LANGUAGE LEARNING AND DEVELOPMENT
https://doi.org/10.1080/15475441.2020.1802279

In Support of Phonological Bias in Implicit Learning

Katya Pertsova and Misha Becker

Department of Linguistics, University of North Carolina, Chapel Hill, North Carolina, USA

ABSTRACT

This paper explores the hypothesis that children pay more attention to phonological cues than semantic cues when acquiring grammatical patterns. In a series of artificial allomorphy learning experiments with adults and children we find support for this hypothesis but only for those learners who do not show clear signs of explicit learning. In particular, learners who cannot verbalize a correct rule after the experiment nevertheless perform above chance on phonological patterns, but not the semantic ones. On the other hand, learners, particularly adults, are more likely to (explicitly) discover and successfully verbalize a rule based on a salient feature of animacy compared to a phonological feature based on the number of syllables. We discuss implications of these results in the context of a distinction between explicit and implicit learning mechanisms and how this distinction relates to the study of phonological bias.

Introduction

When learning a language one has to learn the distribution of various grammatical elements. This presents certain challenges because such a distribution can depend on many factors, some of them phonological, others semantic, and yet others morphosyntactic. For example, the suffixes on finite verbs in Hungarian vary based on the person and number of the subject (morphosyntactic), but also based on the definiteness of the object (semantic), and the phonological shape of the stem (Singer, 1982). In Tsez, membership in a noun-class depends on the animacy and sex of the referent (semantic), but also on some phonological properties of the stems (Gagliardi, 2012). Do learners have any biases that lead them to pay more attention to specific types of cues when learning the distribution of grammatical markers or form-meaning mappings more generally?

There are reasons to think that such biases indeed exist: language acquisition studies show that children appear to privilege phonological cues over semantic ones as discussed later in this paper. Additionally, it is well-known that many grammatical classes of words have arbitrary phonological correlates that speakers are sensitive to. For example, monosyllabic German nouns beginning with [kn], [dr], or [tr] tend to be masculine (Szagun et al., 2007; Zubin & Köpcke, 1984); verbs in English tend to have fewer syllables than nouns and differ in their stress-pattern (Kelly, 1992); Spanish verbs of second conjugation tend to have mid-vowels in the stem (Albright, 2008). Why should semantic or grammatical classes have distinct, seemingly arbitrary phonological properties? One common explanation is that these phonological distinctions among different classes may be a result of a series of independent historical changes which overtime become opaque. A phonological bias could be another

CONTACT Katya Pertsova pertsova@unc.edu Department of Linguistics, University of North Carolina at Chapel Hill, NC
This work would not have been possible without the help from our research assistants Simon Wolf, Brent Eisenbarth, and Evan Emhart. We would also like to thank the following people and audiences for helpful comments and discussion: Jennifer Culbertson, Elliott Moreton, Jennifer Smith, Brian Hsu, members of audience at the LSA 2018 Annual Meeting, associate editor Daniel Swingley, and the anonymous reviewers. For their help with recruiting children participants we thank the staff at the Museum of Life and Science in Durham and Lateasha Camp, the director of an after-school program in Carrboro, NC.

explanation for why such arbitrary phonological patterns develop and persist – if learners are highly attuned to phonological cues early on and such cues are useful in categorizing words into distinct grammatical classes, then it would be likely for such cues to be preserved or even enhanced. Some evidence for this comes from artificial language learning (A.L.L.) studies. For example, Lany and Saffran (2010) showed that toddlers were able to learn a semantically based class distinction when it was accompanied by phonological regularities that they were previously trained on, but a control group could not learn the same semantically based class without additional phonological cues.

In general, we define phonology-over-semantics bias here as a predisposition of the learning system to learn the phonological patterns more easily or better in some way than semantic ones. We make no claims about which aspect of the learning process is responsible for the bias: a systematic learning advantage could result from the differences in the perception of cues from different modalities, or the differences in attention, or the differences in the relative weight of cues during learning, or something else. We can only observe behavior to see whether phonological cues are systematically privileged over semantic ones. If we find evidence for this from multiple sources, such as natural language acquisition, artificial language learning, and typology, then we can be more certain that such a bias is real.

In this paper, we test the existence of the phonology-over-semantics bias by examining how an artificial language pattern involving phonologically- and semantically-conditioned allomorphy is learned by children and adults. We find some evidence for the phonology bias, but only in those participants who were not able to state after the experiment what pattern/rule they were learning. Participants who do correctly verbalize the pattern post-experiment, do well on whatever pattern they are exposed to. For a highly salient feature of animacy, more participants are able to learn the rule explicitly. We connect these results to the literature on explicit/implicit learning in the next section. Additionally, the ability to correctly verbalize a rule correlates with age: adults are much more likely to do so compared to children. Our findings provide a potential explanation of a previous similar study (Culbertson et al., 2019) which reported a phonological bias in children, but a semantic bias in adults. We find that the apparent semantic bias in adults disappears once we exclude explicit learners. Adults who are not aware of the correct rule show the same pattern of behavior as children; namely, they are above chance on the phonological pattern and at chance on the semantic one. Our study raises a question of whether natural semantic features are acquired by children using a more explicit learning strategy compared to the phonological ones. This could in principle be one of the explanations for the existence of the phonological bias, since explicit learning and the declarative memory known to underly explicit knowledge are often said to develop later compared to the implicit learning (Finn et al., 2016; Nelson, 1995).

The paper is organized as follows. First, we go over the background information motivating our study, then we present the experiments and the results, and, finally, we discuss our findings and their implications.

Background

L1 acquisition

Past research on the acquisition of noun class morphology shows a general tendency for children to rely more on phonological and/or morphosyntactic cues to noun categorization than semantic cues (Hebrew, Berman (1985); German, Mills (1986); Czech, Henzl (1975); French, Karmiloff-Smith (1979), Spanish, Pérez-Pereira (1991), Tsez, Gagliardi (2012)). For example, in an experimental study by Pérez-Pereira (1991) Spanish-learning children were tested on their assignment of novel nouns to gender classes based on cues that are predictive of gender in Spanish. The nouns labeled made-up creatures that were displayed with either feminine or masculine gender characteristics (e.g., having long vs. short hair, wearing high heels vs. a man's hat, etc.). Gender was also cued by the phonological form of the word (masculine nouns ending in – *o*, feminine in – *a*) and/or by the article (masculine *el* and feminine *la*). Simplifying somewhat, the main results were that children appeared to place greater importance on the phonological cue of the suffix or the grammatical cue of the article than the natural gender (semantic cue). Additionally, there were no major differences across the age groups (4 to 11 years), though younger children were somewhat more influenced by the phonological cue of the suffix than the syntactic cue of the article when the two were in conflict, compared to older children.

Similar findings are reported in a more recent study on learning Tsez noun-classes (Gagliardi, 2012; Gagliardi & Lidz, 2014). In this study, children and adults were given labels for unfamiliar objects and individuals to test which noun class they would assign them to. In some cases, noun-class cues in a novel word were in conflict: the semantic cue (e.g., feminine referent) was predictive of one class, but the phonological cue (e.g., initial consonant is *r*) was predictive of a different class. Crucially, in adult Tsez semantic cues are statistically more reliable and always trump phonological cues. However, Gagliardi found that children (younger children in particular) disproportionately relied on the phonological cue when inflecting novel nouns, while adults favored the semantic cue.

Even languages that are known to have extensive noun classes with a semantic basis show that children may be encoding noun-class morphology by relying on phonological features. For example, in reviewing the literature on child acquisition of Bantu noun classes Demuth (2003) notes that although children tend to spontaneously produce correct noun prefixes early on, overgeneralization errors are possible with novel nouns. In particular, one experimental study by Kunene (1979) reports that 4–6-year-old Swati speakers had trouble producing plurals of certain novel nouns. Demuth describes their overgeneralization errors as "systematically phonological and morphological rather than semantic." This is consistent with Demuth's claim that Bantu noun classes are, generally speaking, "grammatically productive" but only moderately "semantically productive." That is, synchronic class inclusion (e.g., of loan words) is determined more by phonological features than semantic features. For example, loan words from English or Afrikaans into Sesotho whose initial consonant is used in a noun-class prefix are assigned to that class as long as the first vowel is also compatible with that prefix (e.g., *mo-chini* 'machine' is Class 3/4 because 'mo-' is the prefix for this class, or *se-tulo* 'chair' from Afrikaans *stoel* is class 7/8 because 'se-' is the prefix for that class). Nouns that cannot be classified this way based on their word-initial phonology like *ofisi* "office" are put into the "default" class 9/10 (Demuth, 2000).

The data briefly referenced in this section is suggestive of a phonological bias in first language acquisition. Next, we turn to the question whether such a bias can be detected after learning an artificial language pattern in a lab. However, before discussing A.L.L. experiments, we would like to make an important distinction between implicit vs. explicit learning that such experiments elicit.

Explicit vs. implicit learning

The artificial language learning paradigm is a popular and powerful tool for studying learning biases because it allows experimenters to tightly control relevant and irrelevant factors affecting learning. The flip-side of this advantage is that learning in the lab is very different from learning "in the wild," and so this methodology is often criticized as possibly not being informative about the process of naturalistic language acquisition. In particular, it could be that A.L.L. engages explicit conscious reasoning and problem-solving, rather than the implicit acquisition process most linguistic researchers are interested in. Nevertheless, many A.L.L. researchers, while acknowledging the tradeoffs between ecological validity and experimental control in A.L.L. studies, rarely try to tease apart implicit from explicit learners in their experiments.

The distinction between implicit and explicit learning mechanisms (with some correspondence to different types of memory, declarative and procedural) has been extensively studied in the fields of cognitive psychology with respect to category-learning (Ashby et al., 1998; Maddox & Ashby, 2004; Smith et al., 2015), motor-sequence learning (Cherry & Stadler, 1995; Howard et al., 2004), and in the field of second language acquisition (R. Ellis, 2008; Lichtman, 2012). The existence of these two distinct learning mechanisms available to humans is also supported by neurophysiological studies

(Gureckis et al., 2011; Morgan-Short et al., 2012; Ullman, 2004; Wong et al., 2013). The explicit system is hypothesized to be effortful, and involve conscious deliberation and hypothesis testing which can lead to abrupt learning, while the implicit system is hypothesized to be effortless, automatic, and involve unconscious statistical learning leading to gradual improvement over time.

An important lesson from psychological research on the explicit/implicit distinction is that the two mechanisms can differ in learning biases. For example, a number of well-known experiments show that a pattern that has an exclusive-OR structure is learned better than the one that has a prototype structure when experimental conditions maximally encourage explicit learning (Nosofsky et al., 1994; Shepard et al., 1961), but this advantage goes away or is even reversed under conditions that favor implicit learning (Kurtz et al., 2013; Love, 2002). An example of a prototype-structured pattern defined over three binary features, shape (triangle/square), color (black/white), and size (big/small), is "anything that is one or zero features away from small white square" (i.e., small white square, small white triangle, large white square, small black square), and an example of an exclusive-OR pattern is "either white triangles or black squares." The finding that learning mode can be associated with different biases underscores the importance of distinguishing between implicit and explicit learning in the lab and being aware of how various experimental choices influence learning strategy. Factors that are known to facilitate explicit over implicit learning are experience and formal education (Minda et al., 2008), use of perceptually salient, discrete, separable, and verbalizable features (Kurtz et al., 2013; Nosofsky & Palmeri, 1996), receiving instructions to seek a rule, receiving feedback after every trial (Kurtz et al., 2013), and blocked rather than random presentation of stimuli during training (Noh et al., 2016).

