Morphological Development

Michael Ramscar, [1] Melody Dye, [2] James Blevins, [3] and Harald Baayen [1]

[1] Eberhard Karls Universität Tübingen

[2] Indiana University

[3] University of Cambridge

Abstract

The development of morphological processing has been the focal topic in a debate over the nature of language, learning and the mind in cognitive science. Particular attention has been paid to the systematic nature of children's morphological errors (for example children tend to go through a phase of saying "mouses" as they learn English nominal morphology). Because these errors aren't explicitly corrected, it has been argued that the transition to adult language cannot be explained by learning, and that the acquisition of even relatively simple aspects of grammar must involve innate, language specific mechanisms. We describe the background to this debate, along with some models of the learning of noun morphology based on formal learning theory that generate clear and surprising predictions. In particular, that exposure to regular plurals (e.g. *rats*) can decrease children's tendency to overregularize irregular plurals (e.g. *mouses*). We review empirical results showing that testing memory for items with regular plural labels does lead to a decrease in irregular plural overregularization in six-year-olds, yet an increase in four-year-olds. These models and results indicate that when the learning problem facing children is properly characterized, overregularization both arises and is resolved as a result of the discriminative nature of human learning systems, and the way learning responds to the distribution of evidence in the linguistic environment. The models and results we review indicate that, far than being evidence for language specific mechanisms, the behavior manifest in overregularization bears all the hallmarks of basic learning mechanisms that we share with a number of other animals. We discuss the implications of this for our broader understanding of language and learning, as well as the nature of human communication.

Morphological Development

1. Introduction

Language, whether spoken or written, is the primary means of human communication. Yet exactly *how* language facilitates communication has yet to be explained. This chapter describes our approach to explaining how morphological systems work and what morphological development entails. The approach is rooted in the way minds learn, and is based on clearly and explicitly stated learning mechanisms for which there is a wealth of biological evidence; It is consistent with the principles that govern artificial communication systems (Shannon, 1948); And, unusually, it makes surprising – and successful – predictions about the pattern of morphological development.

Because it is rooted in the discriminative principles of learning, the perspective of language our approach offers is very different to the associative, combinatoric view taken by most researchers. However, we believe its origins make it likely that the perspective it provides will prove to be of fundamental importance to understanding human communication, and consequently the challenges facing children with communicative disorders. Accordingly, in this chapter we describe the principles of learning in detail, along the picture of morphological and linguistic development they give rise to.

2. Language, morphology and development

In thinking about how human communication works, linguists have typically assumed that language facilitates communication by conveying meanings (much as trains convey passengers; Reddy, 1979). Linguistic theories assume that words (and the sub-word units called morphemes) encode (i.e., are associated with) units of meaning, such that the word 'units' is composed from two morphemes (the morpheme 'unit' and the morpheme 's', and the word 'morphemes' composed from three ('morph,' 'eme' and 's').

From this perspective, the task facing a language learner involves learning what the conceptual units of her language are, learning how to associate them with sound units to create morphemes, and figuring out the kinds of morphemes that can be combined to form complex words, along with how sequences of morphemic combinations combine to yield more complex sentences.

Over a century of study has uncovered a number of problems with this approach. Critically, at both a behavioral and neural level, it has been found that learning doesn't work in the associative manner that linguistic theories are wont to imagine. What has traditionally been called "associative" learning is a systemic process that serves largely to *discriminate* rather than associate the details of a learner's internal representation of the world (Rescorla & Wagner, 1972; Ramscar et al, 2010). Further, while humans are perfectly capable of learning to discriminate between events and behaviors, they do so in ways that do not involve the discrete internal representations (i.e., the "units of meaning") that are supposed to provide the stock of combinable elements in combinatorial approaches to language (Wittgenstein, 1953; Ramscar & Port, 2015). Finally (though by no means exhaustively), combinatorial approaches to language are wont to describe meanings as being encoded into verbal and written signals by speakers and writers to later be decoded by readers and listeners: However, the process this envisages violates the basic principles of coding theory, since it assumes that the appropriate meanings of words or morphemes with many potential "senses" can be successfully decoded from signals that do not actually code for these senses. In fact, coding theory is founded on the principle that each discriminable aspect of a message must be encoded in a signal (Shannon, 1948; in the case of written English, the available coding resources do not even consistently discriminate different lexical forms from one another, i.e., the past and present tenses of *read*).

Although most cognitive scientists are familiar with at least some of these problems, most theories of language acquisition and processing remain steadfastly rooted in associative combinatorics, even in the absence of an adequate account of what exactly gets combined, or how the encoding and decoding of meanings is actually supposed to work. In what follows, we describe an account of language learning and processing based on *discrimination learning*. It does not assume that morphemes serve to convey meanings, but rather supposes that they serve to discriminate between meaningful states of affairs in communication. From this perspective the task facing a language learner is that of learning which potential states of the world morphemes may discriminate, and how the distinctions afforded by morphemes can be used in communication. To illustrate the utility of this approach we show how it not only provides a satisfactory account of many phenomena associated with morphological development, but also makes surprising (and successful) predictions about the way that patterns of morphological over-generalization develop and recede.

3. How does a child's minds learn?

Any account of how children master morphological processing seeks to answer two questions: *what* do children learn, and *how* do they learn it? In the recent past, a great deal of linguistic theorizing has proceeded from *what* to *how*. Perhaps unsurprisingly, these theories have struggled to explain *how* the *what* of their theoretical postulates are learned. This has led in turn to a situation where the psychological bases of many theories of language are opaque, and where claims about "innate mechanisms" abound.

We take the opposite tack: starting from how children learn, we consider how this might constrain our conjectures about what children learn. One potential benefit of this approach is that our current scientific understanding of learning is far more advanced than that of language. We share many of our basic learning mechanisms with other animals, and in many domains, animal models have proved invaluable in illuminating the biological and neural structures underlying these mechanisms.

Legend has it that Ivan Pavlov's famous discovery – that ringing a bell before giving dogs food later caused them to salivate whenever the bell sounded – was a felicitous accident.[1] Whatever the exact truth of the matter, his discoveries have given rise to a popular idea: That animals learn to "associate" unrelated events according to the frequency with which a stimulus (a bell) and a response (salivation) are paired. Empirically, however, this naïve view of conditioning as a process that simply tracks co-occurences has long been known to be wrong (Rescorla 1988), as

[1] The bell is actually a myth (Todes, 2014).

have two popular – yet equally false – beliefs about the conditions that produce learning: First, that explicit rewards and punishments are necessary for learning; and second that a co-occurrence between a stimulus and a response is sufficient for learning.

