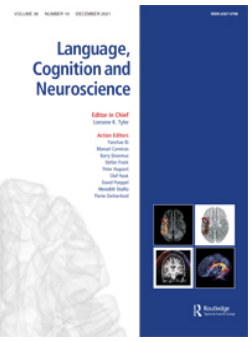

Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

A discriminative account of the learning, representation and processing of inflection systems

Michael Ramscar

To cite this article: Michael Ramscar (2021): A discriminative account of the learning, representation and processing of inflection systems, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2021.2014062

To link to this article: https://doi.org/10.1080/23273798.2021.2014062

Published online: 13 Dec 2021.

LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2021.2014062

A discriminative account of the learning, representation and processing of inflection systems

Michael Ramscar

Department of Quantitative Linguistics, Eberhard Karls Universität Tübingen, Tübingen, Germany

ABSTRACT

What kind of knowledge accounts for linguistic productivity? How is it acquired? For years, debate on these questions has focused on a seemingly obscure domain: inflectional morphology. On one side, theorists inspired by Rumelhart & McClelland's classic error-driven learning model have sought to show how all morphological forms are the products of a single memory-based process, whereas the opposing theories have claimed that irregular forms are processed by qualitatively different mechanisms to rule-governed regulars. This review argues that while the main ideas put forward by Rumelhart & McClelland – that inflectional patterns are learned, and rule-like behaviour emerges from the distribution of forms – appear to be correct, the theory embodied in their model (and those following it) is incompatible with the discriminative nature of learning itself. An examination of the constraints error-driven learning mechanisms impose on theories of morphological processing – along with language learning and human communication itself – is presented.

ARTICLE HISTORY
Received 20 May 2021
Accepted 18 November 2021

KEYWORDS
Error-driven learning; language learning; inflectional morphology; linguistic productivity; computational modelling

1. Inflectional morphology and the nature of language processing

The human capacity for linguistic communication is striking for its productivity and expressivity. People are able to use language to express a seemingly unbounded number of thoughts and feelings, and in the course of expressing themselves they routinely use and understand combinations of words – and even words themselves – that they have never encountered previously. Explaining how this is accomplished is one of the central endeavours of the brain and cognitive sciences, and in recent decades a great deal of the field's explanatory effort – along with a great deal of debate – has been focused on morphological productivity, and in particular the patterns of variation associated with the inflection of nouns and verbs. Pinker (2001) traces the origin of these specific concerns to a single, highly influential article, "The Psycholinguists" by George Miller (1967):

> "For several days I carried in my pocket a small white card on which was typed understander. On suitable occasions I would hand it to someone. 'How do you pronounce this?' I asked.
>
> He pronounced it.
>
> 'Is it an English word?'
>
> He hesitated. 'I haven't seen it used very much. I'm not sure … '
>
> (Miller, 1967, pp. 80–81; in Pinker, 2001)

As Miller notes, although "understander" is an English word, it is also a rare one. In July 2021, a search of the 1.1 billion word Corpus of Contemporary American English found just 13 attested examples, a frequency of around .01 / million words, supporting Miller's contention that it is rare enough for none of his respondents to have encountered it before. Nevertheless, Miller wrote, everyone he asked appeared to understand "understander" in the same way. He argued that this ought to bother psycholinguists, especially those who subscribed to the "behavioristic" theories that had dominated the field up to that time. How could the idea that language is simply a set of vocal responses conditioned on the presence of appropriate stimuli explain these observations? If none of Miller's respondents had seen the word "understander" before, then how could their behaviour have been "conditioned" so as to generate the consistent responses they provided?

Miller's essay sought to both capture the essence of behaviourist theories of language, and to convey the reasoning that had led many psycholinguists to reject these theories by the late 1960s. It also introduced some themes that have since come to dominate attempts to explain linguistic productivity:

CONTACT Michael Ramscar michael.ramscar@uni-tuebingen.de, ramscar@gmail.com

> Although a surprising amount of psychological ingenuity has been invested in [these kinds] of argument[s], it is difficult to estimate [their] value. No one has carried the theory through for all the related combinations that must be explained simultaneously. One can speculate, however, that there would have to be many different kinds of generalization, each with a carefully defined range of applicability. For example, it would be necessary to explain why 'understander' is acceptable, whereas 'erunderstand' is not. Worked out in detail, such a theory would become a sort of Pavlovian paraphrase of a linguistic description. Of course, if one believes there is some essential difference between behavior governed by conditioned habits and behavior governed by rules, the paraphrase could never be more than a vast intellectual pun." (Miller, 1967, pp. 80–82)

A quarter-century or so after Miller dismissed the whole idea that linguistic knowledge might be explained in terms of conditioning, advances in computer technology enabled its plausibility to be re-examined in detail, through the use of learning simulations. Rumelhart and McClelland (1986) proposed a groundbreaking model that, following Miller's example, sought to account for an aspect of English morphological productivity – in this case, the past tense suffix *–ed* – in terms of learning and conditioning; and in doing so, they employed a learning rule that, at its heart, differed little from the standard models of classical conditioning of the day (Rescorla & Wagner, 1972; Stone, 1986).

Although English inflectional morphology is a relatively simple system, its properties capture many characteristics of language as a whole (Seidenberg & Plaut, 2014). For example, when it comes to marking nouns for plurality, inflectional processes appear to be largely systematic in that all but a few English plurals appear to be formed by the addition of an allomorph of the morpheme –s to a singular noun form, e.g. *car-cars, dog-dogs*. Similarly, the overwhelming majority of English verbs appear to be formed by the addition of an allomorph of the morpheme –ed to a root verb form, e.g. *walk-walked, google-googled*. However, despite the existence of these regular patterns, these systems are only "quasi-regular", in that some of their forms deviate by degrees from them, e.g. *knife-knives, mouse-mice, child-children,* or *come-came, feel-felt, teach-taught* (Seidenberg & Plaut, 2014). As with many other aspects of language, the patterns of English inflection are exception-ridden and only partially productive, yet despite this, speakers and listeners are typically able to both effortlessly master their use, and to extend them systematically when generating forms for novel items.