While all learners have access to the explicit and the implicit systems, participants in a lab experiment differ with respect to which of these two systems they rely on more heavily. That is, some people approach an experiment as a puzzle to be solved by trying out different hypotheses, others tend to simply rely on their gut feeling, and, yet, others try to memorize the stimuli. Additionally, different experimental designs and stimuli encourage different strategies. From previous studies we know that the use of different strategies is associated with different behavioral "signatures" that can be used to diagnose what subjects were doing in an experiment. For example, the most prominent sign of implicit learning discussed in Reber (1967)'s seminal paper "Implicit Learning of Artificial Grammars" is gradual, rather than abrupt, improvement in performance coupled with inability of subjects to verbalize any rules or patterns that characterized the training stimuli. In the psychological Artificial Grammar Learning experiments pioneered by Reber, the stimuli are strings of letters or symbols generated by some finite state machine. This experimental paradigm sharply contrasts with many psychological experiments on category learning from the 50s and 60s in which the stimuli and the pattern being learned were simple enough for subjects to reason about explicitly (e.g., "red striped circles or white triangles"). The A.L.L. studies typically used in linguistic research are different in a number of ways from the Reber-style artificial grammar experiments as well as from artificial category-learning experiments. Moreton and Pertsova (2016) investigated whether linguistic A.L.L. experiments involving learning of words that follow a simple phonotactic pattern or categorizing such words as masculine vs. feminine, produce signs of both explicit and implicit learning. They found that explicit learning was particularly prevalent under experimental conditions that involved telling subjects to seek a rule, giving feedback after every trial in a forced choice task, or contrasting members of different categories on each trial (e.g., "which of these two words is feminine and which is masculine?"). Experimental conditions that did not include the above-mentioned characteristics led to a smaller (but still significant) number of people who showed signs of explicit rather than implicit learning. These signs were: self-report of using a rule or a strategy, ability to correctly verbalize a rule post-experiment, bimodal performance in the testing phase with those who stated a correct rule achieving near perfect performance and others staying at chance, abrupt improvement and acceleration of response times during training (in conditions that allowed the measurement of training performance). In general, the subjective measures were correlated with the objective ones: that is, those subjects who reported finding a rule and were able to state it also performed near ceiling and showed abrupt improvement during training (which presumably indicates that they have hit on the right rule). Moreton and Pertsova also found that some subjects use more than one learning strategy. People sometimes reported deliberately seeking a rule, failing to find one, and resorting to intuition; or the reverse: relying on intuitive guessing at first, but then suddenly recognizing a pattern and consciously using that knowledge in the testing. The extent to which the two learning mechanisms interact with each other is somewhat controversial (for different views on this issue in L2 acquisition see Ellis (2008), Paradis (2009), Poldrack and Packard (2003)), but there are reasons to think that they are not mutually exclusive. One type of evidence suggesting that there is some communication between the two systems comes from the literature on the "aha-insights" (e.g., Gilhooly (2016), Hélie and Sun (2012)) which are typically seen as products of implicit reasoning (because they come during a period of rest when one is not consciously working on a problem) becoming available to the explicit system. Aha-moments are also reported to happen after a period of intense conscious deliberation on a complex problem. In general, we remain agnostic about the degree and nature of the interactions between the two learning systems.

Experimental studies relevant to the phonology-bias hypothesis

In this section, we review A.L.L. studies which bear on the question of a phonological bias. Culbertson et al. (2017) explore a learning situation involving conflicting semantic and phonological cues in an A.L.L. paradigm. In their first study with adults they find that early exposure to one of the cues (either semantic or phonological) has an effect of biasing learners to later pay more attention to that cue. They call this effect "staging," and offer it as a possible explanation for the phonological bias in L1 since access to phonological structure arguably precedes access to semantic structure in natural language acquisition (infants have access to some aspects of phonological structure before they know what many words mean). In the absence of staging, adults' preferences were dictated by feature salience. When both a phonological and a semantic feature explained the data equally well (the two features were correlated or "aligned" during training), adults were more likely to learn a phonological rule if the phonological feature had higher salience and the semantic feature had lower salience, and vice versa. Phonological salience was defined as having more phonological segments (having a suffix vs. having a circumfix) and semantic salience was based on typological frequencies of features *animacy, shape, flexibility* as features motivating inflectional class membership (listed in decreasing order of frequency/salience). It is important to note that the two types of salience metrics above are not comparable to each other, and in general there is no independent way of establishing whether a particular phonological feature is more salient than a particular semantic feature aside from observing the difficulty of learning each feature in isolation. When a particular phonological feature is more difficult to learn than a particular semantic feature, this could be due to a general phonological bias or to a feature-specific bias (e.g., feature salience). What the authors show is that there is no across-the-board general advantage for phonological features – it is possible for a semantic feature to be learned better than a phonological one under some conditions. We will see below that in addition to salience and bias, the mode of learning (explicit vs. implicit) can also affect learnability of a pattern.

In another series of experiments Culbertson et al. (2019) tested children (ages 6–7) and adults using a different set of features, the semantic feature being planets vs. aliens, and the phonological feature being words of shape CiCi vs. CaCa (where the second syllable was fixed for each vowel). The training consisted of 96 trials during which children were shown singular and plural forms of novel nouns and asked to repeat them. The plural marker was determined by the semantic and/or the phonological property depending on the condition. In the experiment in which the two types of features were independent (experiment 1), testing consisted of 48 production trials (repetitions of 8 trained and 4 new "untrained" nouns) during which children were asked to produce a plural given a singular and were given correct feedback. In the experiment in which the two features were aligned in the training (experiment 2), there were two production phases: the first one used the 8 trained nouns, included correct feedback, and lasted 24 trials. It was followed by another 48 trial production phase (repetitions of 8 trained and 4 untrained nouns) without feedback which was used to evaluate performance. In experiment 1, children performed equally well in the phonological and semantic conditions, correctly producing plurals on about 70% of the 48 test trials. However, in experiment 2 significantly more children made their decisions based on the phonological cue than the semantic cue, while nearly all adults relied on the semantic cue. This result suggests that children pay more attention to the phonological cues, while adults pay more attention to the semantic ones. The authors were also able to replicate the staging effect with children. They interpreted their results as showing that when two features have equal salience (because they are learned equally well in isolation), a phonological bias can still show up as a preference for phonological cues in a situation of conflict or competition.

However, given possible influence of learning mode these results do not definitively show that the two features had equal salience. It could be that there is a phonology over semantics advantage for implicit learning and the reverse advantage for explicit learning given that most semantic features are more verbalizable than the phonological ones. Likewise it could be that feature salience is different for the two learning modes. If this is the case, it is conceivable that when the two features were tested in isolation from each other, children did equally well on both of them, but they used different learning strategies, or different proportions of children in each condition used one or the other strategy or a different mix of the two strategies. If choice of strategy also affects learning, then it is hard to infer from independent conditions which feature had an advantage without knowing which participants used explicit vs. implicit learning. Interestingly, while in experiment 1 average performance on previously seen trials was around 65% correct, performance on such trials in experiment 2 was 85% correct. One possibility is that in experiment 2 when the two features were in competition, the first production phase in which the 8 trained nouns were presented with feedback could have facilitated transfer of the implicit knowledge to the explicit system by allowing subjects to test their intuitions against feedback and/or to better memorize trained nouns. This knowledge could then have been used in the following phase on which participants were evaluated. (In comparison, in experiment 1 performance was evaluated based on a single production phase which included feedback and novel items.) If the second experiment facilitated more explicit learning, one would expect general improvement in performance. On the other hand, there were no significant jumps in the overall learning curves for trials with feedback in experiments 1 and 2. So the improvement in children's performance on experiment 2 could have also stemmed from the aligned cues making learning easier, or from the combination of aligned cues and the additional training with feedback.

In any case the data from experiment 2 (presented in Figure 1 in a way that allows us to compare each child's performance on aligned and conflicting trials) shows that there was a bimodal distribution of performance suggestive of explicit learning. Namely, most children learned *some* rule well since most of them (14/20) performed above 90% correct on aligned trials and at the same time did well on conflicting trials, choosing either the semantic rule (5 children in the top right corner of the graph), or the phonological rule (9 children in the bottom-right corner of the graph). These children might be the ones who ended up with a more or less explicit knowledge of the rule they were learning or with a very strong implicit intuition. Of interest are the remaining six children. What is striking about them is that although they do not appear to have learned any rule well, they are not performing at chance as a group. Instead, they have a slight preference toward the phonological rule. Testing their mean proportion of "correct" responses on all trials (where correct is taken to be a phonological choice), they perform on average at 63% correct (median = 76%) which is significantly greater than chance according to Wilcoxon one-sample signed rank test ($V = 0$, $p = .01$). Thus, one possibility is that most children in this experiment were at first learning implicitly and, therefore, were biased toward the phonological rule. Some of these children were then able to either strengthen their implicit knowledge of the pattern or to transfer their implicit knowledge to the explicit system during the production phase with feedback. Others remained or reached a slightly above chance level of performance on the phonological cue. Yet others, have discovered a semantic rule, possibly because they were learning explicitly. In contrast, adults, who are better at explicit learning, were more likely to have been learning explicitly from the very beginning and hence were overall biased toward the more verbalizable semantic rule. Of course, there is no way to test this hypothesis without additional measures of implicit/explicit learning.[1] However, this

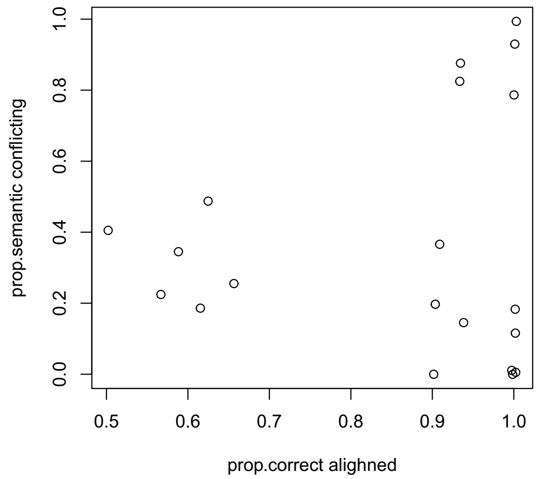

Figure 1. Children's performance (individual means) on aligned and conflicting trials in Experiment 2 from Culbertson et al. (2019). Higher values on the y-axis indicate semantic preference, while lower values indicate phonological preference. Raw data was obtained from the authors' repository at https://datashare.is.ed.ac.uk/handle/10283/3438.

pattern of results – imperfect, but above chance performance of learners on phonological patterns – is very similar to what we find in analogous experiments described in this paper. To foreshadow our results, even adult learners trained on two competing aligned cues end up showing a slight phonological preference if they fail to find and state any rules. On the other hand, semantic features seem to be learned either very well (with explicit awareness) or else not at all, as discussed below.

A number of A.L.L. experiments tested learning of semantic features that are relevant for morphology. For example, Ferman and Karni (2010) tested Hebrew learners on a pattern in which verbs took different suffixes depending on the animacy of the subject. Children (age 8 and 12) and adults were trained for 10 consecutive daily one-hour sessions. Despite such a lengthy training only 12 year-olds and adults were able to generalize the pattern to new items, and all of those who generalized reported explicit awareness of the animacy rule with an abrupt increase in performance following the report of the rule. The 8 year olds, on the other hand, performed at chance. In another study Brown et al. (2018) trained adults and 6 year-olds on a language that used a mix of English and artificial words and instantiated a pattern in which one particle was restricted to animals and another – to vehicles. They found that after four separate training sessions carried out in the span of 4 to 8 days, the only participants who showed above chance performance on new items were those who also correctly verbalized the pattern in the final interview. These were nearly all adults and about half of the children.

Likewise, Moreton and Pertsova (2016) failed to find any robust signs of implicit semantic learning in an experiment in which they tested learning of grammatical gender conditioned by either semantic or phonological cues. The three semantic cues were edible vs. inedible, long vs. round/compact shape, and large vs. small. Six phonological features included three features that the authors deemed to be verbalizable – disyllabic vs. trisyllabic, stress on first vs. second syllable, and all consonants are same vs. different – and three phonological features that the authors deemed less verbalizable, namely stressed vowel is back vs. front, all consonants are labial vs. coronal, and all consonants are stops vs. fricatives. Participants were trained either in an implicit-promoting condition in which they saw a sequence of feminine (or masculine) picture-word pairs, or explicit-promoting condition in which they saw two word-picture pairs and were asked to decide which of them was masculine/feminine and were given feedback on each trial (for exact details and procedure refer to Moreton and Pertsova (2016)).

[1]We cannot rely on the learning curves provided in the paper as a measure of explicit learning because these curves only show the trajectory of responses after most learning has already taken place during the initial 96 trials without feedback.

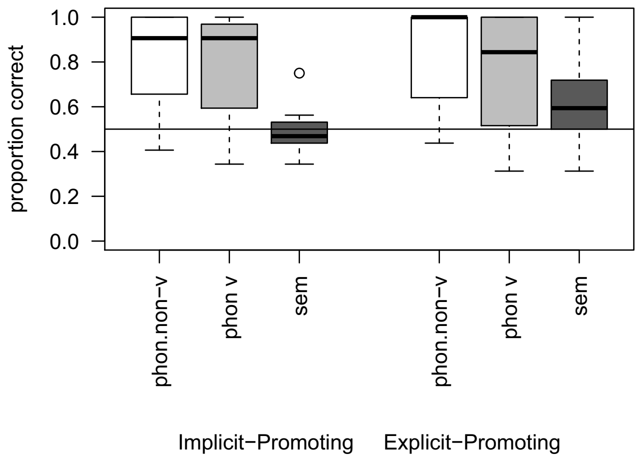

Figure 2. Boxplots for performance on 32 test trials in Experiment 1 from Moreton and Pertsova (2016) for three groups of features: phonological verbalizable, phonological non-verbalizable, and semantic.

Consider the boxplot in Figure 2 that shows subject mean proportion of correct responses for the three groups of features further divided based on the type of training: forced choice with feedback (explicit-promoting) vs. positive examples of one category with no feedback (implicit-promoting). Overall, subjects trained on the phonological cues did significantly better than those trained on the semantic cues regardless of training regime. The only successful semantic learners (N = 5) were in the explicit-promoting condition and were able to state a rule after the experiment. All other semantic learners performed at chance.