Empirically, mere association cannot account for conditioning. For example, if a group of rats is trained on a *Tone-Shock, Tone-Shock, Tone-Shock, Tone-Shock, etc.,* schedule, they will learn to associate a tone with a mild electric shock, and freeze when later tones sound. However, if rats are exposed to an identical number of tone-shock pairings into which a number of tones not followed by shocks are interpolated (i.e., *Tone, Tone, Tone-Shock, Tone-Shock, Tone, Tone, Tone, Tone, Tone-Shock, Tone, Tone, Tone, Tone, Tone-Shock… etc*) they show a different pattern of learning. As the number of tones not followed by shocks increases, tone-shock associations decrease proportionally (Rescorla, 1968).

Given that the *non*-occurrence of shocks after tones influences the degree to which rats condition to tones, it follows that learning must comrpise more than simple counts of positive occurrences of cues with events. To explain how rats learn from the *background rate* of the tones (a process that popular "associative" conceptions of learning cannot explain; Rescorla 1988), modern learning theories suppose that in the two situations just described, the "no shock" trials act to alter rats' implicit expectations about the tones. These theories conceive of learning as a process that serves to reduce uncertainty in the predictive mental models that animals construct out of their experience with their environments (Rescorla 1988). Functionally, learning modulates the value that animals implicitly assign to the usefulness of sensory cues in predicting the events that animals experience. As events unfold, animals' representations of cue values change as a function of the current expectations an animal has learned to form, and the degree to which these expectations led it to anticipate (or fail to anticipate) what actually occurs.

This results in a competitive mechanism in which learning to predict an outcome from one cue has an impact on the uncertainty that drives the learning of other cues. This is best illustrated by *blocking* (Kamin 1969): If a rat learns it will be shocked upon hearing a tone, and then later a light is paired with the tone, learning about the light as a cue to the shock will be inhibited – and thus the rat is unlikely to freeze in response to the light – because the tone already eliminates uncertainty about the shock.

These and other results demonstrate that rats do not learn simple "associations" between stimuli and responses, but rather learn the degree to which cues are systematically informative about events, a process that discriminates cues that are more informative from cues that are less so (Rescorla 1988). Since there are invariably far more uninformative coincidences in the environment than informative ones, it follows that expectations that are wrong have more influence on the shape of learning than expectations that are right (which is why discrimination learning is often described as being *error-driven*).

Finally, we should stress that although we have focused our discussion on the tones, in principle, *everything* in the rat's local environment can potentially influence learning (Rescorla 1988). However, in the same way that rats learn to discount tones as predictive cues the more they appear absent shocks, they also learn to discount other aspects of an experimental environment that have high background rates relative to a relevant outcome. For the sake of simplicity, models and explanations tend to focus solely on potentially informative cues, ignoring cues whose high background rates are likely to render them largely irrelevant in competitive terms. However, it is important to stress that in principle the novelty of a cue is entirely relative, and can only be computed in relation to the other available cues, and a learner's previous experience (Rescorla 1988).

Beginning with Recorla & Wagner (1972) many formal models employing these principles have been devised to fit and predict animal learning effects. Moreover, these models are also routinely (and successfully) applied in the study of decision-making, learning, and response selection in humans (see e.g., Montague et al. 2004; Niv 2009, Schultz, 2010). Experimental results provide reason to believe that children are sensitive to background rates in language learning (Ramscar et al. 2010, 2011, Ramscar at al 2013a),[2] while other studies show that animal models

[2] Although Saffran et al (1996) claim their "statistical learning" results are incompatible with Rescorla-Wagner learning, given a flat cue structure like a stream of syllables the model actually

successfully predict patterns of development in set-size learning (*subitization*; Ramscar et al 2011), and rule-understanding in children (Ramscar et al. 2013b), as well as changing patterns of performance in simple "associative learning" tasks across the lifespan (Ramscar et al 2013c; Ramscar et al 2014). Indeed, the Rescorla-Wagner model is arguably the best-supported model in psychology, with a sizeable (and still growing) body of behavioral evidence arguing in support of its principles (for reviews, see Miller, Barnet, & Grahame 1995, Siegel & Allan 1996; Ramscar et al, 2010; Ramscar et al 2013d).

Given this, and given our aim of using *how* learning works to uncover *what* is learned in morphological development, it is worth noting that the logic of discrimination enshrined in animal learning models suggests that, far from the "blooming, buzzing confusion" of multiple entities described in much of developmental psychology, a newborn learner is best conceptualized as entering the world with a large, undifferentiated set of cues connected to little or no environmental knowledge. Starting from N=1, the set of (more or less individuated) entities in the learner's representation of the world then begins to expand as perceptible variances and invariances in the environment encourage discrimination learning (James, 1890).

4. What is morphology, and how might it be learned?

4.1 The combinatoric approach

Having established in broad terms *how* children are likely to learn morphological systems, we now turn out attention to *what* they learn. That is, we can now consider whether these mechanisms are sufficient to account for morphological development, and if so, how.

We ought to acknowledge here that interest in morphological development extends <u>far</u> beyond the simple concern of understanding, say, plural marking, for its own sake. Over the course of the past quarter century, research on morphological development has been seen "as addressing some of the most important issues in cognitive science" (Seidenberg & Plaut, 2014, p1), largely because:

---

asymptotes at cue weights that approximate the transitional probability between each syllable (Ramscar et al, 2010), which behaviorally is what Saffran et al observed in infants.

> [Morphology has…] three interesting characteristics. First, it is systematic: Most past tenses are formed by adding the morpheme that is spelled -ed and pronounced as in the examples baked, baited, and bared. Second, it is productive: People can readily generate past tenses for novel forms such as nust-nusted or wug-wugged. Third, it is quasiregular (Seidenberg & McClelland, 1989): There is a main pattern but also irregular forms that deviate from it in differing degrees (e.g., keep-kept, run-ran, go-went). Phenomena such as tense on verbs and number on nouns have been taken as simple, decisive demonstrations that grammatical rules are an essential component of linguistic knowledge (Pinker, 1999). Irregular forms exist outside this system of core linguistic knowledge and are learned and generated by other mechanisms such as memorization and association.