Study of the development of these capacities has revealed two traits that are of particular interest for current purposes. First, using simple elicitation techniques, e.g. *This is a Wug. Now there is another one. There are two of them. There are two* ____. Or *Here is a man who knows how to rick; he did the same thing yesterday; he* ____, pioneering researchers such as Berko (1958) have shown that the capacity for the productive use of regular patterns develops early. Children as young as four are able to easily and consistently generate appropriate novel forms, e.g. *wugs* and *ricked*. Second, the development of inflectional processing in children appears to follow a "U-shaped" trajectory. Children who produce *mice* or *broke* in one context might later produce over regularised forms like *mouses* or *breaked* in others, before they go on to exhibit the consistent inflection patterns that characterise adult usage (Brown, 1973; Brown & Hanlon, 1970).

Their quasi-regular nature, their characteristic development patterns, and their relative simplicity mean that the English plural and past tense systems provide tractable (and fruitful) domains for the study of language processing and its development. Indeed, it has been suggested that inflectional morphology is the psycholinguistic equivalent of the fruitfly – a model linguistic system whose properties allow for the detailed implications of broader theories to be tested in vitro (Pinker, 2001).

Rumelhart and McClelland's model was designed to show how, despite objections such as those raised by Miller, learning was indeed sufficient to account for the workings of this system. The model performed a version of the *rick-ricked* task described above: taking a representation of the phonetic form of a verb's present tense as its input, it generated a representation of the phonetic form of its past tense as its output. Critically, it was able to generate both regular and irregular inflectional variants of verbs by means of a single, uniform procedure. Moreover, this procedure also supported the inflection of novel forms, including the over-regularised forms that children produce as they learn to generalise, and the "U-shaped" learning pattern associated with them.

From a theoretical perspective, the Rumelhart and McClelland model's ability to generalise can be seen as its most critical feature (Seidenberg & Plaut, 2014). At the time the model was first presented, people's ability to generate novel forms that were unattested in their experience was often considered to provide *prima facie* evidence for the idea that abstract rules exist as a form of linguistic knowledge, since it was widely assumed that these rules were *necessary* to account for regular patterns of generalisation (see Miller, 1967, above). By contrast, the performance of the Rumelhart and McClelland model offered an alternative explanation for the consistent patterns of generalisation observed in inflection: analogy. The reason that the model generalised the regular + *ed* pattern to mark the past tense of novel verb forms was simply because most of the past tense forms it had been trained on ended in *ed*. Accordingly, this raised the possibility that the patterns of generalisation exhibited by language users might not in fact be the product of "rules", and might instead reflect latent patterns in the distribution of the linguistic forms that speakers had been exposed to and learned from.

Another important contribution of the model was its detail (Seidenberg & Plaut, 2014). Although high-level psycholinguistic theories can posit mechanisms at will, they cannot – in principle – answer questions about whether these mechanisms are actually implemented in the way an intuitive theory supposes. Nor can they show whether the actual implementation of a posited mechanism would in fact behave in a way that is consistent with what is known empirically about our linguistic / psychological capacities. By contrast, implemented models allow for notional mechanisms to be straightforwardly tested against empirical data.

A further positive attribute of computational models is that because they implement mechanisms, they allow for discovery. Running a model can enable the automatic generation of explanations (and predictions) that may be far from intuitively apparent. Prior to the appearance of the Rumelhart and McClelland model, the suggestion that both the regular past tense and its diverse irregular counterparts could be the product of a single mechanism seemed not only highly counterintuitive to many researchers, it was often thought to be impossible (Seidenberg & Plaut, 2014; see also Miller, 1967). By contrast, although the Rumelhart and McClelland model did not explicitly implement rules, it was nevertheless able – to a degree – to generate patterns that corresponded to the "rules" of regular English inflection. Yet, as mentioned above, any rule-like behaviour in the model's output simply emerged as a result of there being regularities present in the distribution of phonetic features in its input: the regular forms and the irregular forms that are the exceptions to these rules were all computed by the same mechanism. (Later developments of it, e.g. Hahn & Nakisa, 2000; Haskell et al., 2003; Joanisse & Seidenberg, 1999; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1993 provided even more detailed accounts of this process, and extended it to other apparent "rules".)

Although the model contains a number of architectural embellishments (such as translators that transformed phonetic features into its own internal representation scheme), its core component is what Rumelhart and McClelland described as a "pattern associator" network. In the simulation process, this network was trained to predict a representation of the phonetic forms of the past tense of each English verb from a representation of the phonetic form of its present tense (or *stem*) form. These were arranged so that all of the phonetic features of the stems were connected to all of the phonological features of the past tenses. In training, the weights on these connections were then adjusted using an error-driven learning rule (a version of the delta-rule, Stone, 1986), which served to adjust the weights on connections in the network in order to find the pattern that best fit the training set.

The model's training proceeded on a trial-by-trial basis. On each trial a set of inputs that corresponded to a past tense form was presented to the model, and the weights on the connections in the network were adjusted by calculating the discrepancy between the intended activation of the outputs (which corresponded to its past tense form), and the actual activation pattern. As training progressed, these discrepancies caused the value of the inputs that led to erroneous output activations to be downgraded, and the value of the input features that led to correct activations to be reinforced. The goal of the learning process was for the network to settle into a pattern in which the weights from a given set of inputs activated only the correct output features for that set, and in practice this meant learning to inhibit the influence of any input features that led to the activation of incorrect output features (this is discussed in more detail in Section 3.1).

Because of the way that forms were represented in the model, any individual input and output unit could take part in the representation of multiple stem and past tense forms. The distributed nature of these representations of the inputs and outputs and the distribution of the sound patterns of English verbs guaranteed that erroneous patterns of activation would occur whenever new patterns were encountered in training. This in turn enabled the model to not only simulate the inflection of novel forms, but also the patterns of over-regularisation observed in children (who, as noted above, often produce forms such as \ *breaked* instead of *broke*). In both cases, the generalisation of the regular pattern occurred as a result of the frequency of the regular + *ed* ending in the distribution of English past tense verbs (and the regular + s ending in the distribution of English plural nouns). Given the combination of the setup of the network, the learning rule and the distribution of input forms, it follows that until the model had explicitly learned to inhibit the activation of the features representing + *ed* when any given pattern was input, the pattern of weights learned by the network would always tend to favour the activation of those features. Accordingly, Rumelhart and McClelland summarised the model's contribution as follows:

> "We have, we believe, provided a distinct alternative to the view that children learn the rules of English past tense acquisition in any explicit sense … a child need not figure out what the rules are, or even that there are rules … A uniform procedure is applied for producing the past tense form in every case. The base form is supplied as input to the past tense network, and the resulting pattern of activation is interpreted as the phonological representation of the past tense form of that verb … " Rumelhart and McClelland (1986, p. 267)

Given that the prevailing theoretical consensus at the time held that human language could not even be learned (Chomsky, 1985), it is perhaps unsurprising that the Rumelhart and McClelland model evoked a lively critical response. Whereas some of these criticisms questioned the broader implications of the model (e.g. if language really is that product of these basic learning mechanisms, then why do children learn languages whereas their pets, which share the same mechanisms and environments, uniformly fail to do so?), other critiques focused on the details of the model itself. Did its performance really tell us much about the way the past tense was actually processed in or learned by the minds of speakers?