To summarize, in all experiments discussed in this section that included a post-experimental debriefing, the only participants who did well on learning a semantic pattern were those who could also explicitly state the pattern afterward. On the other hand, phonological learning is frequently observed in experimental studies even in participants who do not show awareness of the pattern after the experiment (Chan & Leung, 2014; Dell et al., 2000; Moreton & Pertsova, 2016), in infants (Cristia' & Seidl, 2008; Saffran & Thiessen, 2003), and older children (Culbertson et al., 2019; Plante et al., 2010). These results provisionally support a phonology over semantics bias in implicit learning.

Experiments

Two sets of experiments testing the existence of the phonology bias are described below. Experiment 1 was run over the web with adults. This experiment is very similar to Culbertson et al. (2017)'s first experiment, except we used somewhat different features and tested the relative contributions of semantic vs. phonological cues when they were both independent and correlated, while Culbertson et al. (2017) only used the correlated condition. The main hypothesis explored in Experiment 1 was existence of a phonology-over-semantics bias. We hypothesized that the phonological rule would be learned better and faster than the semantic one both when the two features were independent and when they were correlated. Post-hoc we also tested for the interaction between ability to state a correct rule after the experiment and the proportion of correct responses in the testing phase.

The main sign of explicit learning we relied on was ability to state a correct rule after the experiment and bimodality of performance on test trials across subjects. Since our experiments did not include feedback, other measures, such as abrupt improvement in performance and reaction times, were not available to us. It should be noted that subjective reports are not always reliable, but several studies report reasonable correlations between verbal reports and other measures of implicit learning (Moreton & Pertsova, 2016; Ziori & Dienes, 2006). To the extent that such measures provide us with additional information that tends to correlate with specific biases observed in learning different types of pattern, we find them useful. We also recognize that we cannot distinguish between various sub-populations of learners that might be using both implicit and explicit strategies to different degrees, or in different temporal orders. In this paper we will rely on the difference between "correct staters" (those people who state a correct rule post-experiment) and everyone else as a proxy for the difference between those who *successfully used explicit learning at least some of the time during training and/or testing* and everyone else. The latter group will necessarily be quite heterogeneous. It will include subjects who used explicit learning unsuccessfully, those who used implicit learning, or some combination of explicit and implicit learning, as well as those who just guessed randomly. However, it is crucial to note that subjects who use exclusively an explicit strategy, but fail to find a rule, as well as subjects who truly guess randomly are expected to perform at chance. Thus, any above chance performance of non-correct-staters must be due to implicit or incidental learning happening below the level of awareness (possibly even among those who thought they were using a rule).

In the second experiment we applied a similar method with some changes to children (ages 5–11) and a control group of adults who were run in person rather than over the web. In this experiment, we manipulated an additional parameter during training, namely blocked vs. random presentation of trials. The former encourages explicit reasoning, while the latter encourages implicit reasoning (Noh et al., 2016). Both experiments included post-experimental questions querying participants about strategies they used in approaching the task and the rules/generalizations that they found.

The linguistic phenomenon that our artificial languages are modeled on is suppletive allomorphy, a situation in which a morpheme has several allomorphs that are not phonologically related and cannot be derived from the same underlying form, but whose distribution can be phonologically or semantically determined. An example of semantically conditioned suppletive allomorphy is Persian plurals: nouns referring to humans take the plural allomorph -/an/, while all other nouns take the allomorph -/ha/ (e.g., the noun *sar* "head" is pluralized as *sar-ha* when it refers to a body part and as *sar-an* when it refers to a leader or head of a group). An example of phonologically conditioned suppletive allomorphy occurs in Martuthunira (an Australian language), where locative case allomorphs are – *ngka* following disyllabic stems and – *la* following trisyllabic stems (Haspelmath & Sims, 2013).

Experiment 1

This experiment tested adults' ability to learn plural allomorphy that depended on either a phonological property of the stem, a semantic property, or ambiguously on both properties. The allomorphy was instantiated in an artificial language consisting of singular and plural nouns. If the phonological bias exists for all learners and is absolute (phonological cues always trump semantic cues), we would expect that the phonological rule would be easier to learn than the semantic rule, and that when the pattern is ambiguous, subjects interpret it as phonologically-conditioned more often than semantically-conditioned. Another possibility is that subjects learn the phonological and semantic rules equally well when the two are independent of each other, but when they are correlated, they prefer the phonological rule. This result would be similar to what Culbertson et al. (2019) report for children (see the discussion in the previous section). Such a result would then require an explanation of why this bias does not show up when the features are learned independently.

Stimuli and design

The stimuli were nonsense words obeying English phonotactics paired with pictures. The words differed in the number of syllables, one vs. two (a phonological feature) and were paired with one of two sets of pictures. In the first set, half of the pictures depicted animate beings and the other half various inanimate objects (semantic feature 1). In the second set, half of the pictures depicted long and skinny objects and the other half – round and compact objects (semantic feature 2). Each participant was trained using either the first or the second set of pictures, but not both: thus, for some people the two possible conditioning factors for plural allomorphy were number of syllables and animacy, and for others they were number of syllables and shape. Based on previous findings (Culbertson et al., 2017), we hypothesized that animacy would be a more salient semantic feature compared to shape. It should be noted that all three types of features are attested as possible triggers of allomorphy in real languages.[2] The full list of the word stimuli can be found in the Appendix and pictures as well as experimental and analysis files are deposited in the Carolina Digital Repository (at https://doi.org/10.17615/9eqt-m345). For half of the words the plural was marked via the prefix *ro-* and for the other half via the suffix *–ek*. We used a prefix and a suffix to emphasize the difference between the two classes of nouns at the expense of some linguistic naturalness.

There were three main conditions in the experiment (phonological, semantic, and mixed) and within each condition, two subconditions (based on whether the semantic feature was animacy or shape). In the phonological conditions, the choice of the plural allomorph was determined by the phonological feature, and the semantic feature was irrelevant (equally distributed across the two values of the phonological feature). In the semantic conditions, the determining factor was either animacy or shape and the phonological feature was irrelevant. In the mixed conditions, it was ambiguous whether the choice of the allomorph depended on phonology or semantics as only two out of four possible paradigm cells were presented (with the remaining two cells reserved for testing). This is illustrated with example paradigms in Table 1, which shows six total sub-conditions of the experiment.

The example paradigm in Table 1 for the mixed condition with the semantic feature being animacy shows that the suffix *–ek* occurs with monosyllabic animate nouns, while the prefix *ro-* occurs with disyllabic inanimate nouns. Thus, a participant exposed to this incomplete paradigm can either learn a rule based on animacy or based on syllable number (or possibly on both). Which of these rules will be preferred should become apparent based on nouns for which the two rules make opposite predictions (e.g., a monosyllabic inanimate noun). Such nouns appear in the testing phase.

Procedure

During training, the task was to listen to 40 pairs of singular and plural forms (10 from each paradigm cell for the phonology and semantics conditions, or 20 from each of the two cells in the mixed conditions) and to look at the corresponding pictures. The pairs were presented in random order on

Table 1. Example plural paradigms for six conditions.

Phonology			Semantics	
	anim.	inanim.	anim.	inanim.
1 syll	*pof-**ek*** “bees”	*piz-**ek*** “beds”	*pof-**ek*** “bees”	***ro**-piz* “beds”
2 syll	***ro**-bezosh* “cats”	***ro**-sapik* “bikes”	*bezosh-**ek*** “cats”	***ro**-sapik* “bikes”
	long	compact	long	compact
1 syll	*shik-**ek*** “brushes”	*dez-**ek*** “clocks”	*shik-**ek*** “brushes”	***ro**-dez* “clocks”
2 syll	***ro**-pefid* “forks”	***ro**-dezus* “balls”	*pefid-**ek*** “forks”	***ro**-dezus* “balls”
Mixed				
	anim.	inanim.		
1 syll	*pof-**ek*** “bees”	**testing**		
2 syll	**testing**	***ro**-sapik* “bikes”		
	long	compact		
1 syll	**testing**	***ro**-dez* “clocks”		
2 syll	*pefid-**ek*** “forks”	**testing**		

[2]The Martuthunira pattern mentioned earlier provides an example of syllable-sensitive allomorphy. Other examples include Eastern Armenian, in which plural allomorph is *-er* for monosyllabic bases and *-ner* for polysyllabic bases (Haspelmath & Sims, 2013, p. 160), and Estonian genitive plural allomorphs, which are analyzed by Kager (1996) as an interaction between metrical structure and morphology. Both shape and animacy commonly condition assignment of lexemes to noun classes (Corbett, 1991; Seifart, 2005).

a computer screen auditorily and visually with the singular and plural versions side by side (an example trial is shown in Figure 3). The words were played as soon as the pictures appeared. Participants clicked the "continue" button to proceed to the next trial whenever they were ready. The instructions told participants that they would hear singular and plural forms of words in a made-up language, and that their task was to listen to the words and look at the pictures. They were also told that after the training is over they would be shown more words in the same language and asked to choose one of two possible plural forms for each word.

In the testing phase, 24 new items (6 from each paradigm cell) were presented in random order. A participant heard the singular form and saw two pictures side-by-side showing the singular-plural pair. Below, two possible plural forms appeared, one with *ro-* and one with *-ek*. Participants were instructed to pick "the best" option. The pairing of words and pictures, the order of trials, the assignment of allomorphs to specific feature values, and assignment of stimuli to training vs. testing were all randomized for each subject. A screenshot from the testing phase appears in Figure 4.

When the testing phase was over, participants filled out a short debriefing questionnaire in which among other things they were asked to check a box for the strategy that they used during the experiment (intuition vs. a rule) and whether they could state any rules or generalizations that helped them make decisions in the testing phase. Participants' responses on the last question were coded by the first author as correct or incorrect. A response counted as correct if it indicated that the choice in the testing phase was made based on syllable or sound/letter count or word-length for the phonological condition and based on animacy or shape for the semantic conditions. Subjects whose responses were coded as correct will be referred to as *correct staters* in the subsequent discussion. Participants were native speakers of English[3] run over the web, using Amazon's Mechanical Turk as the recruitment

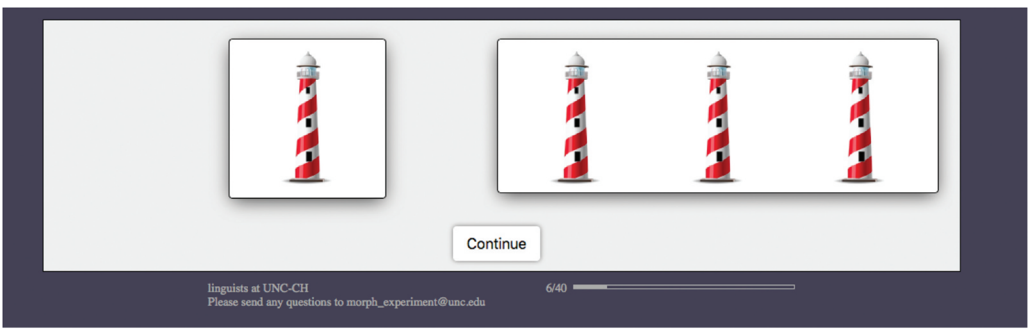

Figure 3. Example training trial (words were presented auditorily). Customized version of Experigen software (Becker & Levine, 2013) was used for running the experiment.

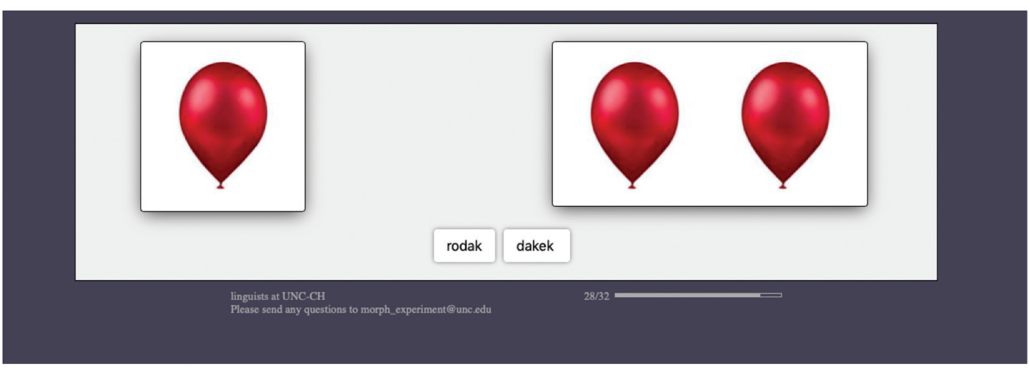

Figure 4. Example testing trial. Subjects hear the base-form "dak" and have to chose the correct plural form from the two options.

method (Buhrmester et al., 2011). A total of 142 Mechanical Turk workers participated in the experiment, which lasted on average 25 min and for which they were paid 2.00 USD.