Rumelhart and McClelland's (1986) model offered an alternative view of this last point. Taking the phonological form of a verb's present tense as input, it generated the phonological form of its past tense as output using a uniform procedure for all tenses. It also supported the generation of past tense forms for novel verbs. The model and the various claims made about it caused controversy and launched a debate that has generated an enormous body of research on a range of morphological phenomena.

For our current purposes, however, the agreements in this debate are more relevant than its disagreements: Almost all the participants in this debate accept that the *what* of morphological development is a means of composing and decomposing associative morphemes. Hence, in the case of English plurals, they assume that a child learns a morpheme that associates the concept mouse with the word *mouse*, an association between the concept mice and *mice*, an association between the concept rat and *rat*, and an association between the concept for sets of objects (excluding multiple mouses) and a morphemic gesture characterized linguistically as a terminal sibilant on a noun, and often written as +*s*, etc.

Yet, as we described above, the *what* of the "associative" learning process has actually been found to be **discriminative**. The "association-tracking" at the heart of the processes imagined in this debate is actually an evolutionary twist on Sherlock Holmes' dictum: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth." When a rat learns to associate a tone with a shock, the "association" is what is left over after learning has systematically weighed every other potential source of information and found it wanting.

Consistent with this picture of learning, researchers studying human categorization have found that at a behavioral and neurological level, human performance is best accounted for by models that don't actually contain pre-established (or even determinable) categories, but instead treat categorization as an active process of discriminating between more or less appropriate category labels (or other affordances and behaviors) in context (see Ramscar & Port, 2015 for a review). Similarly, the non-discrete nature of meanings could be described as the closest thing that philosophy that has to offer to a fact (Wittgenstein, 1953;Quine, 1960). All of which indicates that the idea of the "associative morpheme" – a discrete mapping between a discrete unit of meaning and a discrete linguistic unit – is incompatible with what we know about human learning, about the nature of meaning, and about the computational properties of the human categorization system.

What is more, if we accept the message of this evidence, and allow that morphological systems do not involve a discrete set of mappings between a set of units of meaning and the lexical forms of a language – which we term lexemes – then a significant peculiarity of this debate – and indeed, of the way linguists think about morphology more generally – becomes apparent. In the Rumelhart and McClelland (1986) model (and most subsequent models of morphological development) the child is envisioned as learning a combinatorial, transformational rule that, for example, adds the English past tense morpheme +ed to a verb stem in order to transform an uninflected form into a past tense form. Accordingly, the model's training set was devised to teach this transformation: uninflected forms are repeatedly turned into past tense forms, as if the learning environment comprised mature speakers going about saying, *"go-went, walk-walked, speak-spoke, talk-talked, etc."*[3]

However, this characterization of the learning environment is dubious in the extreme. Mature speakers do not wander around giving extensive tutorials on the nature of supposed transformations, saying, *"go-went, walk-walked, etc.*" They talk about what interests them, in context. This means that children rarely, if ever, hear *"go"* then *"went*" in close proximity. Instead, they hear "Wanna go see a movie?" on Tuesday, and "Oh no! I forgot to put the trash out before we went to the movies last night," on Wednesday.

In other words: 1., Children don't learn in the way envisaged by the combinatoric account of morphology; 2., The units it supposes that they learn are highly implausible; And 3., Children learn morphological systems without actually ever actually encountering morphemes being transformatively. Thus it seems possible, perhaps even likely, that a different approach to the conceptualization of morphological processing might yield a better account of *what* develops, as well as *how*.

4.2 The counter-intuitive appeal of discriminative morphology

We begin our formulation of this alternative account by considering the very things that the combinatoric story gets wrong: the nature of learning, human knowledge representation, and the learning environment. As we noted above: learning is discriminative; categorization is an active, context driven process that serves to discriminates between lexemes (and other more or less discrete affordances and behaviors); and children hear lexemes used in contexts that offer precious little evidence for transformations at all.

To illustrate how these constraints might influence what a child actually learn givens these constraints, we will take English nominal morphology as an example, and consider how these factors might give rise to the typical patterns of over-regularization children exhibit while learning the nominal system (a fuller

---

[3] Although Rumelhart and McClelland (1986) use a discriminative learning rule analogous to Rescorla-Wagner, they describe their model as a pattern associator, and adopt a combinatoric view of morphology.

description of this approach, along with code to produce the simulations described below, can be found in Ramscar et al 2013d).

In English, correct irregular plural marking is particularly difficult to acquire (Ramscar & Dye 2011), even in comparison to the more commonly studied case of past tense marking. This reflects the nature of the input. For verbs, while irregulars are rare as types, they tend to have high token frequencies, such that the 40 most frequent verb forms are all irregular (Davies 2009). Moreover, in the Reuters Corpus (Rose et al 2002) just three irregular verbs (*be*, *have* and *do*) account for fully a quarter of the attested verbs forms, with past tense verb forms outnumbering base or present tense verb forms. Thus, in learning past tense inflection, children are likely to encounter more past inflected than uninflected verb forms, and of these, more irregular than regular past inflections.

Plurals are different: Children mainly encounter singular forms, and when they do encounter plurals, they are likely to be regular. In the Reuters Corpus, only around 30% of nouns occur in their plural form, and of these, the overwhelming majority of types and tokens are regular. While this makes the learning problem for plurals substantively more difficult than the past tense, the two problems may not be different in kind: As with the past tense, children's irregular plural production follows a 'U-shaped' developmental trajectory, such that children who have been observed to produce "mice" in one context still frequently produce overregularized forms such as "mouses" in another (Arnon & Clark 2011).

Besides assuming that native languages are not learned in formal teacher-pupil settings, but by hearing them used in context, our account of nominal development also assumes that lexical and morphological systems are systematically informative, such that the combined value of both positive and negative evidence favor a mapping between the experiences a child has with the class of objects described by a given noun, and the noun itself. For example, a child learning the lexeme "mice", will hear the word used in a way that makes it most informative about mice, or depictions of them, and must learn to associate the appropriate cues in the environment—mouse-things—with the lexeme (Quine, 1960, Wittgenstein, 1953). Conceptually, this assumption reflects the idea that adult speakers use language in informative ways, and hence, that a mouse ought to be more informative about the English lexeme “mouse”, and mice more informative about the lexeme “mice”, than vice versa (otherwise, mice could equally mean mouse, and mouse could equally mean mice). Accordingly, when a child is asked to name a picture of mice, the child is able to say *mice*, because learning has discriminated a set of mappings between the mice-relevant semantic dimensions of the child’s experience and the more or less discrete gestural/phonetic form *mice*.