Pinker and colleagues (Marcus et al., 1992; Pinker & Prince, 1988) argued that in practice, Rumelhart and McClelland's explanation of U-shaped developmental sequence was inadequate. Critically, it appeared that the U-shaped learning the original Rumelhart and McClelland model exhibited was simply a reflection of the way that the input of regular and irregular items had been manipulated in the model's training regime. The model was initially trained on mostly irregular forms, and then at a discrete point in time, training switched to mostly regular forms, and it is clear that both this pattern and rates at which items were introduced in the training set was far from reflective of children's actual experience (Pinker & Prince, 1988).

Perhaps more importantly, critics of the model also described a range of phenomena associated with English inflection that seemed to show that a set of systematic constraints applied to regular – but not irregular – inflection. It was argued that not only was the existence of these constraints entirely incompatible with the account put forward by Rumelhart and McClelland, but that they also provided concrete evidence for the existence of rules themselves (see e.g. Alegre & Gordon, 1996; Kim et al., 1991, 1994; Gordon, 1985; Pinker, 1999, 2001; Pinker & Prince, 1988; Prasada & Pinker, 1993; see also Legate & Yang, 2007; Lignos & Yang, 2018; Marslen-Wilson & Tyler, 2007). One of the more important of these phenomena is what Pinker (2001) describes as *systematic regularisation*: the tendency for some irregular verbs and plurals to regularise in certain contexts. For example, in English, although the plural of *life* is *lives*, the compound noun *low-life* has the plural form *low-lifes* ("all of my daughter's friends are low-lifes") not *low-lives*. Similarly, when it comes to verbs, a city encircled by ordinance is *ringed* with artillery, not *rang* with artillery. Pinker argued that these examples reveal a flaw in Rumelhart and McClelland's model because they clearly indicate that representations of phonetic features cannot be the only inputs to the inflection system. If a given input, say *life*, can come out the other end of the inflection process as either *lifes* or *lives*, then some other factor must be serving to influence this.

Pinker and colleagues (Pinker, 1999, 2001; Pinker & Prince, 1988) argued that words that underwent automatic regularisation all shared a formal property, namely that they were *headless* (the links between an irregular root stored in the lexicon and its past tense or plural form had been broken, Kiparsky, 1982; Selkirk, 1982; Williams, 1981). They further argued that the existence of systematic regularisation (which revealed that the inflection process was sensitive to whether words and compounds were headless, and as a corollary, whether words were nouns or verbs) provided support for a more traditional account of inflectional morphology. Linguists has long thought that inflection process comprise two processes, with the formation of past tenses for irregular forms relying on rote memorisation, and regular forms being generated by rules. However, because the results of *WUG*-like tasks were incompatible with a simple story about rote memory storage (the past tense of *SPLUNG* is typically generated for the novel verb stem *SPLING*, Pinker, 1991), Pinker and colleagues proposed a modified "dual-route" theory in which the idea of a simple memorised list of irregulars was upgraded to include an associative memory component akin to the one proposed by Rumelhart and McClelland (Kim et al., 1991; Pinker, 1999, 2001; Pinker & Prince, 1988; Prasada & Pinker, 1993).

This modification enabled the account to both explain the analogical processing of irregular verbs and nouns, while also allowing it to maintain that regular forms were processed by an abstract rule, which, according to the theory, could be shown to be distinct from analogical processes in memory because it acted as a default in instances when memory failed, or when irregular forms were blocked for grammatical reasons (Pinker, 1999; Pinker & Prince, 1988). The dual-route theory also made testable empirical claims. Notably, it argued that phonological and grammatical information were the only factors that were relevant in the processing of inflections:

> "inflection is an isolable subsystem in which grammatical mechanisms can be studied in detail, without complex interactions with the rest of language. It is computed independently of syntax, the subsystem that defines the form of phrases and sentences … [and] is also insensitive to lexical semantics … " Pinker (1997, p. 531)

2. Context and the routes to inflection: where the past tense debate went wrong

If the dual-route account was correct, then it clearly posed an insurmountable problem, not only for the Rumelhart and McClelland model, but also for the approach it embodied. Accordingly, it is notable that alternative suggestions for the causes of what Pinker and colleagues described as "systematic regularization" did exist. MacWhinney and Leinbach (1991; see also Harris, 1992) had argued that "semantic stretching" (where the patterns of usage of a word root and its past tense form in context is sufficiently dissimilar that it weakens the semantic similarities between them) might result in the past tenses of verbs that are usually irregular becoming regularised.

To empirically test this proposal, Ramscar (2002) conducted a series of elicited inflection tasks, the results of which revealed that the context in which a novel verb was encountered could significantly influence the forms participants then produced as its past tense. Critically, these experiments showed that generation of both irregular *and* regular forms appeared to be influenced by context. Moreover, by manipulating the context in which novel verb forms like SPLING were presented, Ramscar showed that participants could be systematically influenced to produce either the irregular form SPLUNG or the regular form SPLINGED. The finding that context / semantics could affect inflection clearly seemed to show that inflection patterns were *not* solely influenced by grammatical and phonological information. It also raised a question: was it really the case that systematic regularisation was driven by speakers' sensitivity to headless forms, or did speakers' tendency to regularise forms like *low-lifes* simply reflect the effect of the different contexts in which these forms occurred on speakers' semantic representations of them (MacDonald & Ramscar, 2001)?