Results

The results were analyzed for proportion of correct responses on testing trials. Based on post-questionnaire responses we included the distinction between correct staters vs. others into our set of dependent variables. It should be noted that the latter group included subjects who did not state any rules as well as subjects who stated incorrect rules. To give the reader an idea of what participants said in their questionnaire responses, here are some representative examples.

(1) a. Successful explicit learners (correct staters): *"It seems living things followed the 'ek' pluralization, while inanimate objects followed the 'ro' pluralization"; "Short words usually ended a certain way and long words also ended a certain way."; "If it was round or square the plural was ro before the word. If it was long/thin its plural was ek at the end."*

b. Others:

(i) Unsuccessful explicit learners: *"For things that need electrical energy to operate it will end with the letters ek, and for things that need human operations then those words start with ro"; " . . . tried animacy but quickly dropped that then tried listening for final consonant, then tried listening for the last vowel, avoided repeat vowels."*

(ii) Intuitive, implicit learners: *"Just followed what sounded like the original"; "just tried to listen and learn"; "one that sounded most similar was probably it"; "it was mostly what sounded proper"*

First, let us look at test performance in the four conditions in which the phonological and semantic features were independent. These conditions are: Phon(anim) (phonological condition with syllable length being the relevant feature, and animacy being the irrelevant feature), Phon(shape) (phonological condition with syllable length being the relevant feature, and shape being the irrelevant feature), Anim (semantic condition with animacy being the relevant feature and syllable length being the irrelevant feature), and Shape (semantic condition with shape being the relevant feature and syllable length being the irrelevant feature). Figure 5 shows the mean proportion of correct responses on 24 test trials.

The data were analyzed using mixed-effects logistic regressions using the lme4 package in R (Bates et al., 2015). First, individual models for each condition were fit with participant as a random effect and no fixed effects to determine if performance in each condition was below or above chance level of 0.5. Intercepts in all four models were positive and significant indicating that adults were able to learn the distribution of allomorphs and apply it to novel words in all conditions, although to various degrees of success (see Table 2).

Our first hypothesis was that learning in the phonological conditions should be better than learning in the semantic conditions. This hypothesis was not confirmed. The easiest rule to learn was animacy, followed by syllable number, followed by shape. Thus, there is no apparent across the board advantage for phonology over semantics. It could be that this order is dictated by feature-specific bias or salience. However, given that bias and/or salience (detected by comparing ease of learning in independent conditions) can differ for implicit vs. explicit systems, in the next series of analyses we explore the effects of learning strategy. In the post-experimental questionnaire, 42 out of 91 participants in independent conditions said they followed a rule, and 35 of them stated a correct rule. Consider Figure 6 which in addition to group means includes by-subject means distinguished by subjective report (stating a correct rule, reporting rule-use but no correct rule, and reporting intuition use). It is clear that those who stated a correct rule performed at or near ceiling, while those who did not state

[3]Being a native speaker of English was listed as one of the prerequisites for participating in the study. Also, in the demographic post-questionnaire we asked participants to list their native language and any other languages they speak. We excluded those who did not list English as their native language.

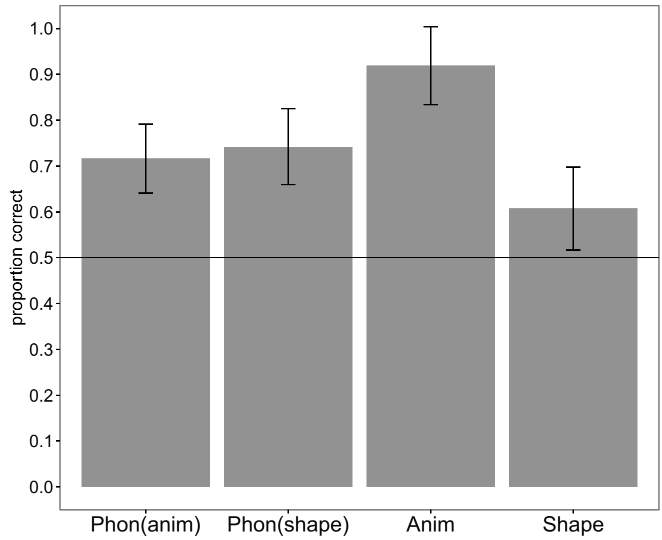

Figure 5. Group means of performance in the test phase of phonology and semantics conditions.

Table 2. Intercept coefficients and corresponding predicted probabilities of correct responses from mixed-effects logistic regressions for each condition.

condition	intercept coeff.	predict.prob.	p-value
Phon(anim)	1.35	0.79	<0.01
Phon(shape)	1.28	0.78	<0.01
Anim	5.27	0.99	<0.01
Shape	0.52	0.63	=0.01

a correct rule performed generally worse. There also appears to be a difference between phonological and semantic conditions for participants who did not state a correct rule (confirmed in the regression model below).

The rule-users had a bimodal distribution of performance – they performed either near ceiling or

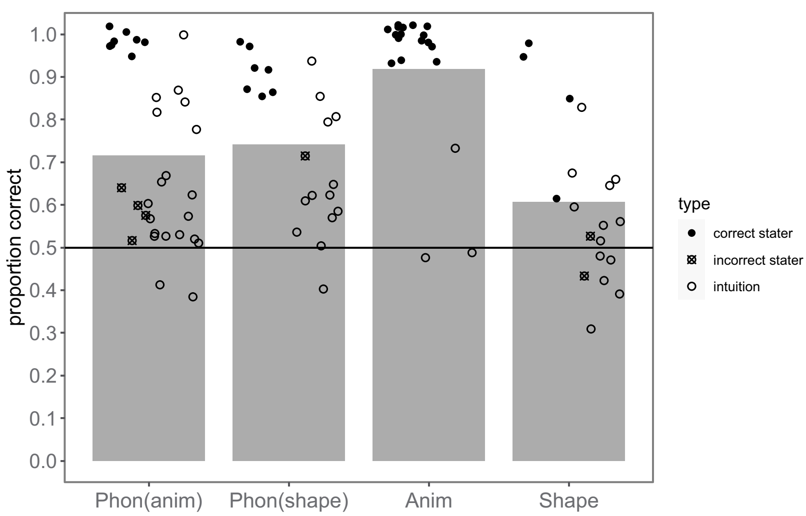

Figure 6. Average by-subject proportion of correct responses on test trials of independent conditions. Plotting symbols: Black circle = stated correct rule; crossed circle = reported rule-use, but did not state correct rule; white circle = reported relying on intuition.

closer to chance – which is typical for explicit learners. On the other hand, self-reported intuition users had a more monomodal distribution of performance. To test whether performance was above chance for intuition vs. rule-users, we ran separate mixed-effects logistic regression models for these sub-groups with subject as a random effect. The results in Table 3 show that the only group of subjects who performed at chance were intuition users who were trained on a semantic pattern, three of them were in the animacy condition and 13 in the shape condition.

To statistically test for the effect of stating a correct rule, we ran a mixed effects logistic regression model with fixed effects being feature type and stating correct rule (*corr.stater* in Table 4). A preliminary analysis showed that the difference between Phon(*anim*) and Phon(*shape*) was not significant, which means that the irrelevant semantic feature did not have an effect on learning the phonological rule. Therefore, for simplicity, these two conditions are collapsed in the model.

The significant intercept indicates that subjects performed above chance in the reference cell, "Phon, non-corr.stater." That is, even those who did not report finding a phonological rule or stated a wrong rule performed significantly above chance. Non-staters in the animacy and shape conditions performed worse, although this difference was only significant for the shape condition and not the animacy condition, likely because the latter only had three non-staters, so the effect may not be reliably estimated. There was a significant improvement in performance when switching to *corr.stater* in the phonology condition, and an even greater improvement for correct staters in the animacy condition (the interaction between *corr.stater* and *anim*). This shows that subjects who verbalized a correct rule after the experiment were among the top performers especially if they learned the animacy rule. The non-significant interaction of *corr.stater* and *shape* does not allow us to conclude that phonological correct staters were on average different from shape correct staters (although again this effect may not have been reliably estimated because there were only 4 correct staters in the shape condition).

We now turn to the mixed conditions in which phonological and semantic features were conflated or aligned during training. Based on the results in the independent conditions, we expect that subjects would be more likely to learn the phonological rule when it competes with a rule based on shape (because learning in the phonological condition was better than in the shape condition), and the animacy rule otherwise (because the animacy condition had best performance overall). However, we expect that most animacy learners will be correct staters and that there should be more of them compared to shape-learners. Additionally, if unsuccessful explicit learners and other non-staters are implicitly learning the phonological pattern, we expect that they will still show a slight preference for phonological choices over the semantic ones.

Table 3. Intercept coefficients and corresponding predicted probabilities of correct responses from mixed-effects logistic regressions for intuition vs. rule-users trained on semantic and phonological patterns.

condition	N	intercept coeff.	predict.prob.	p-value
Phon, intuition	32	0.65	0.66	<0.01
Phon, rule	21	2.87	0.95	<0.01
Sem, intuition	16	0.19	0.55	0.15
Sem, rule	22	4.6	0.99	<0.01

Table 4. Estimates for the fixed effects of the mixed-effects logistic regression (with subject as a random effect) for conditions in which semantic and phonological cues are independent.

	Coeff.	Std. Error	z value	Pr(>|z|)	
Intercept ref: phon, non-corr.stater	0.6	0.11	5.52	0	***
anim	−0.42	0.39	−1.08	0.28	
shape	−0.42	0.2	−2.11	0.03	*
corr.stater	2.65	0.32	8.36	0	***
corr.stater*anim	1.84	0.71	2.57	0.01	*
corr.stater*shape	−0.92	0.53	−1.73	0.08	

Recall that half of the test trials in the mixed conditions were of the same type as the training trials (following Culbertson et al. (2017) we call them “aligned”), that is, they were from the two cells in the paradigm in which semantic and phonological features make the same prediction. The other half of the trials were from the unseen cells (“conflicting”), trials for which phonological and semantic features make opposite predictions. Figure 7 shows group means of performance on all test trials in mixed conditions. For aligned trials, the y-axis in this figure can be interpreted as the average proportion of correct responses (responses that agree with the training data), but for conflicting trials, the y-axis shows the proportion of responses that would be correct given a semantically based rule. We see that on aligned trials subjects performed relatively well, and that they tended to favor the semantic rule when animacy was competing with syllable number, but not when shape was competing with syllable number. This pattern confirms the expectations based on the independent conditions.

To see by-subject performance broken down based on stating a correct rule see Figures 8 and 9. In these figures each point represents a subject with x-axis showing proportion of correct responses for that subject on aligned trials, and y-axis showing proportion of responses following a semantic rule on conflicting trials. One can see that correct staters cluster into two groups: those in the upper right corner learned and stated the semantic rule, and those in the lower right corner learned and stated the phonological rule. There were two correct staters in the animacy condition, however, who surprisingly performed around or below chance. One of them likely applied the rule backwards as can be seen from their low proportion of correct choices on both aligned and conflicting trials. The other stater was more or less at chance on both types of trials. It is possible that the rule only occurred to them after the testing phase was over. The fact that most correct staters were in the animacy rather than shape condition is consistent with the results from independent conditions. Looking at the data of subjects who did not state a correct rule (Figure 9), we can notice that most of them occupy the bottom-right corner of the plot, indicating that they performed above chance on the aligned trials and tended to make phonological choices on conflicting trials (recall that low proportion of semantic choices on conflicting trials is equivalent to high proportion of phonological choices). In fact, several of these subjects perform almost perfectly according to the phonological rule.

We did not use logistic regression to analyze the data from the mixed conditions because in these conditions some people acquired a phonological rule, some a semantic rule, and some did not acquire any rule at all, resulting in a pattern of means that was not normally distributed and had large variance. Instead, we performed a number of by-subject analyses. First we compared proportion of participants

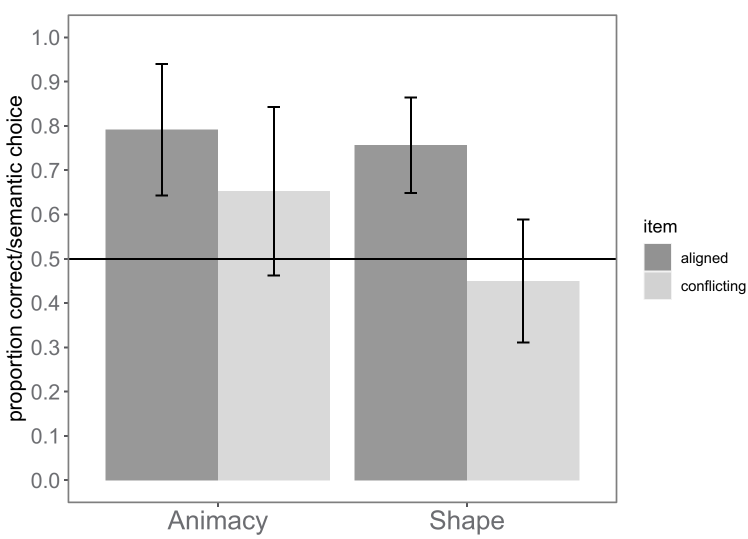

Figure 7. Group means of performance in the test phase of mixed conditions. Error-bars show 95% confidence intervals.