At the outset of word learning, all and any kind of “stuff” in the world will seem potentially informative about concrete nouns such as mouse and mice. Thus learning to discriminate the correct cues to each form will involve discriminating the “particular stuff” that is best associated with any given singular and plural form of a noun (e.g., mousiness in the case of *mice* and *mouse*) from other kinds of stuff. At the same time, learning to discriminate *singulars* from *plurals* will require discriminating the specific stuff that best predicts a singular as opposed to a plural (i.e. the presence of multiple mouse objects ought to cue *mice,* as opposed to a single mouse object which ought to cue mouse).

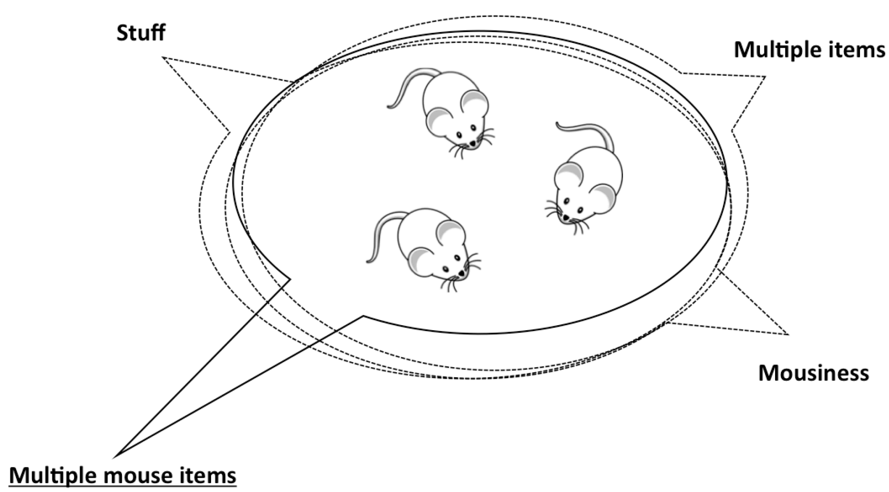

Figure 1: Four semantic-dimensions that will be consistently reinforced by a child’s exposure to the word “mice” in the context of mice. (While these dimensions are separated for explanatory clarity, as far as the learning mechanisms are concerned, they could equally be values on continuous perceptual dimensions.)

To illustrate how these distributional differences influence the learning of English nominal morphology, Figure 1 depicts four environmental dimensions that will reliably and consistently covary with the irregular plural form “mice.” Although these semantic dimensions will co-occur with the word “mice” at the same rate, their covariance with other singular and plural nouns will differ. Because of these differences in background rate, the error associated with each dimension will vary in kind. Accordingly, although early in learning, generic cues like ‘stuff’ will receive positive reinforcement when “mice” is encountered in context, their ubiquity will also cause them to produce a high degree of error. Indeed, compared to more uniquely informative cues, they will occur far more often in contexts in which “mice” is not heard. Thus, the influence of the more generic cues will wane over the course of learning, as learners converge on multiple mouse- items as the best cue to “mice.”

In learning, expectations errors arises as a function of experience, which means that the pattern of reinforcement and unlearning of the semantic dimensions in Figure 1 depends heavily on their distribution in the learning environment. While the set of singular and plural lexemes that are usually classed together and called ‘regular’ plurals in English are distinguished by a number of subtle differences (such as different sibilant allomorphs), broadly speaking, regular plurals are far less discriminable from one another than irregular plurals, particularly in terms of how they discriminate plurality from singularity. In regular plural lexemes, plurality is uniformly denoted by the presence of some form of final sibilant. By contrast, irregulars employ a variety of means of discriminating singular and plural forms. This guarantees that irregular plural lexemes are, at once, less similar to their singular forms than regular lexemes (e.g., dog/dogs vs. child/children), and are also less similar to other irregulars (foot-feet vs. child children).

At this point, it is worth expanding a little on what we mean by a *lexeme*: In the model of learning we describe here, a child is seen as learning to discriminate the semantic (and lexical) cues to a <u>system</u> of phonetic and lexical contrasts simultaneously in context (Ramscar, Dye & Klein, 2013; Ramscar & Port, 2015). As such, the degree to which any given phonetic lexical contrast has itself been discriminated will depend very much on the current state of the learner. Accordingly, the discrete lexemes we describe here should be taken to represent a simplification for descriptive and modeling purposes of what we assume to be a far more continuous system in reality. Accordingly, for example, it might be helpful for some theoretical purposes to represent the state of the plural system mastered by a young learner as comprising irregular forms, regular stems and a single lexeme marking the regular plural contrast +*s*, whereas in other circumstances, the regular plural contrast might be more appropriately by positing different lexemes for each regular plural allomorph. In other circumstances, for example, in modeling the adult ability to discriminate and respond to prosodic differences in the stems of regular forms (Baayen et al.. 2003, Kemps et al, 2005), well learned regular plural and singular forms might best be modeled individuated lexemes.

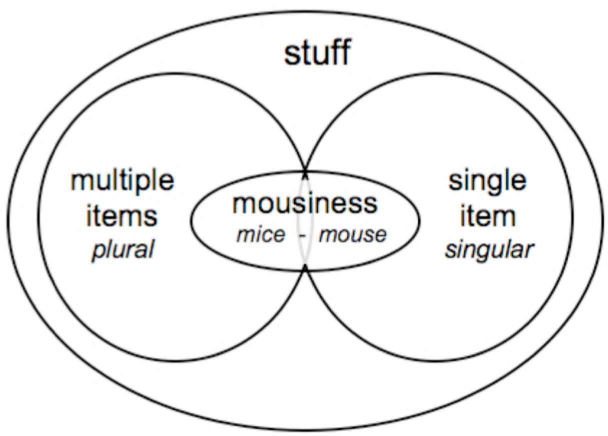

Figure 2: The relative specificity of the four dimensions as cues to plural forms. While the less specific cues ('stuff' and 'mousiness') will receive positive reinforcement early in learning, their ubiquity will cause them to produce more error than the uniquely informative cues. As a result, the influence of the less specific cues will wane as experience grows.