To answer this question Ramscar (2002) conducted a series of experiments that pitted the predictions of a semantic account of homophone inflection against those of the dual-route account advanced by Pinker et al. (Kim et al., 1991; Pinker, 1997, 1999, 2001). The dual-route account predicted that the regularisation of irregular sounding verb stems was driven by grammatical sensitivity. It maintained that the reason why the past tense of to "ring the bell with a hammer" is rang, whereas the past tense of "ring the city with artillery" is ringed is because people analyze the latter as being derived from a noun (Kim et al., 1991). Accordingly, the theory predicted that speakers would automatically regularise any verb that they perceived as being derived from a noun (i.e, analyzed as being headless).

However, Ramscar (2002) found that a first set of participants' perception of the "grammatical origins" of verbs was a poor predictor of a second set of participants' preferences for irregular versus regular past tense forms of homophone verbs in context. By contrast, a set of ratings of the semantic similarities between the forms of verbs in context taken from third set of participants did serve to accurately predict the form preference ratings. These findings were then subjected to further test in a series of reading time experiments (Ramscar et al., 2013) which revealed no dissociation in processing – reading times for regular forms were influenced by context in the same way as irregular forms – a finding that runs counter to the dual-route theory's claim that regular inflections were processed by a separate, context insensitive system.

Still further support for the idea that *all* inflected forms are subject to the influence of context was provided by Ramscar and Dye (2011), who examined the role of context on plural inflection, and in particular the effects of context on the behaviour of regular and irregular plural forms in compounds, another domain that seemed to offer support for the dual-route theory. When people are asked to judge the acceptability of regular and irregular plurals in compounds, they seem to treat them differently, e.g.:

rat-eater

mouse-eater

*rats-eater

mice-eater

When native speakers judge the acceptability of these compounds, they tend to feel that only *rats-eater is ill formed. Pinker (1994, 1999) argues that this preference provides further evidence form the formal operation of rules on inflection, since it derives naturally from the dual-route theory's account of about rule based plural inflection and compounding. This account rests on four assumptions, two of which are basic to the dual-route theory, and two of which come from rule based accounts of compounding (Seidenberg & Plaut, 2014):

(1) Only singular nouns and irregular plural forms are stored in memory.
(2) All regular plurals are generated by rule.
(3) Compound formation is also governed by rules (Kiparsky, 1982; Siegel, 1979).
(4) There is a strict order in the way that rules are implemented in the mental parsing and production system, with compounding rules being applied *before* the plural rule.

These assumptions make it relatively easy to generate an explanation for the intuitive appeal of the X-eater examples above: If compounding occurs before inflection, and regular plurals are *not* stored in memory, then this system can produce *rat-eater*, *mouse-eater* and *mice-eater,* but in cannot produce **rats-eater*, and which accounts for why the latter sounds weird. What is more, a series of studies designed to further explore this system (Gordon, 1985) revealed that while young children regularly produce forms like *mice-eater* in elicitation tasks, they do *not* produce analogous forms such as *rats-eater*. Instead, in exactly the same contexts that children produce *mice-eater* to describe a monster that eats mice, they overwhelmingly produce *rat-eater* to describe a monster that eats rats.

In an extension of these studies, Alegre and Gordon (1996) then examined the interpretation of noun phrases such as *red rat eater*. This noun phrase is ambiguous in that it appears to support two interpretations: [*red rat*] *eater* (an NP/N structure, in which something is an eater of red rats) and *red* [*rat eater*] (an ADJ/NP structure, in which that rat eater is red). However, because the structure of the system described above rules out the formation of the compound noun *rats eater*, it follows that if people are sensitive to this, then the phrase *red rats eater* ought not to be ambiguous. It should only be interpreted as the NP/N, [*red rats*] *eater*. Consistent with this, when Alegre and Gordon examined young children's interpretations of the noun phrases *red rat eater* and *red rats eater* they found that their preference for the NP/N interpretation – [*red X*] *eater* – increased markedly when *X* was the plural form, *rats*.

To examine whether these findings really did support the level-ordering account, Ramscar and Dye (2011) conducted a series of experiments that not only examined regular noun phrases but also, critically, noun phrases containing <u>irregular nouns</u>, such as *red mice eater,* which Alegre and Gordon had not tested. If the level ordering account of regular compound interpretations is correct, it follows that interpretations of noun phrases containing irregular plurals ought *not* to show the same patterns of bias for noun phrases containing regular plurals. This is because the theory maintains that the singular and plural forms of irregulars like *mouse/mice* are both stored in lexical memory, and it thus follows that the explanation for why *red rats eater* should be interpreted as an NP/N rather an ADJ/NP – because *rats eater* is not a legitimate product of the system – does <u>not</u> apply to *red mice eater*.

Ramscar and Dye's results showed that for both adults and very young children, the same patter of preferences for the NP/N structure held regardless of whether regular or irregular plurals occurred in these compounds, which suggests that Alegre and Gordon's original findings were not in fact driven by the *regularity* of the plurals in compounds. Moreover, and perhaps most importantly, Ramscar and Dye also showed that people's biases about whether compounds containing plurals should be given an NP/N or an ADJ/NP interpretation could be easily manipulated by changing their components – e.g. *brave soldiers list* versus *long soldiers list* – or by changing the contexts in which they were presented.

In other words, whenever they were tested empirically, the claims and evidence that had been put forward to support the claim that regular and irregular forms were processed differently from one another – which would count against a straightforward account of morphological processing based on learning – did not stand up to scrutiny (see Seidenberg & Plaut, 2014, for a similar conclusion).

However, the results reviewed above suggest more than that. They do not only show that the processing of both irregular and regular forms *can* be influenced by context; they seem to suggest that the process of inflection is influenced by context *everywhere*.

3. Form relies on context: motivations for a discriminative account of the learning, representation, and processing of inflection

The findings reviewed in the previous section are clearly inconsistent with the dual-route account put forward by Pinker & colleagues (Pinker, 2001, 1999). Critically, however, they are hardly consistent with the single-route account proposed by Rumelhart and McClelland either. Rather, they appear to raise some critical questions about the analyses that gave rise to the Rumelhart and McClelland model in the first place. First, while Rumelhart and McClelland describe the learning component of their model as a "pattern associator", it is far from clear whether the functioning of a two-layer network implementing a version of the delta-rule is accurately described as "an associative learning model" at all (this point is discussed in more detail below). Second, although the model treats inflection as a process in which a mechanism takes a stem form and transforms it into an inflected form, one might reasonably ask whether this is really the best way of conceptualising this process from a learning perspective, especially from the perspective of a model that incorporates the role context and semantics play in inflection.