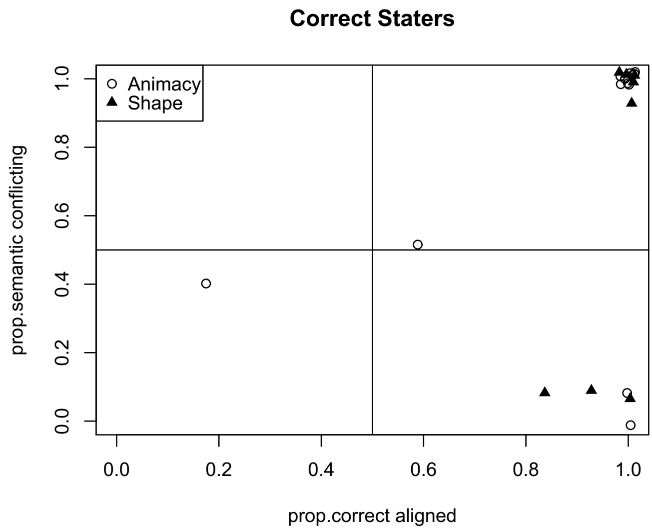

Figure 8. By-subject means on aligned and conflicting trials in mixed conditions for correct rule-staters.

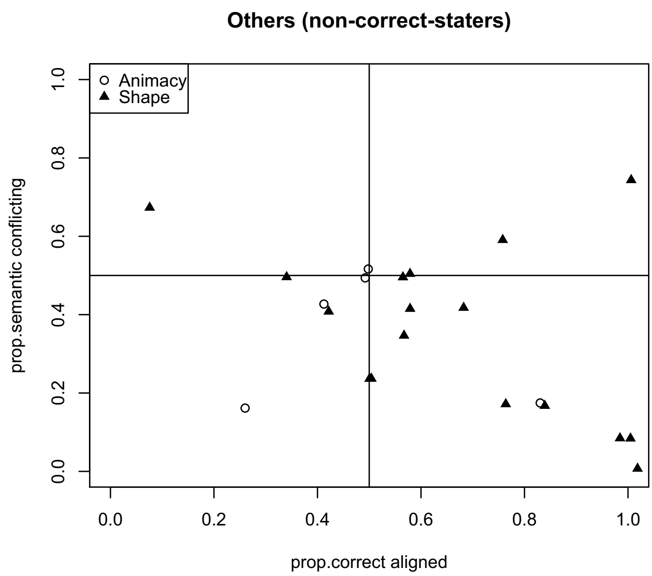

Figure 9. By-subject means on aligned and conflicting trials in mixed conditions for those who did not state a correct rule.

that followed the semantic rule when the semantic feature was “animacy” vs. “shape”. A participant was considered to follow a semantic rule if their proportion of semantic choices was above the upper limit for the 95% confidence interval for the mean of 0.5 (i.e., there’s less then 5% chance that they would get this many correct responses by guessing). This limit is 0.7, or at least 17 correct responses out of 24 test items. The limit was calculated as 0.5 + (1.96*se), where *se* is standard error for the sampling distribution of proportions for samples of size 24 which equals $\sqrt{(0.5 * (1 - 0.5))/24}$. Recall that based on the results of the independent conditions, we expected that more people would pick up on the animacy rule compared to the shape rule. This hypothesis was supported by a one-tailed proportion test – more participants responded according to the semantic rule when the semantic feature was animacy (9/18 = 0.5) rather than shape (6/25 = 0.24; $\chi^2 = 3.1$, $p = 0.04$). We can also test this hypothesis by comparing the proportion of correct staters in these two conditions. The difference is again statistically significant, with a greater number of people in the animacy condition finding a semantic rule (11/18 = 0.61) compared to the shape condition (5/25 = 0.2; $\chi^2 = 5.9$, $p = 0.01$).

We also expected that participants would be more likely to learn the phonological rule when it competes with the shape rule compared to the animacy rule. To test this hypothesis, we compared the proportion of people in the animacy vs. shape conditions whose behavior on the testing trials could be characterized with 95% confidence as following a phonological rule (see the explanation above for how this number was calculated). There were three such participants in the animacy condition (3/18 = 0.16) and eight in the shape condition (8/25 = 0.32), with the difference not being statistically significant by Fisher's exact one-tailed test ($p = 0.4$). As for correct staters, there were two of them who found the phonological rule in the animacy condition and three in the shape condition (again not a significant difference). Therefore, our expectation that we would see more phonological learning when the competing semantic feature is shape rather than animacy was not confirmed.

Finally, subjects who did not state a correct rule appear to have a tendency toward the phonological choice. To test if this tendency is in fact significant, we ran a one-sample one-tailed non-parametric Wilcoxon signed rank test on the proportion of phonological choices for non-staters which confirmed that the median proportion of these choices was greater than what is expected by chance (V = 164, $p = 0.002$). This result is similar to what we found in the independent conditions. Recall that in independent conditions, non-staters performed above chance when learning a phonological rule, but at chance when learning the shape rule.

Summary and discussion of experiment 1

The results of Experiment 1 in the conditions in which the two features were independent are largely consistent with the results in the conditions in which the two features were correlated. In both cases we see that an animacy-based pattern was the easiest to learn. In fact, nearly everyone in the animacy conditions performed close to ceiling and an overwhelming majority were also able to verbalize the animacy rule after the experiment, suggesting that this pattern was acquired mostly explicitly. Relatively few people were able to learn the shape-based rule either when shape was the only relevant feature or when it was correlated with the phonological feature. Overall, it appears that at least for explicit learners, the animacy-based rule was the easiest to learn, followed by syllable number, followed by shape. This result is consistent with previous findings that animacy is easier to learn than shape (Culbertson et al., 2017).

Most interestingly, however, in independent conditions self-reported intuition-users and non-correct-staters performed above chance when learning a phonological pattern despite their inability to state this pattern. Likewise, in the mixed conditions those who did not state a correct rule still showed a slight phonological preference in their performance. We take this as evidence of implicit learning happening in the phonological conditions. However, it is also possible that some of the high performing non-staters in the phonological conditions were actually explicit learners who for some reason did not verbalize the rule. These high performers include two learners in the independent conditions and three learners in the mixed conditions, all of whom performed above 95% on test trials. For what it is worth, all of these participants indicated that they relied on intuition in their questionnaires. Additionally, excluding these participants from consideration does not change the overall results. In contrast to the phonological conditions, we found no evidence of implicit learning in the semantic conditions. However, for the feature of animacy chances of finding such evidence were low simply because most animacy learners correctly verbalized the rule.

The pattern of responses we observed in Experiment 1 is consistent with the existence of a phonological bias in implicit learning. However, it is also consistent with an alternative explanation based on feature-salience. In particular, a reviewer suggests the following possibility. If learning progresses along a path that begins with implicit learning and gradually reaches explicit awareness, then our results may indicate that subjects move faster along this path when learning an animacy pattern, slower when learning the phonological regularity based on syllable number, and slowest of all when learning a shape-based rule. That is, the most salient feature, animacy, is learned quickly and learners reach explicit awareness; the feature of intermediate salience, syllable number, shows mixed behavior – some subjects reach awareness and others are still at the implicit stage; finally, the least salient feature, shape, is not learned well at all. In principle, we find that this explanation is plausible but it entails that we should be able to observe some signs of early implicit learning for a feature like animacy for example, if the training phase is made shorter or more difficult. We would also expect to observe implicit semantic learning in children, who are arguably worse than adults at transferring knowledge to the explicit system. It so happens that our next experiment does not confirm these expectations.

Experiment 2

In the second experiment, we tested children and a control group of adults in person, rather than over the web. We also used a slightly different design than in Experiment 1. The goal of this experiment was to see whether children would behave differently than adults when learning the same pattern of allomorphy in an artificial language, and whether we can replicate some of the findings of Experiment 1 with non-virtual participants. Our hypothesis was that children would be more likely than adults to show a phonological bias because such bias has been reported in natural L1 acquisition. Additionally, if learning mode has an effect on phonological bias, we expect to see an interaction between learning difficulty and measures of explicit/implicit learning.

Stimuli and design

The stimuli and the task for this experiment were similar to those of Experiment 1 with the following differences. We only used one phonological and one semantic feature, number of syllables and animacy, respectively. The plural morpheme had two allomorphs as before which were both suffixes (rather than a suffix and a prefix). The stimuli were rerecorded using a child-directed style of speech with exaggerated F0 by a linguistically trained native speaker of English.

Given that in Experiment 1 the same advantages that were observed in the independent conditions were also observed in the mixed condition (namely, the animacy advantage for correct staters, and the phonology advantage for others), in Experiment 2 we planned to first test the independent conditions, and reserve the mixed conditions for the scenario in which both features were learned equally well independently. In other words, we expect that if a bias for a particular group of learners shows up in the independent conditions, it will also show up in the mixed conditions. To foreshadow the results, we did find the animacy advantage for explicit learners and the phonology advantage for implicit learners in the independent conditions (consistent with Experiment 1) and, therefore, we did not run the mixed condition.

Another big change introduced in Experiment 2 was two different ways of ordering stimuli in the training mode: random order as in Experiment 1 and blocked order in which all words with one of the suffixes were presented together as a block, followed by all words with the other suffix. The blocked order was introduced for two reasons: first, we hoped that it would make learning easier for children. Second, the blocked order has been hypothesized to facilitate a different learning strategy compared to random order. Random order of stimuli facilitates implicit learning because it places a greater demand on working memory and is, therefore, hypothesized to be more difficult for the explicit system which relies on working memory and executive attention (Noh et al., 2016). On the other hand, blocking all items of the same category together allows subjects to focus on within category similarities and makes hypothesis testing easier. Thus, any potential differences between explicit and implicit learners in Experiment 2 could be enhanced by the difference in stimuli order.

The final difference introduced in Experiment 2 was that the contrast between animate and inanimate trials was cued using videos for animals and static pictures for objects. We feel that this made the animacy contrast in our experiment more interesting for children and more natural since in real life animacy is cued by volitional movement, among other things (Poulin-Dubois et al., 1996; Spelke, 1991; Woodward et al., 1993).

Four experimental conditions were created by crossing two experimental factors: semantic vs. phonological cue and the training mode, blocked vs. random. This resulted in 4 conditions: semantics blocked (S-block), semantics random (S-rand), phonology blocked (P-block), and phonology random (P-rand). In the blocked mode, 10 training stimuli with one of the suffixes were presented first, followed by one generalization trial, followed by the next 10 stimuli with the other suffix and the corresponding generalization trial. For example, in the semantic condition a training phase could look like this: a participant sees 10 animate nouns which all form the plural with the suffix *-um*. Then the participant is tested on one new animate noun (using the same forced-choice task as in the testing phase), and this is followed by another training block of 10 inanimate nouns, whose plurals are formed by a different suffix, *-ek*. After this there is one more generalization trial for inanimates and a new screen explaining that in the next phase of the experiment more new words and pictures will follow, and the participant will have to choose the best word for a picture from the two given alternatives. In the random training mode, 20 training stimuli appear in random order. The testing phase was the same for both modes and followed the same structure as in Experiment 1. It consisted of 12 new items, six of each type in random order (e.g., animate vs. inanimate).

Procedure

The experiment with children was run at the Museum of Life and Science in Durham, NC and at a local afterschool program in Carrboro, NC. The locations presented some challenges: for example, we were not able to perfectly control the level of noise during the experiment and ensure that all children were tested in the same conditions. On the other hand, running the study in a museum allowed for more rapid recruitment and provided a novel activity for museum visitors to participate in. It should also be noted that we could not detain museum visitors for a long time. This constraint together with children's lower attention span meant that we had to shorten our study to last no more than 10–15 minutes. This explains the lower number of training stimuli in Experiment 2 compared to Experiment 1 (20 vs. 40 training items, respectively).

Before starting the experiment we briefly talked to children about other languages they might know. We then explained to them that in this activity they would be learning a new foreign language that has different ways of saying "more than one of something." They were given a practice trial demonstrating the task and an opportunity to ask questions. During training children saw singular – plural pairs as in Experiment 1 and they were encouraged to repeat them aloud. If they chose not to do that, the experimenter repeated the words for them before going to the next trial. At the very end of the experiment, subjects were asked how they made their choices during testing. Their responses were then coded based on two criteria: "stated some explicit rule, pattern, generalization or tendency", and "stated a correct rule/pattern."

Seventy-one monolingual English-speaking children between ages of 5 and 11 participated in the experiment (conditions were roughly similar in the distribution of ages: the means across conditions ranged between 7.4 and 7.9 years old). Ten children were excluded due to response bias (always picking the same suffix – four children, or always picking the last item – six children),[4] leaving 61 children in the experiment. A control group of 61 adults (some were parents of the children and some were college students) were run using the same procedure. One of them was excluded due to response bias. All participants were volunteers and received a small token gift for their participation.