Figure 2 shows how the various potential semantic cues to *mice* overlap relative to a fairly simple set of lexemes (comprising irregular forms, regular stems and a single lexeme marking the regular plural contrast +*s*) in an idealized young learner, and illustrates why the structure of the lexicon necessitates that learning to correctly produce "mice" will require unlearning a number of irrelevant semantic cues. As a child learns from this distribution of lexemes, the semantic dimension "multiple items" will initially be reinforced whenever plural forms are used to talk about any sets of object in context. And in this distribution, most noun forms employ largely the same phonetic form to denote kind semantics (e.g., *rat* is strongly associated with rat semantics). It is the presence or absence of a terminal sibilant (+*s*) that then discriminates between singularity and plurality when the largely common noun stem is used. This means that initially, whenever a set of objects is to be described, the child's language model will predict that a form should have a final sibilant.

Meanwhile, whenever mouse objects are talked about, the relationship between "multiple mouse objects" and *mice* will be reinforced as well as "multiple items." This means that in other contexts where plurals are used, "multiple items" will not only lead the child to implicitly expect a final sibilant, it will also cause *mice* to be expected. When the child does not hear *mice*, this will lead to error, causing the relationship between "multiple items" and *mice* to be downgraded in the child's language model. Similarly, whenever *mice* are talked about, the child's internal representation of the relationship between "multiple items" and a terminal sibilant will be downgraded. Over time this will gradually increase the association between "multiple mouse objects" and *mice* (in mice contexts) while decreasing the association between "multiple objects" and *mice* (in non-mice contexts).

So far we have discussed what the expectations of a "naïve" observational learner are. How might we expect this to influence production? Here we can envisage two scenarios: In the first, the learning simply reinforces the behavior a child produces, so that when the noise in her underlying model of the world results in her saying "mouses," mouses is reinforced as a response. In the second, production is also driven by a child's model of her intended behavior (what she has observed), but attempts to perform a behavior reinforce the underlying model, not whatever noisy behavior that emerges from it (Gläscher et al, 2010; model-based learning allows one to explain why practice can lead ballroom dancers towards the right steps, rather than reinforcing treading on toes).

Behaviorally, of course, we know English speakers often do go through a period of saying "mouses" in childhood, but as they grow older, they come to produce only the adult form, “mice”. In the model we have described, we assume that this reflects an interim state that arises from the distributional properties of the learning environment, and is resolved by further sampling from this distribution. Which leads us to an advantage of models: they generate unambiguous formal predictions.

Ramscar & Yareltt (2007) present a computational simulation which reinforces a model of the world rather than the overt behavior it would generate, and which predicts that eliciting over-regularized forms from children will actually cause them to over-regularize less. In a series of experiments, Ramscar & Yareltt then showed that children exhibited the same behavior: when seven-year-old children repeatedly produced the same plurals across blocks of trials, their rates of over-regularization went down in later blocks. That this occurred even where children were given positive feedback on the incorrect forms that they produced further supports the idea that learning reinforces children’s models of the world, and not necessarily behavior.

Ramscar et al (2013d) show that when the challenge facing a child language learner is explicitly set up in the way shown in Figures 1 and 2, the distribution of forms and semantics in English inevitably leads to what has been described as “U-shaped” performance in plural production: Mastery of correct irregular forms is preceded by a phase when both correct and incorrect irregular plurals are produced. Moreover, Ramscar et al show that in this model, the ultimate elimination of the interference from sibilant final forms that leads to over-regularization is driven by error caused by the expectation of *mice* when the semantics of **regular** forms are experienced in a lexical context. That is, as described earlier, the same non-discriminative semantic dimension that causes children to expect a sibilant final form in an irregular context – leading to over-regularization – causes them to expect irregular forms in regular contexts, which gradually causes the non-discriminative semantic cues to be unlearned as cues to irregulars, reducing over-regularization.

This results in an unambiguous prediction: engaging children at an appropriate stage of development in a task invoking the semantics of regular forms in a lexical context ought to bring about a reduction in over-regularization. To test this, Ramscar et al first pretested children on a task that elicited both regular and irregular plural and singular forms, and then one group of children then performed a color related task to which the semantics of regular forms are irrelevant, and the other performed a memory task in which the same regular plural forms from the elicitation task along with regular lures (i.e., a task that depends upon the lexicon – see Ramscar & Port, 2015 – in which that the semantics of regular forms could confidently be expected to be experienced). The children were then post tested using the elicitation task.

Consistent with the detailed predictions of the model, the color task had no effect on over-regularization. By contrast, younger children tested in the memory condition showed a small but significant increase in over-regularization. However, the same exposure to the semantics of regular plural forms brought about a large and significant **decrease** in over-regularization in the older children (Ramscar et al 2013a), just as the model predicts. In other words, the pattern of children's over-regularization and their retreat from it is exactly as one would predict given the way that they learn, the forms that they learn, and the way that these forms are distributed semantically.

5. The scientific appeal of discrimination learning

Because children reliably go through a period in which they over-regularize – saying mouses – and because they reliably stop doing this without encountering any explicit instruction to do so, it has often been argued that their behavior presents a logical puzzle: why *do* they stop (Pinker 1984, 2004; Pinker & Prince 1988)?

The models and the results we have described show that the patterns of over-regularization behavior exhibited by developing English speakers are not puzzling. Over-regularization is caused by the distributional properties of the English learning environment, and in typically developing children, it resolves itself without any need for explicit correction because the same processes and circumstances that give rise to the problem ultimately resolve as sampling and learning increase. (Indeed, because learning reinforces children's models of the world, explicit feedback about behavior often has little obvious effect on what children actually learn; Ramscar & Yarlett, 2007.)

In the light of the debate that has surrounded this phenomenon, it seems worth noting that our explanation does not "solve" the puzzle of over-regularization as it has previously been posed. There are two reasons for this: First, because, insofar as this puzzle has been framed as a "logical problem", it is only actually a problem if one ignores some well-established facts about animal learning (Rescorla & Wagner, 1972), or if one assumes that infants are somehow less capable learners than the animals these principles were derived from – namely, rats (Pinker, 2004).