From this latter perspective, it is important to note that not only do many of the results described so far indicate that context is a critical determinant of form, but it is of course also the case that in learning, children are rarely if ever exposed to the transformations that the model is set up to learn. Rather, it seems clear that children mainly, perhaps even only, encounter individual forms, whether they are "inflections" or "stems", in context. That is, the idea that root forms are transformed into inflected forms – a cornerstone of the Rumelhart and McClelland model – is a theoretical analysis taken straight from generative models in linguistics. Yet, from a learning perspective, the idea that children learn to transform root forms into inflected forms actually makes very little sense at all, a point that becomes especially apparent when one considers how contextual information might be incorporated into a model of inflection learning. If we assume that children *learn* transformational rules, then it follows that we must assume that the contexts in which children learn supply evidence for these rules. That is, we must assume that children learn the grammar of English by listening to their caregivers reciting the various root forms followed by their various inflections. However, if there is one thing that proponents of all of the theories described so far agree on, it is that this is not what children are exposed to in learning, and that this is not how children learn. Accordingly, given the disparity between what the Rumelhart and McClelland model (and the many models of inflection based on the same conceptual analysis that have followed it) learns on one hand, and the actual learning task children are faced with, one might reasonably ask whether these models really give us much insight into what children actually learn as they master the processes of inflection (and language) in context.

So what do children learn? To begin to address this very big question, it seems helpful to break it down into two smaller ones. First, how are the capacities of learners best characterised? What kind of learning mechanisms can we reasonably attribute to language learners? Second, what kind of processes can we plausibly envisage these learning mechanisms supporting? That is, what might a model of the inflection process derived from the way morphological patterns are learned in context actually look like?

Since answering the second of these questions clearly relies on the answers to the first, it will be addressed later, in the section following this one. With regards to the first question – how are the capacities of learners best characterised – it is clear that humans share their basic learning mechanisms with other animals, and as a result, animal models have enabled us to gain considerable insight into the neural and psychological processes that govern learning. The evidence from these models suggests that biological learning mechanisms are best characterised in error-driven terms (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Schultz, 2006). For example, the classic Rescorla and Wagner (1972) model of animal learning uses a computation of the discrepancy between a learner's expectations and reality to modify the weights in a network connecting a set of predictive cues to a set of expected outcomes in trial-by-trial learning in much the same way as input forms are connected to output forms in the Rumelhart and McClelland past tense model. Indeed, the Rescorla-Wagner learning rule is simply the linear form of an earlier rule proposed by Widrow and Hoff (1960; see Stone, 1986), and this in turn is formally equivalent to the delta-rule used by Rumelhart and McClelland (Sutton & Barto, 1981).

3.1. Learning: associationism and the discriminative turn

In order theoretically characterise the way that delta-rule learning might capture what it is that children learn as they come to master the use of words in context, Ramscar et al. (2010) conducted a conceptual and empirical analysis of the functional role that error-driven learning might play in this task, taking the Rescorla and Wagner (1972) learning model as an example. This model was conceived and developed as an associative learning model, its goal being that of offering a formal description of the way that classical conditioning experiments were thought to show how animals learned to associate a set of perceptual / conceptual primitives with events in the world. However, Ramscar et al. (2010) argued that although the developers of error-driven learning models – e.g. Rumelhart and McClelland, and Rescorla and Wagner – had traditionally thought about modelling these tasks in *associative* terms, a detailed analysis of the function of delta-rule learning suggested that for most purposes, the rule was actually best conceptualised as describing a *discriminative* learning mechanism (this point also applies to the error-driven learning algorithms found at the heart of other connectionist / neural network models; Ng & Jordan, 2002, as well as Bayesian models of conditioning, e.g. Daw et al., 2008).

Before describing the logic of discrimination enshrined in the mechanisms implemented in these models, it is worth noting that the term *discrimination learning* is used many ways in the literature, and that this can potentially lead to confusion when it comes to understanding these mechanisms (Hoppe et al., 2021):

(1) The term *discrimination learning* has been widely used in the animal learning literature since the early part of the twentieth century. Consistent with the behaviourist principles that dominated theory at this time, it was (and still is) used in this sense in a mechanism-neutral way to describe the fact that objectively, both animals and humans were able to learn different responses to different stimuli. Accordingly, *discrimination learning* simply meant learning to associate response A with stimulus B, and response C with stimulus D (Rescorla & Wagner, 1972).
(2) In machine learning, the term *discriminative model* was introduced to provide a concrete, mathematical and conceptualisation of one possible way of learning in classification problems. Generative models – which *discriminative models* are typically contrasted with – learn the data that generates a set of labels. By contrast, discriminative models are defined in terms of their capacity to learn to maximise the conditional probabilities of output units given input units (Ng & Jordan, 2002). This definition is once again neutral with regards to the mechanism. While most classification problems in which discriminative models are employed also tend to implement the discriminative algorithms discussed below, they need not necessarily employ these algorithms.
(3) Finally, *discriminative learning* can be applied to the algorithm implemented in error-driven learning models (Ramscar et al., 2010). This is because in most learning situations, these algorithm enforce *cue competition*, a process that serves to discriminate against or in favour of the units that serve as inputs – by re-weighting the influence of individual units – according to how informative they are about different outputs (this process in described in detail below).

As with the misconception that holds that delta-rule learning is "associative", these different notions of discriminative learning have important implications for the way the learning process is conceptualised.

For example, it has long been known that simple *association rates*, the frequencies with which stimuli (or cues) are associated with responses (or outcomes) are incapable of explaining basic conditioning. Rather, learning in animals has been shown the product between an interaction between the rates at which cues and outcomes co-occur, and two other related factors: the *background rates* of cues (how often a given cue occurs in the absence of a given event) and *blocking* (the prior predictability of an outcome in a context in which it co-occurs with a cue). Because all three of these factors are captured by the delta rule, what is actually learned by models that implement it (and its variants) depends on an interaction between these factors as training unfolds. Overall, the association rates between cues and outcomes will tend to increase the weights on the links between them in learning, while the effects of blocking and background rates tend to inhibit or even decrease the value of these weights (see Ramscar et al., 2010, 2013b for reviews).