[4]It is very likely that children who always picked the last item did so because the label for the last item was slightly closer to the plural picture due to the words being centered on the screen (see Figure 4). In fact one child explicitly stated this as a reason for always picking the last item.

Results: adults

Let us first examine the results from the control group of adults. The group and individual means of the proportion of correct responses in the test phase are displayed in Figure 10.

Adults performed above chance in every condition (see Table 5 for intercepts of mixed-effects logistic regression models with no fixed effects). As in the first experiment, adults performed better when learning animacy-based allomorphy compared to one based on syllables. The majority of participants in the animacy conditions correctly pluralized more than 90% of words in the testing phase, and all of these top-performers were correct staters. This time, however, with a much shorter training phase there were more animacy participants who did not state a correct rule compared to Experiment 1, and these participants performed at chance (see analysis below).

To test whether stating a correct rule had an effect on performance in each condition, we also ran individual mixed-effects models for each condition with a single predictor, correct stater. The results of these models (summarized in Table 6) showed that the effect of *corr.stater* was significant for all conditions except for phonology-blocked which only had three correct staters who all performed 100% correct, so the model could not estimate variance for this group. The intercepts of these models (corresponding to participants who did not state a correct rule) were not significant except for the phonology random condition ($\beta = 0.36$, predicted prob. = 0.59, $p = 0.03$). That is, non-correct-staters performed at chance in all conditions except for phonology-random condition in which they were slightly but significantly above chance.

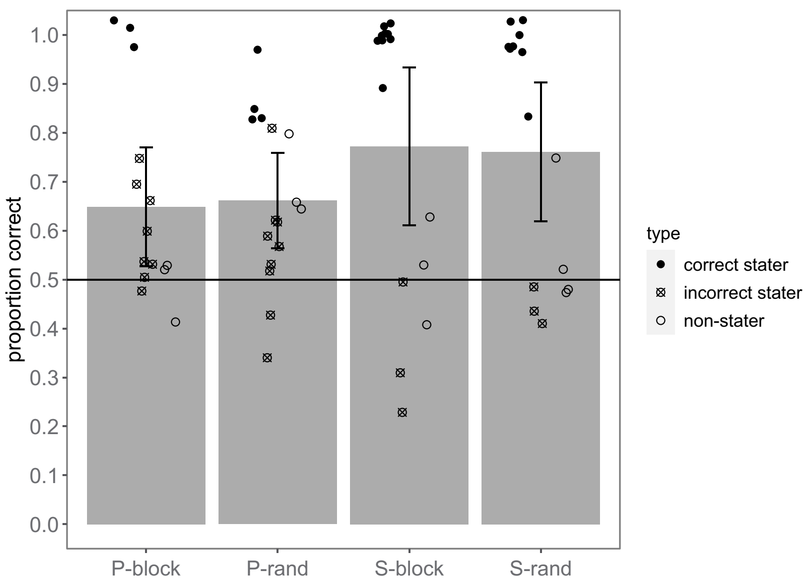

Figure 10. Adults. Group and individual means of correct responses in the test phase. Plotting symbols: Black circle = stated correct rule, crossed circle = stated an incorrect rule, white circle = didn't state a rule. Error bars show 95% confidence intervals.

Table 5. Intercept coefficients and corresponding predicted probabilities of correct responses (from mixed-effects logistic regressions for each condition).

condition	intercept coeff.	predict.prob.	p-value
P-block (phonology-blocked)	0.73	0.68	0.01
P-rand (phonology-random)	0.72	0.67	<0.01
S-block (semantics-blocked)	3.00	0.95	0.01
S-rand(semantics-random)	2.17	0.90	<0.01

Table 6. Fixed effects coefficients from mixed-effects logistic regressions for each condition with a single fixed effect, correct stater.

condition		intercept coeff.	st.error	p-value	num. of correct staters
P-block:	intercept	0.21	0.68	0.22	
	corr.stater	not reliable	not reliable		3
P-rand:	intercept	0.36	0.67	0.03 (*)	
	corr.stater	1.58	0.46	<0.01 (*)	4
S-block:	intercept	−0.22	0.23	0.34	
	corr.stater	4.89	1.03	<0.01 (*)	9
S-rand:	intercept	0.04	0.21	0.82	
	corr.stater	3.8	0.75	<0.01 (*)	8

Table 7. Experiment 2: Estimates of the fixed effects of a mixed-effects logistic regression model (with subject as a random effect) for adults.

	Estimate	Std. Error	z value	p	
Intercept ref: *P*-rand, non-corr.stater	0.39	0.15	2.51	0.01	*
blocked	−0.19	0.19	−1.00	0.31	
sem	−0.37	0.20	−1.85	0.06	.
corr.stater	1.68	0.47	3.55	0	**
blocked*corr.stater	2.21	1.09	2.03	0.04	*
sem*corr.stater	1.89	0.75	2.50	0.01	*

To test how the overall performance was influenced by the relative contribution of the training mode (blocked vs. random), feature type (phonology vs. semantics), and stating of the correct rule we fit a mixed-effects logistic regression model with the above factors as fixed effects and subject as a random effect. The fixed effects of this model are reported in Table 7. The significant intercept confirms that performance in the baseline condition (P-rand, non-corr.stater) was above chance. Performance does not significantly change with a switch to P-block (the effect of *blocked*), but decreases with a switch to S-rand (the effect of *sem*) although this decrease is not quite significant (p = 0.06). Performance increases significantly for P-rand correct staters (the effect of *corr.stater*), as well as for P-block correct staters, and for S-rand correct staters (the interactions). Thus, overall stating a correct rule is associated with better performance in the test phase. As far as by-participant analysis, significantly more participants stated a correct rule in the two semantics conditions (17/30 = 0.56) compared to the two phonological conditions (7/30 = 0.23, χ^2 = 5.6, p = 0.01).

These results replicate the results of Experiment 1 run over the web: there are more correct staters in the animacy conditions compared to the phonological conditions. Non-correct-staters performed at chance, except in the phonological condition when the stimuli were presented in random order.

Results: children

Children's results are shown in Figure 11. Unlike adults, children performed above chance only in two conditions as evidenced from the intercept coefficients of condition-specific models with no predictors: P-rand (β=0.45, predicted prob. = 0.61, p = 0.004) and S-block (β=0.96, predicted prob. = 0.72, p = 0.031). Most children, when asked how they chose plural forms in the test phase, said something along the lines of "I don't know," "I'm so smart," or "[I picked] whatever sounded right." Fourteen children stated some sort of a rule referring to a connection between the stimuli and their responses, with eight of them stating the correct rule. All high performers in S-block stated the correct rule, and 7 out of 8 children who stated a correct rule were in the S-block condition (one was in S-rand). No child stated a correct rule or performed at ceiling in the phonological conditions.

Child correct staters tended to be somewhat older (mean age of 8.75 compared to the mean age of non-staters = 7.5). Figure 12 shows the age distribution for correct-staters vs. others. Here are a few examples of correct rules stated by children: "Animals are one sort of way and like chairs are another

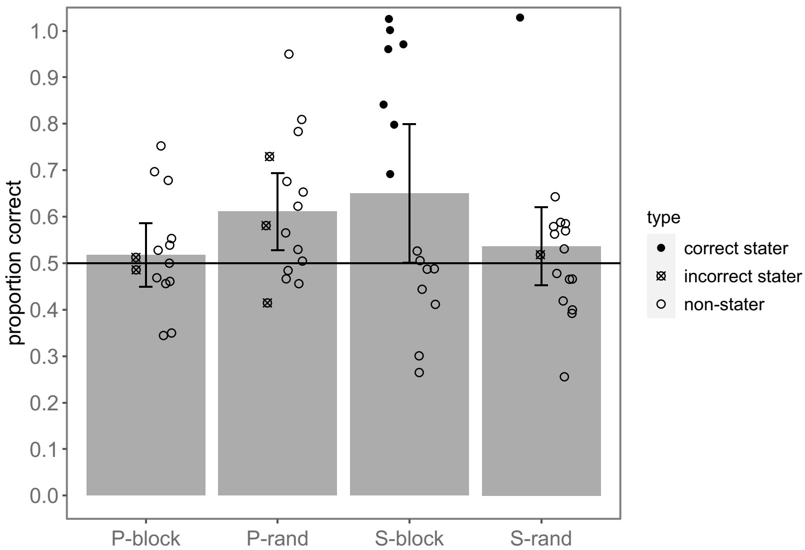

Figure 11. Children. Group and individual means of correct performance in the test phase.

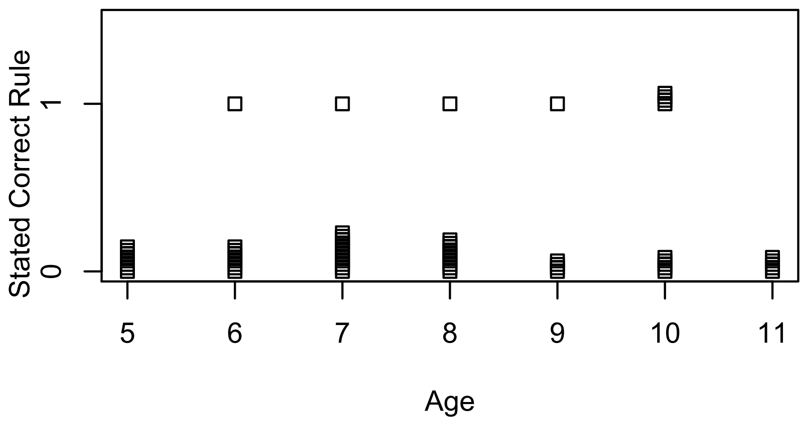

Figure 12. Number of correct staters by age.

sort of way" (9 y.old); "animals had -um at the end and things had -ek" (8 y.old); "-ek meant a tribe, -um meant more than one." (6 y.old).[5]

Since stating a correct rule was something children did almost exclusively in only one condition, the effect of this factor cannot be disentangled from the effect of S-block. To deal with this, we ran two separate models: the first one does not distinguish between staters and non-staters and the second one excludes correct staters from the analysis. Thus, the first model (summarized in Table 8) analyzes children's overall performance regardless of whether they showed conscious awareness of what they learned. The only significant effect in this model was the interaction between blocked training mode and semantics, indicating that performance in the S-block condition was significantly better than in

[5]This child appears to have learned a collective vs. individuated plural distinction, interpreting the animate suffix as collective since most of the videos for plural animate nouns showed groups of animals, while pictures for inanimate objects showed multiple versions of the same object. We counted this response as correct because it was still based on a semantic distinction, and was consistent with the presented data.

Table 8. Experiment 2: children (all). Estimates of the fixed effects of a mixed-effects logistic regression model (with subject as a random effect).

	Coeff.	Std. Error	z value	Pr(>|z|)	
Intercept ref: *P*-rand, 5y.old	−0.30	0.47	−0.64	0.52	
age	0.098	0.054	1.80	0.07	.
blocked	−0.35	0.28	−1.23	0.21	
sem	−0.31	0.27	−1.11	0.26	
blocked*sem	0.89	0.39	2.22	0.03	*

Table 9. Experiment 2: children (non-correct-staters). Estimates of the fixed effects of a mixed-effects logistic regression model (with subject as a random effect).

	Coeff.	Std. Error	z value	Pr(>|z|)	
Intercept ref: *P*-rand	0.43	0.14	3.23	0.001	**
blocked	−0.35	0.16	−2.13	0.03	*
sem	−0.40	0.16	−2.45	0.01	*

P-rand. This result is not surprising given that S-block had a lot of top-performers who were all correct staters.

The second model excluded correct staters from the analysis. This model resulted in "singular fit" in which the variance of the random effect was estimated to be 0, suggesting that subjects did not consistently vary across different subconditions. There are varying opinions on what to do in such cases. Singmann and Kellen (2017) suggest that it is more conservative to accept a model with some degenerate parameters (variance close to 0) than to remove random effects. They say "It is clear that a model with such problematic or degenerate parameters is not the most adequate from a purely statistical standpoint, but it can nevertheless be a reasonable solution if the focus is ultimately on the fixed effects." Taking this suggestion, we ignored the degenerate random effect. In the full model, only the effect of semantics is significant ($\beta = -0.43$, $z = -2$, $p = 0.04$), the effect of blocking is marginal ($\beta = -0.38$, $z = -1.76$, $p = 0.08$), and the interaction between the two, as well as *age* are not significant predictors. Removing the interaction and the age predictor leads to a model with a slightly better fit (based on AIC and BIC values) which is reported in Table 9. According to this model, performance in P-rand (the reference cell) is significantly above chance, and changing the cue to "semantics" or training mode to "blocked" results in significant decrease in performance.

