differences in prior language experience across experiment participants can affect grammar learning in our experiment paradigm.

Beyond contributing to theoretical debates on implicit/explicit language learning, our findings may be relevant for teaching praxis in illustrating a crucial distinction between L2 grammar learning that is based on understanding of underlying encoded meanings vs. learning that is based on exploiting aspects of the task design that allow learners to produce correct answers without necessarily attending to the target meanings directly (e.g., systematically choosing the verb "are" instead of "is" because the preceding noun ends in *-s*, without understanding that this suffix denotes plurality). We are enthusiastic about the translational potential of drift-diffusion modeling for language teaching praxis, e.g., by suggesting how educators might (at different times and for different short-term teaching purposes) intentionally exploit vs. avoid features of classroom task design that invoke the kind of low-level learning processes we describe here.

Acknowledgements

We wish to thank Krishna Bhatia, Will Camacho, Hannah Chaddah, Gayatri Chavan, Sarah Hassan, Amber Lewis, and Victoria Ogunniyi for their contributions to data collection for this project. We also thank Dr. Laura Batterink for her generosity in sharing materials and ideas relevant to this project. We thank Joshua Buffington, Dr. Jennifer Cabrelli, Dr. Philip Hamrick, Dr. Irene Finestrat, Dr. David Miller, and Dr. Kim Potowski their for additional insights. This project was funded by a Dingwall Foundation Fellowship in the Neural Foundations of Language, by a National Science Foundation Doctoral Dissertation Improvement Award (#1941189), by an American Philosophical Society John Hope Franklin Fellowship, and by a Travel and Research Grant from the Institute for Citizens & Scholars (formerly Woodrow Wilson National Fellowship Foundation).

References

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128*, 32-55.

Batterink, L. J., Oudiette, D., Reber, J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169-179.

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory, 22*(11), 544-556.

Ellis, N., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition, 32*(4), 553-580.

Faretta-Stutenberg, M., & Morgan-Short, K. (2011). Learning without awareness reconsidered: A replication of Williams (2005). In *Selected proceedings of the 2010 Second Language Research Forum*. Somerville, MA: Cascadilla Proceedings Project.

Grey, S., Williams, J. N., & Rebuschat, (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, *36*(4), 611-645.

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.5.0. https://CRAN.R-project.org/package=emmeans

Leow, R. (2015). *Explicit learning in the L2 classroom*. New York: Routledge.

Leung, J. H., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually derived meanings. *Studies in Second Lg. Acquisition*, *33*(1), 33-55.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940-967.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psych. Science*, *9*(5), 347–356.

Rebuschat, P., Hamrick, P., Sachs, R., Riestenberg, K., & Ziegler, N. (2013). Implicit and explicit knowledge of form-meaning connections: Evidence from subjective measures of awareness. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition*. Honolulu: University of Hawaii Press.

Rose, M., Haider, H., & Büchel, C. (2010). The emergence of explicit memory during learning. *Cerebral Cortex*, *20*(12), 2787-2797.

Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonom. Bulletin & Rev.*, *12*(2), 195-223.

Singmann, H. (2020). *Reanalysis of Ratcliff and Rouder (1998) with Diffusion Model and LBA*. Online R vignette. https://cran.rstudio.com/web/packages/rtdists/vignettes/reanalysis_rr98.html#diffusion-model-analysis

Singmann, H., Brown, S., Gretton, M., & Heathcote, A. (2020). *rtdists: Response Time Distributions*. R package v0.11-2. https://CRAN.R-project.org/package=rtdists

Vinding, M., Lindeløv, J. K., Xiao, Y., Chan, R. C., & Sørensen, T. A. (2018). Volition in prospective memory: Evidence against differences in recalling free and fixed delayed intentions. https://psyarxiv.com/hsrbt/

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 475-482.

Williams, J. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.