As was described in relation to the Rumelhart and McClelland model earlier, as delta-rule learning progresses on a trial-by-trial basis, the values of the weights on a link between a cue and an outcome are increased when the cue occurs with an outcome that is not already fully predicted, and devalued when it leads to prediction error (when an outcome that does not occur is predicted). The actual value of these changes is a function of what has already been learned, and it decreases in proportion to the degree to which outcomes are predicted by prior learning so that when outcomes are fully predicted, no learning occurs. This last feature enables the delta-rule to capture the idea that the goal of learning is the reduction of uncertainty, since it follows that if learning about something decreases an individual's uncertainty about it, they will have less to learn about that particular thing.

One result of the way the simple factors described above interact in the learning process is that they typically force cues to compete for predictive value as part of a fully connected system of cues and outcomes. *Cue competition* typically results in the formation of strong positive weights between cues that produce little or no error for a given outcome in training, and strong *negative* weights between inputs that do lead to prediction errors. This means that in practice, the learning mechanism described by the delta rule results in a process that emphatically does *not* simply learn to associate cues with outcomes (as Rumelhart and McClelland's description of a two-layer network as a "pattern associator" might imply). Rather, the outcome of the learning process is a set of network link values that *discriminate* in favour of more reliable inputs and against less reliable inputs. A further important result of this process is that it inevitably follows that many elements of the input patterns fed to the rule in training will in fact be entirely *disassociated* from output patterns that they co-occur with.

Accordingly, although the Rescorla and Wagner (1972) learning rule was originally proposed as part of an elemental model couched in associative terms (see also Ellis, 2006; Miller et al., 1995; Siegel & Allan, 1996), Ramscar et al. (2010) argue that because it actually implements an error-driven learning mechanism, for most purposes (an exception is described below) it is best understood by re-conceptualising learning as an *discriminative process* that reduces a learner's uncertainty about events in the world by learning to ignore them.

These considerations have significant implications for the way that learning is conceptualised. In particular, the nature of learning appears to suggest that there are strong constraints on what children can be expected to learn as they master the use linguistic of forms in context. To illustrate one of the more critical of these constraints Ramscar et al. (2010) conducted a series of analyses and simulations that show how, as a consequence of the role that cue competition plays in error-driven learning, the temporal structure of information can play a crucial role in determining whether or not discriminative learning actually occurs. The nature of this constraint can be best illustrated by first considering the effects of learning in a context in which a set of complex stimuli predict a set of discrete elements (i.e. when a large cue set is used to predict a smaller set of outcomes), and then comparing it to its inverse, a context in which a set of discrete elements predicts a set of complex stimuli (when a small set of cues is used to predict a larger set of outcomes; see Figure 1).

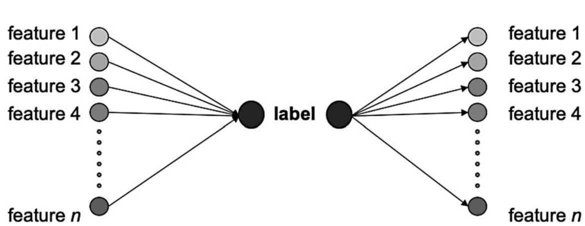

Figure 1. The possible relations (links) between a discrete label (e.g. a word or an affix) and the features of a high-dimensional context / object as events and labels occur in learning / time. Feature-to-Label learning (left) will facilitate cue competition between features, and abstraction of the informative dimensions that predict labels. Given that only a single cue occurs, competition is impossible in a Label-to-Feature relationship (right), which will simply facilitate learning of the probabilities of the features given the label.

In the first of these two learning scenarios, the *features* of things (events, objects, actions, etc.) in the world serve as cues to the forms used to talk about them *labels* (Feature-Label learning), an information structure that naturally allows for features to compete as cues to labels (Figure 1, left). To explain why, consider a child being shown one of the objects in Figure 2, and being told, "Look! A wug", If we assume that the child learns in the way error-driven learning models suppose, their mind will reinforce all of the features of the object to "wug". This means if they later encounter another identical object, then given its shape they will (implicitly) expect it to be a wug. However, it is important to note here that if they then hear "Look! It's a wug", then although this will strengthen the connection between this set of features and "wug", it will *not* help them to learn how to use "wug" appropriately. This is because given their current knowledge, and given the overlap in the features of all of the objects shown in Figure 2, they will assume that the objects in centre panel are also wugs, when in fact they are nizzes.

To learn to discriminate wugs from nizzes, our child will (implicitly) need to make some prediction errors. Suppose they next encounter a niz, the object in the centre panel of Figure 2, and hear, "Look! It's a … " Given their prior experience, they will be expecting to hear "wug". However, because the expectation that they will hear "wug" is erroneous (they actually hear "niz"), they will learn to devalue the features of wugs that they erroneously supposed to be cues to "niz" (namely their highly salient but uninformative body shapes). That is, they will learn that they are less likely to hear "wug" when this feature is present than they had supposed. This process will cause value to shift from features that produce more error to those that produce less: the less salient wug features will be implicitly strengthened as a cues to "wug" simply because the value of the wug body shape feature has been devalued. This in turn means that despite the fact that they never heard the word "wug" in this context, the child's understanding of wugs will have actually improved after they learned about "nizzes". And because the converse will happen next time they hear a wug described in a similar context, our child will soon learn to discriminate the right conditions in which to expect (and use) the labels "wug" and "niz".

However, when these relationships are reversed (see Figure 3), such that labels serve as cues (Label-Feature learning), cue competition becomes somewhat problematic. This is because only one cue will be present in speech at any given time, and – as is hopefully obvious – a single cue cannot compete with itself. If a cue is reinforced in isolation, then the *shift* in value