And second, because once one understands how learning works, it becomes clear that the way the problem has been posed in the past makes little sense. Learning is not a process that "associates" "units" of forms and meanings in the way that morphologists have traditionally imagined (Ramscar & Port, 2015), but is instead a discriminative process in which systems of associations are learned implicitly as the result of the process of dissociating anything and everything else. Learning has evolved to allow our selves and other animals to make sense of the world by reducing its dimensionality and complexity in order to highlight what is relevant (Trimmer et al, 2012).

While it is not always easy to intuit exactly how this process allows us to map our semantic model of the world onto the system of forms in a language, it is possible to model this complex process computationally, and gain an understanding in the way illustrated above. By contrast, not only does it make little sense from a learning perspective to imagine that language learners are faced with the task of acquiring rules that transform or agglutinate "form units" that map to "meaning units" as morphologists have traditionally imagined, but, as the numerous claims that have been forward for this or that aspect of linguistic processing being innate over recent years would seem to attest, conceiving of how this kind of combinatoric system actually works in any kind of detail appears to be all but impossible.

6. Discrimination learning and implicit morphology

It is often assumed that regularity is a desirable or normative goal for morphological systems, and that irregular paradigms represent deviations from the uniform patterns that systems (or their speakers) strive to maintain. Such an assumption is challenged, however, by phenomena like suppletion, in which an inflected stem-change produces a phonologically unrelated allomorph (e.g., mouse/ mice), rendering patterns of form-meaning mappings unpredictable. A discriminative perspective makes precisely the opposite assumption).

In the discriminative models we have described, the difference between overtly suppletive forms such as mice/mouse and more regular forms such as rat/rats is that the former serve to accelerate the rate at which a speakers' representation of a specific form/meaning contrast becomes discriminated from the form classes that express similar contrasts. The logic of these models is that all learning serves to increase the level of suppletion in a system of form-meaning mappings.

From this perspective, 'suppletive' irregular forms like *mice* are not categorically different types, as morphologists have been wont to imagine, but are rather merely extreme instances of the system of discriminative contrasts that linguistic communication relies on. Moreover, when one examines language use from this perspective, it is clear that these systematic contrasts are ubiquitous at the sub-phonemic level. Thus for example, while combinatorial inuitions have misled linguistics in assuming that the forms of language are constructed out of an alphabet of phones (Port & Leary, 2005), it has become clear from studies that, for example, the duration and fundamental frequency of the 'cap' in *captain* differs systematically as compared to the morphologically unrelated *cap*. Moreover, analyses have shown that the nature of linguistic contrasts is such that even so-called homophones – like *time* and thyme – appear to be gesturally and acoustically distinct (Gahl, 2008).

Consistent with the models we have described, the patterns of contrast observed in the lexicon are also observable at a morphological level: Baayen et al. (2003) found that a sample of speakers produced Dutch nouns with a longer mean duration when they occurred as singulars than when they occurred as the stem of the corresponding plural. Kemps et al. (2005) show that speakers are sensitive to these prosodic differences, finding that "acoustic differences exist between uninflected and inflected forms and that listeners are sensitive to them" (Kemps et al. 2005:

441). Plag et al. (2014) observed similar contrasts and sensitivities to them in a study of phonemically identical affixes in English.

It thus follows that from a discriminative perspective, it is the regularity in morphological systems that stands in need of explanation. Discrimination learning suggests a solution here as well. Unlike derivational processes, inflectional processes are traditionally assumed to be highly productive, defining uniform paradigms within a given class. Lemma size is thus not expected to vary, except where forms are unavailable due to paradigm 'gaps' or 'defectiveness'. However, corpus studies suggest that this expectation is an idealization. Many potentially available inflected forms are unattested in corpora, and as corpus sizes increase, sampling does not converge on uniformly populated paradigms, but rather they reinforce classes and develop longer tails (Blevins et al, 2015).

In order for a collection of partial samples to allow the generation of unattested forms, the forms that speakers do know must be organized into systematic structures that collectively enable the scope of possible variations to be realized. These structures thus correspond to lexical neigbourhoods, whose effects have been investigated in a wide range of psycholinguistic studies (Baayen et al. 2006; Gahl et al. 2011). From the present perspective, these neighbourhoods are not independent dimensions of lexical organization. Instead they constitute the creative engine of the morphological system, permitting the extrapolation of the full system from partial patterns. Regular paradigms thus enable language users to generate previously unencountered forms in the same way as we describe the learning and workings of regular plural morphology in the model above: Regular forms are not the product of an explicit rule, or of any kind of explicit grammatical knowledge, but rather they are *implicit* in the distribution of forms and semantics in the language as a system.

7. Discriminative language learning

Traditionally, linguistic and social learning have often painted as being in opposition to "mere associative learning" (e.g., Tomasello, 2003), and most scientists and practitioners likely still believe this is the case. However, these contrasts ultimately rely on a very faulty understanding of what associative learning actually is (Rescorla, 1988). As we described above, associative learning is in fact a discriminative process, and critically, this process is also inherently systematic. While the systematicity of learning is simple to state, the explanatory tools it provides are both subtle and powerful. For example Ramscar et al (2013c) have shown that changes across the lifespan in adults' ability to learn the association of arbitrary pairs of words such as *jury* and *eagle* across, at first glance an almost prima facie example of a combinatorial process, can be accurately modeled and predicted if this is conceived of in terms of more systematic discrimination learning in the lexicon. Critically, it has been shown empirically that learning frequently co-occurring pairs (*lock–door*) differs little with age, whereas learning pairs like jury-eagle becomes increasingly difficult. From a discriminative perspective, the fact that the latter co-occur extremely rarely causes them to become negatively associated by learning in the lexical system as a whole, and it is this that causes them to become harder to associate as experience increases. That is, although intuitively learning to pair *jury* and *eagle* appears to be a combinatoric process, it turns out that actual behavior in the task is best explained modeling the lexicon as a heavily interconnected system in which all lexical items are related by complex patterns of co-occurrence, and in which learning causes items to become associated or dissociated from one another as experience increases (i.e., in the same way as we treated morphology in our model above). When paired associate learning is modeled as a process that serves to reduce and ultimately reverse the systematic dissociations that ordinary experience induces adults to learn, i.e., experience in which *jury* is not informative about *eagle*, the changing ability of adults to learn to pair of *jury* and *eagle* at various ages an be predicted with surprising accuracy (Ramscar et al 2013c).