Summary and discussion of Experiment 2

Experiment 2 replicated most of our findings of Experiment 1: participants are more likely to show ceiling or near ceiling performance on the animacy rule compared to the syllable-based rule; however, this advantage is entirely due to correct staters. As before, there were fewer correct staters in the phonological conditions compared to semantic conditions (in fact, there were no phonological correct staters among children).

Recall that results of Experiment 1 are consistent with two different explanations. First, it could be that explicit learners are biased toward the animacy rule, while implicit learners are biased toward the phonological rule. The other explanation is that learning involves gradual progression from implicit learning to explicit awareness and that the animacy feature is more salient than the phonological feature and so progresses faster along this path, resulting in better overall learning of the animacy rule coupled with the ability to state it. However, if explicit learning results from prior implicit learning, we would expect to see some signs of implicit semantic learning early on. But results of Experiment 2 showed that adult and child non-correct-staters in both S-rand and S-block conditions performed at chance. In other words, animacy learners either get the rule and are able to state it, or they do not get it at all – there is no intermediate stage. On the other hand, adult and child phonology learners who did not state the correct rule performed slightly but significantly above chance when the stimuli were presented randomly. Thus, it is likely that our experiment participants used mostly an implicit learning strategy when presented with a phonological pattern, but not when presented with a semantic one. Together Experiment 1 and Experiment 2 provide converging evidence that a phonological rule based on syllable number or word-length can be learned implicitly even after very brief exposure. Given this difference in the pattern of responses to the phonological vs. semantic cues we find it implausible that learning of these two types of rules proceeds along a similar path. Our conclusion is also consistent with results of other experiments discussed in the Background section that used different semantic and phonological features and found no evidence of semantic implicit learning.

Another finding of Experiment 2 was the effect of training mode on learning which was particularly pronounced for children: learning of the semantic rule was almost exclusively observed in the blocked condition but not the random condition, while learning of the phonological rule was only observed in the random condition, but not the blocked condition. The negative effect of blocking in the model in Table 9 is consistent with the expectation that implicit learning is not facilitated by blocked training. The fact that age was not a significant predictor in this model suggests that implicit learning may be less affected by age differences compared to explicit learning (recall that in the previous model in which correct staters were included, age was nearly a significant predictor). For adults the difference between blocked and random modes was less pronounced. It was only manifested in that non-staters in P-rand performed significantly above chance while non-staters in P-block were at chance (see Table 6), although the difference between these two conditions was not significant. Overall, it appears that the random order of the stimuli facilitated implicit learning, but only for the phonological feature. We suspect that the different effects of blocking for semantic vs. phonological features for children had to do with the fact that phonological awareness is not as developed in children as it is in adults and so the phonological feature was not available to the system of conscious reasoning. On the other hand, children are known to establish at a very young age a conceptual distinction between animals and inanimate things (see Gelman & Opfer, 2002 for an overview of the literature). By 1 year of age children show evidence of differential categorization of animate and inanimate entities (Mandler & McDonough, 1993), and by age 3 or 4 years they can reason in an adult-like way about typical properties of animates (Becker, 2006; Massey & Gelman, 1988). Thus, children were able to take advantage of the conceptual animate–inanimate distinction in a condition that maximally promotes explicit reasoning (blocked presentation). In a condition that promotes implicit reasoning (random presentation), children failed to learn both the animacy and the syllable rule explicitly. On the other hand, since adults have greater phonological awareness and are better at explicit reasoning, some of them were able to learn the phonological rule explicitly. Overall the main difference between children and adults is that adults are more likely to achieve explicit awareness in all conditions. In other respects adults and children are similar – both can imperfectly learn the phonological pattern (in contrast to the semantic one) when experimental conditions promote implicit learning, and both are more likely to learn the animacy rule explicitly compared to the syllable-based rule.

General discussion

The experiments presented here aimed to confirm the existence of a phonology-over-semantics bias and test whether the same bias is observed in adults and children. In the process we discovered that it was important to be cognizant about differences between explicit and implicit learning. While it is not possible to perfectly diagnose the extent to which a participant in an A.L.L. study is using one or the other learning mode (or, possibly, both modes), we observed a different pattern of performance from correct staters vs. all others. We hypothesize that above chance performance of those subjects who could not correctly verbalize a rule post-experiment must be due primarily to implicit learners. At least some subjects in this group must have been learning implicitly (even if they were deliberately searching for a rule) to bring up the overall performance of this group above chance level.

Our main finding can thus be summarized as follows: we found evidence that the syllable-based regularity was learned better than an animacy-based one by non-correct-staters (both adults and children), while the reverse was true for correct staters. Phonological learning was observed in the presence of hallmarks of implicit learning, in particular random presentation of stimuli, inability to state a rule post-experiment, performance that is above chance but not at ceiling. On the other hand, semantic learning was only observed for those participants who could correctly state a rule. A key question is whether this finding supports the existence of a phonological bias, particularly given that we only tested three specific features (and only two in Experiment 2). Is there a reason to think that similar results would be found with different semantic and phonological features? More research is necessary to shed light on this question. However, past studies with different semantic and phonological features are consistent with what we found. Namely, implicit phonological learning in A.L.L. has been reported for adults (Chan & Leung, 2014; Moreton & Pertsova, 2016), children (Culbertson et al., 2019; Plante et al., 2010), and infants (Chambers et al., 2003; White et al., 2008) while evidence for implicit semantic learning is hard to find. Coming back to the Culbertson et al. (2019) study, recall that they found a phonological bias for children and a semantic bias for adults. These results are consistent with the phonological bias hypothesis on the assumption that adults are more likely to learn explicitly than children. More specifically, the fact that adults appeared to have a semantic bias in that study could be explained if the distinction between planets vs. aliens had higher salience for the explicit learning system compared to the phonological distinction based on vowel quality. To sum up, we cautiously interpret the results of the experiments reported here to be in line with the previous literature, namely, we take our findings to support the existence of a phonology bias, but only for implicit learners.

The fact that explicit learners did so well on the animacy rule raises another question: could there be a semantic bias in explicit learning? Although it is true that there were more explicit learners in the animacy condition compared to the phonological conditions in all of our experiments, it is also true that in Experiment 1 more adults learned the syllable-based rule explicitly compared to the shape-based rule. Also, in Moreton and Pertsova (2016) there was no advantage for semantic features (shape, length, edible) over phonological features even in the explicit-promoting condition – surprisingly many subjects were able to verbalize the phonological rules, although they often did that by listing specific segments rather than referring to features. This suggests that how well a pattern is learned explicitly may depend more on how salient or familiar the contrasts underlying the pattern are to the explicit system (meta-linguistic thinking about sounds in terms of letters is common for adults). Overall, the fact that some semantic cues, like animacy, may be learned better than a phonological cue, but others, like shape, worse than the same phonological cue is not consistent with a blanket semantics-over-phonology bias for explicit learners. The effect of feature salience on learning is tightly linked to the question of bias investigated here: it is possible that salience differs for the explicit and implicit learning systems. It is also not clear how salience should be defined and whether there is any independent measure of salience for comparing a phonological feature to a semantic one. Given that both salience and bias can be affected by learning mode, it is particularly important for future A.L.L. studies to distinguish between participants that may be using different learning modes.

Another caveat is that it is not entirely clear whether the bias observed in A.L.L. studies is the same bias observed in natural language acquisition, or whether it is due to the fact that typical A.L.L. experiments fail to trigger implicit learning of semantic cues. Real-life semantic contrasts are arguably more complex than what can be gleaned from the pictures or videos. That is, while nonsense words following a specific phonotactic rule closely resemble natural language phonological patterns, an image may not be a good representation of the rich information required to convey a conceptual contrast. On the other hand, we know that blind children can successfully acquire language and do not typically have difficulties with grammatical and syntactic structures even though the contextual language-external input they receive is impoverished (Landau et al., 2009). It is still possible that some other task or a different measure of success would be better at detecting implicit semantic learning compared to the production or forced-choice tasks used in many linguistic A.L.L. studies (see

Paciorek and Williams (2015) for a review of experiments that were more successful at showing learning of semantic distinctions). Once we can demonstrate that a particular task reliably tracks implicit semantic learning, then we can use this task to reassess the comparison of semantic vs. phonological pattern-learning.

A separate question is whether the kinds of A.L.L. tasks linguists use tell us something about linguistic learning (e.g., of allomorphy distribution) or about category-learning more generally, and whether the two types of learning are significantly different from each other. We remain agnostic about whether the phonological bias extends to non-linguistic domains or tasks and leave this question for future research.

If the phonology bias does indeed affect both children and adults and it is the same bias responsible for the phonological preferences found in first language acquisition this would support theories in which L2 acquisition (most closely modeled by A.L.L. studies) is constrained by the same cognitive biases and mechanisms as L1 acquisition, such as the Full Transfer Full Access model (Schwartz & Sprouse, 1996). This holds in so far as the implicit learning mechanism, thought to predominate in L1 acquisition, also plays some role in L2 acquisition (Krashen, 1982). The finding is also consistent with proposals that implicit learning is developmentally invariant (Amso & Davidow, 2012). At the same time, the fact that adult learners overall performed better in the animacy condition and there were more adult explicit learners than child explicit learners, supports the view that explicit learning plays an important role in adult L2 learning as discussed in DeKeyser (2008) and R. Ellis (2008).

If the phonological bias in implicit learning is real, where does it come from? Culbertson et al. (2019) offer several possible explanations. First, the formal contrast that is cued by a phonological or a semantic property is itself marked phonologically (in our case via two suffix allomorphs), so it follows that the cue and what is being cued are located in the same modality for phonologically-conditioned patterns and in different modalities for semantically-conditioned patterns. It may be that modality-internal dependencies are easier to spot for the implicit system. A prediction of this hypothesis is that we should observe better semantic learning if the contrast that is being cued is exclusively semantic in nature. For example, one could design an (admittedly unnatural) artificial language in which only two-syllable nouns could be subjects of stative verbs and compare it to a language in which only nouns referring to furniture could be subjects of stative verbs. If generalizations within a domain are preferred to generalizations across domains, we would predict that the second language should be easier to learn than the first (provided we can figure out how to successfully detect implicit semantic learning). Another offered explanation for the phonological bias is the higher reliability of phonological cues. In particular, phonological cues are typically present in the string expressing the pattern being learned, while the semantic cues may not be – an animal is not always present when an animate suffix is used; thus, animacy may have to be inferred from word-meaning or from context. However, it is not the case that phonological cues are always reliable either. In fact, they are rarely if ever absolutely reliable because all rules have exceptions and many phonological regularities are only statistical tendencies. For instance, in Russian there are several diminutive allomorphs that are partially related to a number of arbitrary phonological properties such as presence of hiatus in the stem, location of stress, presence of initial consonant clusters, and so on (Gouskova et al., 2015). None of these properties or their combination perfectly predict allomorph choice and some are only marginally correlated with different allomorphs, but native speakers are still implicitly aware of many of these correlations as shown in Gouskova et al. (2015) production study with existing rare and novel nouns. A third explanation of the phonological bias is the early availability of phonological cues compared to semantic cues. That is, we know that even newborn infants are already sensitive to some phonological distinctions in their native language (Ramus, 2002; Shi et al., 1999), but the first sensitivity to semantic distinctions is not reported until infants are about 6 months of age (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012). Culbertson et al. (2017) and Culbertson et al. (2019) try to recreate this sort of advantage in their studies through the effect of "staging" by withholding from participants either the phonological or semantic portion of the input early in the experiment. They find that this manipulation does indeed affect performance for both children and adults, with participants' choices on conflicting trials being skewed toward better learning whatever feature they were exposed to first. Culbertson and colleagues reasoned that early availability of phonological cues explains phonological bias in L1 and in experiments with "staging," but they did not expect to observe it in the absence of staging and, so, were surprised that children showed a phonology bias. However, it is not implausible that early availability of phonological information might in essence make phonological cues more salient than semantic cues across the lifespan if the way that the learning system is trained and optimized at early stages has long-lasting effects. We would also like to suggest that a possible reason for why phonological, but not semantic information is available to children earlier is that perhaps recognizing and learning semantic cues requires greater reliance on a mode of learning that is to some extent explicit or requires greater involvement of declarative memory (this memory has been traditionally thought to underly explicit knowledge; see an overview in Ullman and Lovelett (2018)). For example, learning the referential meaning of words has been proposed to involve some explicit reasoning (N. C. Ellis, 1994) and shared attention (Tomasello, 1995). More can be said about what counts as "explicit" since different researchers and research traditions define this term differently. But if learning of semantic cues involves some degree of explicitness (although, perhaps not of meta-linguistic awareness) or reliance on a neurally distinct system of memory and attention which matures slower than the implicit system, this would provide yet another explanation for a phonology-over-semantics bias. Greater interaction among the diverse fields that have been exploring various versions of multiple-learning systems (explicit/implicit, procedural/declarative, rule-based/similarity-based) and the implications of these theories for language and human cognition would be beneficial for shedding more light on the study of cognitive biases.