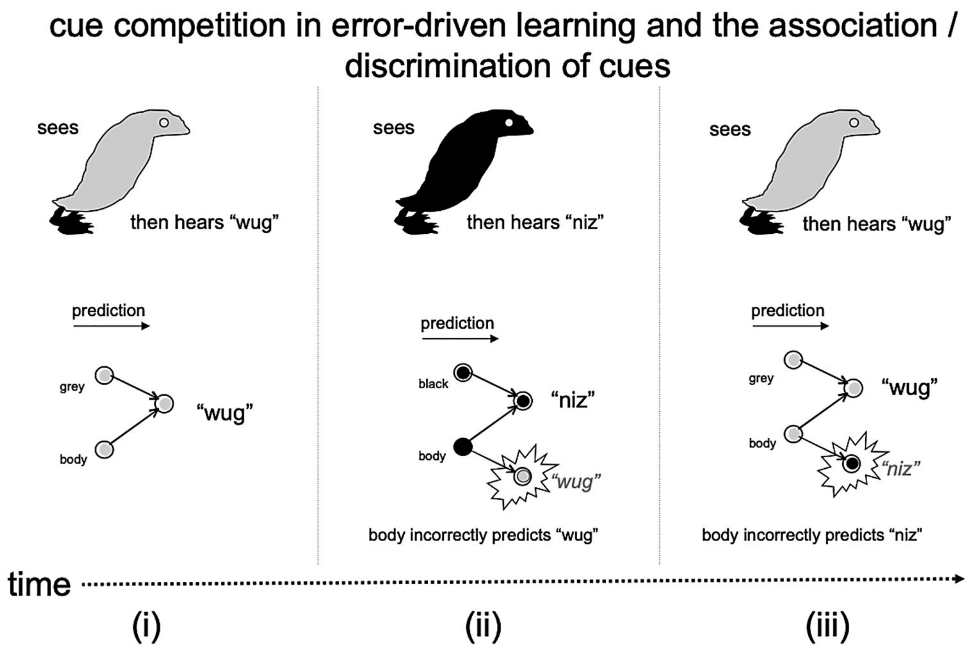

Figure 2. When different sets of features predict different labels, the non-discriminating features will be dissociated from the labels as a result of cue competition.

between competing cues described in the scenario above cannot occur, simply because there are no competing cues for it to lose value to. Because of this, the value of an isolated cue will simply rise and fall in isolation. (It is perhaps worth noting here that although labels themselves also comprise different features, in language these sub-features do not tend to correlate with objects and events in the world in meaningful or systematic ways, which means that cue competition among these features is unlikely to result in the learning of either stable or informative patterns.) Accordingly, it follows that Label-Feature learning will not be discriminative. It will simply lead to the learning of the likelihood of each feature given the label instead. (Again, it should be noted that this is very much a theoretical analysis – in the real world, when it comes to actual word learning, it

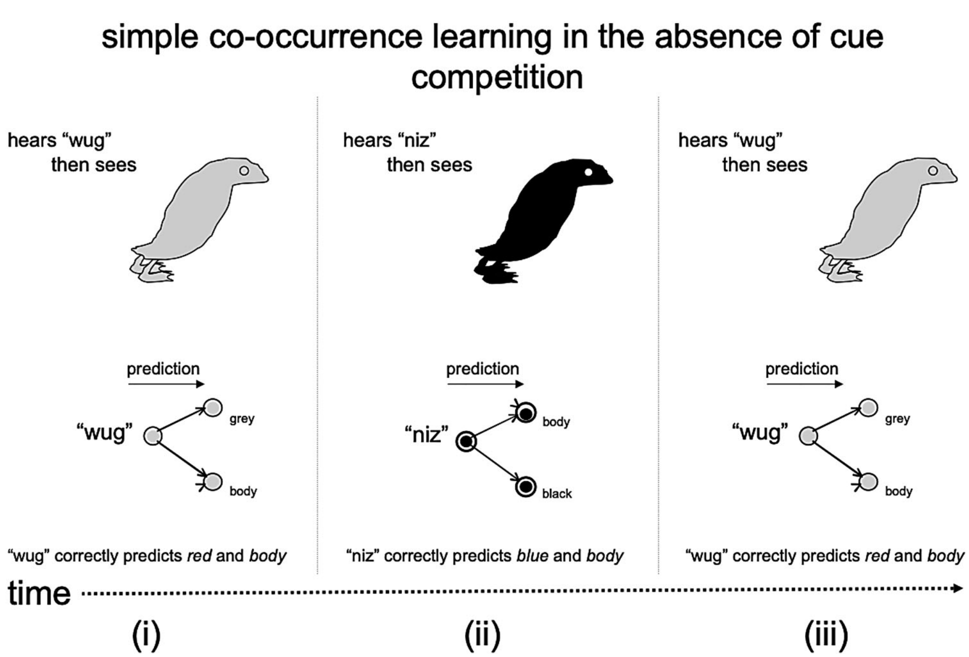

Figure 3. When single labels predict sets of features in isolation, learning will simply result in the conditional probability of each feature given each label being learned.

is likely that repetition and other discourse factors will make the temporal relationships between labels and the world more complicated than this idealisation suggests.)

Ramscar et al. (2010) conducted a series of delta-rule learning simulations to confirm the predictions of these analyses, which were then further explored in a series of empirical studies. In the first of these, adult participants were trained in a rapid presentation paradigm on categories that comprised the novel "fribbles" shown in Figure 4. Each fribble comprised a number of features that included a highly salient body shape that was not a defining feature for classification purposes. Critically, the body shapes of the fribbles were distributed systematically across the three categories, so that 75% of the members of one category and 25% of the members of another category shared the same body shape.

While this meant that body shape was unhelpful to learners trying to determine the cues to category membership, the fribbles also possessed other, less salient features that were more helpful in this regard. To successfully learn the categories, participants had to learn to ignore (i.e. unlearn) the uninformative body shapes and focus on the less salient fribble features, a process that would be facilitated by the cue competition process described above: the greater level of prediction error generated by the fribbles' body shapes as compared to their other features would enable an error driven learner to shift their weighing of the values of the available cues from the fribbles' bodies to those other features.

Learning the fribbles as cues to discrete labels, such as *wug* or *dep* (Feature-Label learning) thus allowed for competitive learning amongst the co-varying cues present in the fribbles. Ramscar et al. predicted that this would allow participants to learn to discriminate the informative features and categorise the fribbles successfully. Consistent with this, participants given Feature-Label training were subsequently able to classify low and high-frequency exemplars with a high degree of success in testing (Figure 5). However, when the temporal arrangement of labels and fribbles was reversed, so that the process was one of learning to predict the features of each fribble from a label (Label-Feature learning) it was predicted that because cue competition could not occur classification performance ought to be poor. As Figure 5 also shows, consistent with the absence of an information structure that facilitated the unlearning of the uninformative dimensions in the category structures, participants trained with labels as cues to objects failed to learn to categorise the low-frequency items.