Although this view of learning and language is rather more complex that our intuitions about learning tend to suppose, as the model we described above illustrates, by highlighting the systematic nature of language learning, it sheds new and productive light on the many apparent puzzles that arise when the learning of aspects of linguistic systems are considered in isolation. It is because of this that we believe that ultimately this approach will yield fruitful methods for understanding how and when language learning goes awry: As Quine (1960) noted, learning language requires that a child master not only the relationship between a system of conventionalized sound and meanings relationships, but also that the child learn how to use this system to communicate. Being able to do so appears to hinge on learning to share subjectivity; the child must somehow learn to comprehend the shared point of view of her community (see also Tomasello, 2003; Wittgenstein, 1953).

Exactly how human infants come to discriminate the "intersubjectively available cues as to what to say and when" (Quine, 1960) is an incredibly complex task. A discriminative account of communication grounded in learning theory allows us to frame questions about the way that children learn the sets of conventionalized cues that underpin languages (Wittgenstein, 1953) in ways that are tractable and amenable even to formal description. For example, if communicative conventions—and language—are the product of learning, why is language apparently solely the preserve of humans? What is *special* about human learners?

One part of the answer to this question appears to lie in the difference between the way that the brains of humans and other animals develop, and its impact on learning. Like many other primates, humans are born with immature brains. Birth is followed by synaptogenesis (the proliferation of synapses) followed by an extended pruning period (synaptic elimination). Brain development in humans, however, is markedly different from that of other primates. In monkeys, the postnatal development of the brain occurs at the same rate in all cortical areas. In contrast, human cortical development is uneven: Synaptogenesis in the visual and auditory cortex peaks a few months after birth, while the same developments occur later in the prefrontal cortex, which doesn't fully mature until late adolescence (Ramscar & Gitcho, 2007; Thompson-Schill, Ramscar, & Evangelia, 2009).

One important behavioral consequence of delayed prefrontal development is young children's inability to select behaviors that conflict with prepotent responses (Ramscar, et al 2013b). In adults, prefrontal control mechanisms bias responses and attention according to goals or context, selectively maintaining task-relevant information and discarding task-irrelevant information (Ramscar & Gitcho, 2007;

Thompson-Schill, Ramscar, & Evangelia, 2009). The absence of this capacity in young children can be illustrated by contrasting their performance with that of adults on biased selection tasks, such as guessing the hand an M&M is in (the hands are biased 25:75). Children up to age 5 tend to overmatch, fixating on the high-probability ‘‘good’’ hand. After age 5, however, a probability matching strategy emerges (Derks & Paclisanu, 1967). This is a rare instance in which children’s inability to think flexibly is an advantage (probability matching actually reduces the number of M&Ms won by children over 5 years old and adults).

Another area of learning in which cognitive flexibility may well prove disadvantageous is in the process of learning even a “simple” morphological system like that of English noun inflection. Linguistic knowledge is, in its essence, conventional. In the presence of a linguistic cue, a social animal needs to be able to understand or response appropriately given the context context. For this to happen, linguistic signals, must be both conventionalized and internalized (Wittgenstein, 1953). Learning the system cues that yields appropriate understanding is far more likely to happen if learners are unable to filter their attention during the course of learning. Given a similar set of cues and labels to learn, learners will tend to sample the environment in much the same way, and thus it is more likely that they will come to have similar expectations regarding the relationship between cues and symbols.

In contrast to children, adults struggle to master linguistic conventions (including English noun morphology; Johnson & Newport, 1989). This may reflect an inevitable handicap brought about by their increased ability to selectively respond or attend to the world and cues in it. Because development increases the complexity of the human learning architecture, allowing learners to filter their attention in learning, it is likely that it also dramatically reduces learner’s ability to learn conventions by naively sampling in the way we described above (Ramscar & Gitcho, 2007; Thompson-Schill et al., 2009). Put simply, the greater variety there is in what adults attend to during learning, the less conventionality there will be in what adults learn. Conversely, the less that children are able to direct their attention in learning, the more what they learn will be shaped by their immediate physical, social, and linguistic environment, and the more their learning about common regularities in that environment will be conventionalized (see also Finn et al, 2013, 2014; Hudson Kam & Newport, 2005, 2009; Singleton & Newport, 2004).

Although the Rescorla-Wagner model is a useful model of at least some of the brains discriminative learning processes, it is clear that there is far more to learning than error monitoring. Humans, especially adult humans, are not the passive observers of the environment that Rescorla-Wagner idealizes them to be. Our understanding of exactly how attention and learning trade off against one another, and how this affects what gets learned as frontal regions mature is, as yet, in its infancy. However, the learning and control processes we describe are amenable to computational and biological modeling, and progress is being made in this regard (Ramscar et al, 2013c). Yet, to emphasize the point that idealized model can be useful even when incomplete, consider that Triesch, Teuscher, Deák, & Carlson (2006) have shown how gaze following emerges naturally from the learning mechanisms we have described here, provided that a child has access to a caregiver that tends to look at things in a ways the infant finds informative. It is clear a model such as this is far more likely to lead to an understanding of how this kind of social development goes wrong, and how it impacts on other aspects of learning, than the many current theories that propose that gaze following is the result of an unspecified innate mechanism. As Box & Draper (1987) famously noted, while models are always wrong in the limit, they can still be very useful. We hope that more of our colleagues will be inspired by this approach in the future.

7. References

Arnon, I., & Clark, E. V. (2011). Children's word production is facilitated in familiar sentence-frames. Language Learning and Development, 7(2), 107-129.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. Journal of Memory and Language, 55(2), 290-313.

Baayen, R. H., McQueen, J. M., Dijkstra, A. F. J., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R. H.

& Schreuder, R. (eds.), Morphological structure in language processing, 151, 355-370.

Blevins, J, Milin, P & Ramscar, M (2015), Zipfian discrimination. In proceedings, Networds Final Conference: Word Knowledge and Word Usage. Pisa. Italy

Box, G. E. P. & Draper, N.R. (1987). Empirical Model-Building and Response Surfaces, New York: Wiley

Davies, M. (2009). The 385+ Million Word Corpus of Contemporary American English (1990-present). International Journal of Corpus Linguistics 14: 159-90

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. Journal of Experimental Psychology: Human Perception and Performance, 28(1), 218.

Derks, P. L., & Paclisanu, M. I. (1967). Simple strategies in binary prediction by children and adults. Journal of Experimental Psychology, 73(2), 278–285.