References

Albright, A. (2008). How many grammars am i holding up? Discovering phonological differences between word classes. In *Proceedings of the 26th West Coast Conference on Formal Linguistics Somerville, MA* (pp. 1–20). Cascadilla Proceedings Project.

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: Item frequencies, relations, and cognitive flexibility. *Developmental Psychobiology*, *54*(6), 664–673. https://doi.org/10.1002/dev.20587

Ashby, F. G., Alfonso-Reese, L. A., Turken, U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481. https://doi.org/10.1037/0033-295X.105.3.442

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry*, *37*(3), 441–456. https://doi.org/10.1162/ling.2006.37.3.441

Becker, M., & Levine, J. (2013). *Experigen, an online experiment platform*. http://becker.phonologist.org/experigen.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Berman, R. (1985). The acquisition of Hebrew. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 255–371). Lawrence Erlbaum Associates.

Brown, H., Smith, K., Samara, A., & Wonnacott, E. (2018). *Semantic cues in language learning: An artificial language study with adult and child learners*. https://doi.org/10.31234/osf.io/7hq2c.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69–B77. https://doi.org/10.1016/s0010-0277(02)00233-0

Chan, R. K., & Leung, J. H. (2014). Implicit learning of l2 word stress regularities. *Second Language Research*, *30*(4), 463–484. https://doi.org/10.1177/0267658313510169

Cherry, K. E., & Stadler, M. E. (1995). Implicit learning of a nonverbal sequence in younger and older adults. *Psychology and Aging*, *10*(3), 379. https://doi.org/10.1037/0882-7974.10.3.379

Corbett. (1991). *Gender (Cambridge textbooks in linguistics)*. Cambridge University Press.

Cristia', A., & Seidl, A. (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development*, *4*(3), 203–227. https://doi.org/10.1080/15475440802143109

Culbertson, J., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language*, *92*, 343–358. https://doi.org/10.1016/j.jml.2016.08.001

Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitiveity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language*, *95*(2), 268–293. https://doi.org/10.1353/lan.2019.0031
DeKeyser, R. (2008). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). John Wiley and Sons.
Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(6), 1355. https://doi:10.1037//0278-7393.26.6.1355
Demuth, K. (2000). Bantu noun class systems: Loanword and acquisition evidence of semantic productivity. In G. Senft (Ed.) *Systems of nominal classification* (pp. 270–292). Cambridge University Press.
Demuth, K. (2003). The acquisition of Bantu languages. In D. Nurse & G. Philippson (Eds.), *The Bantu languages* (pp. 209–222). Curzon Press.
Ellis, N. C. (1994). *Implicit and explicit learning of languages*. Academic Press.
Ellis, R. (2008). Explicit knowledge and second language learning and pedagogy. In L. van Lier and D. Corson (Eds.) *Encyclopedia of language and education* (pp. 1901–1911). Springer.
Ferman, S., & Karni, A. (2010). No childhood advantage in the acquisition of skill in using an artificial language rule. *PLoS One*, *5*(10), e13648. https://doi.org/10.1371/journal.pone.0013648
Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, *142*, 212–220. https://doi.org/10.1016/j.jecp.2015.09.027
Gagliardi, A., & Lidz, J. (2014). Statistical insensitivity in the acquisition of tsez noun classes. *Language*, *90*(1), 58–89. https://doi.org/10.1353/lan.2014.0013
Gagliardi, A. C. (2012). *Input and Intake in Language Acquisition* [PhD thesis]. University of Maryland.
Gelman, S., & Opfer, J. (2002). Development of the animate–inanimate distinction. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 151–166). Blackwell.
Gilhooly, K. J. (2016). Incubation and intuition in creative problem solving. *Frontiers in Psychology*, *7*, 1076. https://doi.org/10.3389/fpsyg.2016.01076
Gouskova, M., Newlin-L- Ukowicz, L., & Kasyanenko, S. (2015). Selectional restrictions as phonotactics over sublexicons. *Lingua*, *167*, 41–81. https://doi.org/10.1016/j.lingua.2015.08.014
Gureckis, T. M., James, T. W., & Nosofsky, R. M. (2011). Re-evaluating dissociations between implicit and explicit category learning: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *23*(7), 1697–1709. https://doi.org/10.1162/jocn.2010.21538
Haspelmath, M., & Sims, A. (2013). *Understanding morphology*. Routledge.
Hélie, S., & Sun, R. (2012). *Implicit cognition in problem solving* [Unpublished manuscript]. http://ccn.psych.purdue.edu/papers/chap_3.pdf
Henzl, V. (1975). Acquisition of grammatical gender in Czech. *Papers and Reports on Child Language Development*, *10*, 188–198.
Howard, D. V., Howard, J. H., Jr, Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: Effects of level of structure, adult age, and extended practice. *Psychology and Aging*, *19*(1), 79. https://doi.org/10.1037/0882-7974.19.1.79
Kager, R. W. J. (1996). On affix allomorphy and syllable counting. In U. Kleinhenz (Ed.), *Interfaces in phonology*. *Studia grammatica* (Vol. 41, pp. 155–171). Academie Verlag.
Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge University Press.
Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, *99*(2), 349. https://doi.org/10.1037/0033-295X.99.2.349
Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
Kunene, Euphrasia Constantine Lwandle. (1979). *The acquisition of SiSwati as a first language*: *a morphological study with special reference to noun prefixes, noun classes, and some agreement markers* (Ph. D. thesis). University of California, Los Angeles.
Kurtz, K. J., Levering, K. R., Stanton, R. D., Romero, J., & Morris, S. N. (2013). Human learning of elemental category structures: Revising the classic result of Shepard, Hovland, and Jenkins (1961). *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(2), 552–572. https://doi.org/10.1037/a0029178
Landau, B., Gleitman, L. R., & Landau, B. (2009). *Language and experience: Evidence from the blind child* (Vol. 8). Harvard University Press.
Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *21*(2), 284–291. https://doi.org/10.1177/0956797609358570
Lichtman, K. M. (2012). *Child-adult differences in implicit and explicit second language learning* [PhD thesis]. University of Illinois at Urbana-Champaign.
Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, *9*(4), 829–835. https://doi.org/10.3758/BF03196342

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, *66*(3), 309–332. https://doi.org/10.1016/j.beproc.2004.03.011

Mandler, J., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, *8*(3), 291–318. https://doi.org/10.1016/S0885-2014(93)80003-C

Massey, C., & Gelman, R. (1988). Preschooler's ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology*, *24*(3), 307–317. https://doi.org/10.1037/0012-1649.24.3.307

Mills, A. (1986). *The acquisition of gender: A study of English and German*. Springer-Verlag.

Minda, J. P., Desroches, A. S., & Church, B. A. (2008). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(6), 1518–1533. https://doi.org/10.1037/a0013355

Moreton, E., & Pertsova, K. (2016). Explicit and implicit processes in phonological learning. In J. Scott & D. Waughtal (Eds.), *Proceedings of the 40th Boston University Conference on Language Development (BUCLD 40)*. Cascadilla Press.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*(4), 933–947. https://doi.org/10.1162/jocn_a_00119

Nelson, C. A. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, *31*(5), 723. https://doi.org/10.1037/0012-1649.31.5.723

Noh, S. M., Yan, V. X., Bjork, R. A., & Maddox, W. T. (2016). Optimal sequencing during category learning: Testing a dual-learning systems perspective. *Cognition*, *155*, 23–29. https://doi.org/10.1016/j.cognition.2016.06.007

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing modes of rule-based classification learning: A replication and extension of shepard, hovland, and jenkins (1961). *Memory & Cognition*, *22*(3), 352–369. https://doi.org/10.3758/BF03200862

Nosofsky, R. M., & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. *Psychonomic Bulletin & Review*, *3*(2), 222–226. https://doi.org/10.3758/BF03212422

Paciorek, A., & Williams, J. N. (2015). Semantic implicit learning. In P. Rebuschat (Ed.) *Implicit and Explicit Learning of Languages* (pp. 69–89). John Benjamins Publishing Company. https://doi.org/10.1075/sibil48

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins Publishing Company.

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*(3), 571–590. https://doi.org/10.1017/S0305000900011259

Plante, E., Bahl, M., Vance, R., & Gerken, L. (2010). Children with specific language impairment show rapid, implicit learning of stress assignment rules. *Journal of Communication Disorders*, *43*(5), 397–406. https://doi.org/10.1016/j.jcomdis.2010.04.012

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. https://doi.org/10.1016/S0028-3932(02)00157-4

Poulin-Dubois, D., Lepage, A., & Ferland, D. (1996). Infants' concept of animacy. *Cognitive Development*, *11*(1), 19–36. https://doi.org/10.1016/S0885-2014(96)90026-X

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, *2*(1), 85–115. https://doi.org/10.1075/arla.2.05ram

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*(3), 484. https://doi.org/10.1037/0012-1649.39.3.484

Schwartz, B. D., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40–72. https://doi.org/10.1177/026765839601200103

Seifart, F. (2005). *The structure and use of shape-based noun classes in Miran˜a (North West Amazon)* [PhD thesis]. Radboud University Nijmegen.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*(13, Whole No. 517), 1–42. https://doi.org/10.1037/h0093825

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants? Sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21. https://doi.org/10.1016/S0010-0277(99)00047-5

Singer, I. (1982). *Simplified grammar of the Hungarian language*. Ludgate Hill.

Singmann, H., & Kellen, D. (2017). An introduction to mixed models for experimental psychology. In D. Spieler, & E. Schumacher (Eds.), *New methods in neuroscience and cognitive psychology* (pp. 1–39).

Smith, J. D., Zakrzewski, A., Herberger, E. R., Boomer, J., ssica, L., Roeder, J., Ashby, F. G., & Church, B. A. (2015). The time course of explicit and implicit categorization. *Attention, Perception & Psychophysics*, *77*(7), 2476–2490. https://doi.org/10.3758/s13414-015-0933-2

Spelke, E. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 133–170). Lawrence Erlbaum Associates.

Szagun, G., Stumper, B., Sondag, N., & Franik, M. (2007). The acquisition of gender marking by young german-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language*, *34*(3), 445–471. https://doi.org/10.1017/S0305000906007951

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432–444. https://doi.org/10.1111/j.1532-7078.2011.00084.x

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Lawrence Erlbaum Associates, Inc.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T., & Lovelett, J. T. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, *34*(1), 39–65. https://doi.org/10.1177/0267658316675195

White, K. S., Peperkamp, S., Kirk, C., & Morgan, J. L. (2008). Rapid acquisition of phonological alternation by infants. *Cognition*, *107*(1), 238–265. https://doi.org/10.1016/j.cognition.2007.11.012

Wong, P. C. M., Ettlinger, M., & Zheng, J. (2013). Linguistic grammar learning and DRD2-TAQ-IA polymorphism. *PLoS One*, *8*(5), e64983. https://doi.org/10.1371/journal.pone.0064983

Woodward, A., Phillips, A., & Spelke, E. (1993). Infants' expectations about the motion of animate versus inanimate objects. In *Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society, Boulder, CO* (pp. 1087–1091). Lawrence Erlbaum.

Ziori, E., & Dienes, Z. (2006). Subjective measures of unconscious knowledge of concepts. *Mind & Society*, *5*(1), 105–122. https://doi.org/10.1007/s11299-006-0012-4

Zubin, D. A., & Köpcke, K.-M. (1984). Affect classification in the german gender system. *Lingua*, *63*(1), 41–96. https://doi.org/10.1016/0024-3841(84)90031-7

Appendix. Stimuli

Nonsense words from which experimental stimuli were selected.

1. One syllable words:

   [zɑb], [vəb], [sad], [fɪd], [dɛz], [gəz], [kaʃ], [tɛʃ], [dək], [kak], [gav], [bɪv], [sɑf], [bəf], [kɑs], [ʃɛs], [pɪp], [zəp], [vɛʃ], [vɑk], [fɪs], [fət], [zɪf], [zɛk], [sɑb], [səf], [bak], [bɛʃ], [paf], [pɪz], [dəf], [dɑs], [tab], [tət], [ʃɛt], [zɑt], [ʃɪk], [ʃɑp], [kɪv], [kɛb]

2. Two-syllable words (stress is always on the first syllable):

   [disɑb], [bɛpəd], [fisad], [dɛzaʃ], [bakas], [fədiz], [vogəv], [vɛbət], [fasɪp], [fɛkab], [zaʃɪb], [zitəv], [soʃək], [sɑpɪk], [pɛfɪd], [pisak], [tikəb], [tɛsəf], [ʃɑtoz], [ʃɛkɛf], [kufis], [kozɑb], [pəfɛv], [bosav], [sɑpɪf], [kusəd], [dɑzəs], [tikoz], [dɛzɪt], [ʃabɪt], [sɑpɪʃ], [fikəʃ], [zaʃat] [pɛkɪv] [bɛzaʃ] [dɑpəs] [ʃəfɪk] [təsəd] [kifək] [zɛbəz]