Finn, A. S., Lee, T., Kraus, A., & Kam, C. L. H. (2014). When It Hurts (and Helps) to Try: The Role of Effort in Language Learning. PloS one, 9(7), e101806.

Finn, A. S., Kam, C. L. H., Ettlinger, M., Vytlacil, J., & D'Esposito, M. (2013). Learning language with the wrong neural scaffolding: the cost of neural commitment to sounds. Frontiers in systems neuroscience, 7.Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. Journal of Memory and Language, 66(4), 789-806

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. Language, 84(3), 474-496

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron, 66(4), 585-595.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. Language Learning and Development, 1, 151–195.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. Cognitive Psychology, 59(1), 30–66.

James, W. (1890). The principles of psychology. New York: H. Holt and Company.

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

Kamin L.J. (1969). Predictability, surprise, attention, and conditioning. In: Campbell & Church (eds). Punishment and Aversive Behaviour. Crofts

Kemps, J. J. K., Rachèl, Ernestus, M., Schreuder, R. & Baayen, R. H. (2005). Prosodic cues for mor- phological complexity: The case of Dutch plural nouns. Memory & Cognition 33(3), 430-446

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. Nature, 431(7010), 760-767.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. Psychological bulletin, 117(3), 363.

Niv, Y. (2009). Reinforcement learning in the brain. Journal of Mathematical Psychology 53.139-154.

Pavlov, I. (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex, trs. Gelb Anrep. London: Oxford University Press.

Pinker, S. (2004). Clarifying the logical problem of language acquisition. Journal of Child Language 31.949-953.

Port, R. F., & Leary, A. P. (2005). Against formal phonology. Language, 927-964.

Plag, I., Homan, J. & Kunter, G. (2014). Homophony and morphology: The acoustics of word-final S in English. Ms, Heinrich-Heine-Universitat, Dusseldorf

Quine, W. V. O. (1960) Word and object. Cambridge, MA: MIT Press

Reddy, M. J. (1979). The conduit metaphor. Metaphor and thought, 2, 164-201.

Ramscar, M. & Port, R (2015) Categorization (without categories). In E. Dawbroska & D. Divjak (Eds.), Handbook of Cognitive Linguistics. De Gruyter Mouton.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P. & Baayen, R.H. (2014) The myth of cognitive decline: Non-linear dynamics of lifelong learning. Topics in Cognitive Science, 6, 5-42.

Ramscar, M., Dye, M. & Klein, J. (2013a) Children value informativity over logic in word learning. Psychological Science, 24(6), 1017-1023

Ramscar, M., Dye, M., Gustafson, J.W., & Klein, J (2013b). Dual routes to cognitive flexibility: Learning and response conflict resolution in the dimensional change card sort task. Child Development, 84(4), 1308-23.

Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013c). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. The Mental Lexicon, 8(3), 450-481.

Ramscar, M., Dye, M. & McCauley, S. (2013d) Error and expectation in language learning: The curious absence of 'mouses' in adult speech. Language, 89(4), 670-793.

Ramscar M.J.A., and Baayen, R.H. (2013). Production, comprehension, and synthesis: a communicative perspective on language. Frontiers in Psychology 4:233. doi: 10.3389/fpsyg.2013.00233

Ramscar, M. (2013) Suffixing, prefixing, and the functional order of regularities in meaningful strings. Psihologija, 46(4), 377-396.

Ramscar, M. & Dye, M. (2011) Learning language from the input: Why innate constraints can't explain noun compounding. Cognitive Psychology, 62(1), 1-40.

Ramscar M, Dye M, Popick HM, O'Donnell-McCarthy F (2011) The enigma of number: Why children find the meanings of even small number words hard to learn and how we can help them do better. PLoS ONE 6(7): e22501. doi:10.1371/journal.pone.0022501

Ramscar, M.J.A., D.G. Yarlett, M.W. Dye, K. Denny & K. Thorpe. (2010). The Effects of Feature-Label-Order and their implications for symbolic learning. Cognitive Science. 34(6), 909-957

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. Trends in cognitive sciences, 11(7), 274-279.

Ramscar, M., & Yarlett, D. (2007). Linguistic Self-Correction in the Absence of Feedback: A New Approach to the Logical Problem of Language Acquisition. Cognitive Science, 31(6), 927-960

Rescorla R.A. (1968). Probability of Shock in the Presence and Absence of CS in Fear Conditioning. Jouranl of Comparative & Physiological Psychology, 66, 1-5

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. American Psychologist, 43(3), 151-160

Rescorla, R.A. & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black & Prokasy (Eds.), Classical Conditioning II: Current Research and Theory. New York: Crofts

Rose, T., Stevenson, M., & Whitehead, M. (2002, May). The Reuters Corpus Volume 1. In LREC (Vol. 2, pp. 827-832).

Rumelhart, DE & McClelland, JL (1986) On learning the past tenses of English verbs In: Parallel distributed processing: explorations in the microstructure of cognition, vol. 2 Pages 216-271

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. Science, 274(5294), 1926-1928.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. Behavioral and Brain Functions 6.1-9.

Seidenberg MS & Plaut DC (2014) Quasiregularity and its discontents: the legacy of the past tense debate. Cognitive Sciience,38(6):1190-228.

Seidenberg, MS & McClelland, JL (1989). A distributed, developmental model of word recognition and naming. Psychological review, 96(4), 523.

Shannon, CA (1948) A mathematical theory of communication. Bell Systems Technical Journal 27(3): 379−423

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. Psychonomic Bulletin & Review, 3(3), 314-321.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American sign language from inconsistent input. Cognitive Psychology, 49, 370–407.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control when a little frontal lobe goes a long way. Current Directions in Psychological Science, 18(5), 259-263.

Todes , D.P. (2014) Ivan Pavlov: A Russian Life in Science. Oxford: Oxford University Press.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press

Tomasello, M (2008) Origins of Human Communication. Cambridge: MIT Press.

Triesch, J., Teuscher, C., Deák, G.O., & Carlson, E. (2006). Gaze following: why (not) learn it? Developmental Science, 9 (2), 125–147

Trimmer, P. C., McNamara, J. M., Houston, A. I., & Marshall, J. A. (2012). Does natural selection favour the Rescorla–Wagner rule?. Journal of theoretical biology, 302, 39-52.

Wittgenstein, L. (1953). Philosophical investigations. London: Blackwell

.