Ramscar et al. (2010) then applied this analysis to a long-standing puzzle regarding children's learning of colour words. While even infants can distinguish basic colour categories (Bornstein et al., 1976), it has long been noted that children go through surprisingly long period in which their application of color words is inconsistent and haphazard (Sandhofer & Smith, 1999). Moreover, children's behaviour exhibits a similar, even parallel pattern when it comes to learn number words (Ramscar et al., 2011). From a discriminative learning perspective, the problems children face in learning to use colour and number words are very similar. Although they will encounter three bears or brown bears, they will never independently encounter "a three" or "a brown". Rather, because these words necessarily occur in ambiguous contexts (Figure 6) children are faced with the problem of learning to discriminate the appropriate cues to a given word in a given context. This problem is analogous to the learning of the feature cues to the category labels in the experiment described above, which suggests in turn that if children learn to use context to predict the different forms of individual number or colour words, they will be able to solve this problem by using prediction error and cue competition to discriminate the environmental features – differences in numerosity, or hue – that reliably cue different number and colour words.

Accordingly, the analyses described above suggest that post-nominal constructions ("this bear is brown") will be far more likely to facilitate the discrimination of the appropriate cues to colour words than pre-nominal constructions ("this is a yellow ball"), because if a child has already learned what a ball is (and children master nouns long before colour words), the features of a highlighted ball can be used to predict the colour word "yellow". Subsequent encounters with other balls in context constructions ("this is a green ball") will then enable cue competition to discriminate the features of balls that reliably predict colour words (hues) from the features that are not reliable predictors of colour words (size, texture, roundness etc.).

By contrast, child encountering a pre-nominal construction ("this is a yellow ball") will be in a position akin to that of the Lable-Feature-learners in the fribble experiment above. Because colour words will serve as cues to nouns in context, the structure of this learning situation will not facilitate cue competition. Confirming this, an empirical study of colour word learning revealed that while training with post-nominal constructions significantly improved the accuracy and consistency of two-year olds' responses to questions involving colour words, pre-nominal training had no effect on their performance (Ramscar et al., 2010).

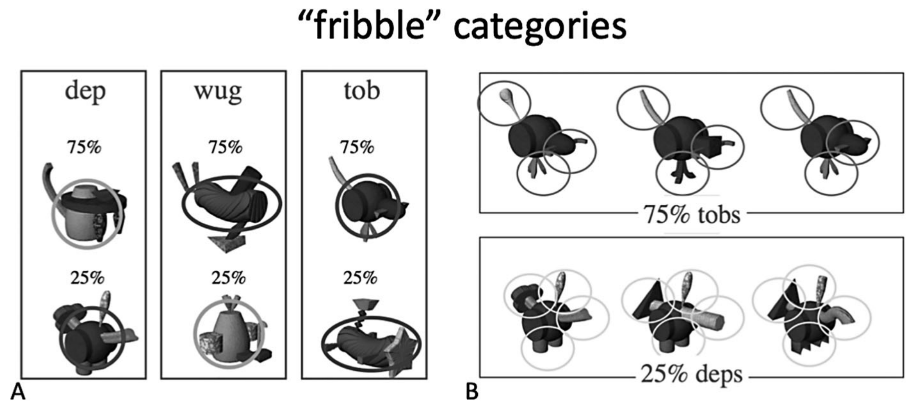

Figure 4. The category structures employed by Ramscar et al. (2010). Note that the fribble body types (circled in panel **A**) do not discriminate between the categories. Accordingly learners must learn to inhibit (unlearn) these features in order to successfully learn both the low frequency and high frequency subcategories. The features learners need to positively weigh in order to successfully discriminate between the low-frequency "dep" and high-frequency "tob" subcategories are circled in panel B. (Stimulus images courtesy of Michael J. Tarr, Carnegie Mellon University, http://www.tarrlab.org/)

As a further demonstration of the role of information structure in discrimination learning, Ramscar et al. (2011) showed how these analyses can also account for some of the difficulties children have learning to use number words (as Figure 6 shows, colour and number word learning have much in common). Training children on 2, 4 and 6 in a Feature-Label configuration (children were shown a set of objects, e.g. bears, and asked, "What can you see? Bears. There are four") not only improved children's ability to discriminate sets of 2, 4 and 6, it also improved their ability to discriminate sets of 3, 5 and 7. This latter improvement, which occurred despite the fact that only 2, 4 and 6 were ever shown in training, is of course consistent with the analysis of a child learning to use "wug" above, which emphasised the fact that although reinforcement is important in learning, the discrimination of the correct cues to labels ultimately relies on prediction error. Further consistent with the analyses above, Ramscar et al. (2011) found that training children in Label-Feature configuration ("What can you see? There are two balls") did not improve their performance.

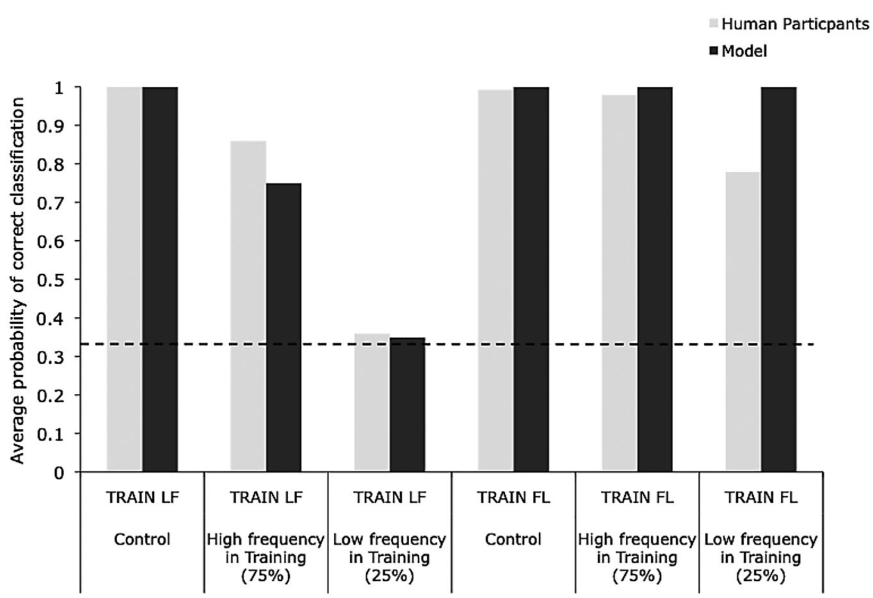

Figure 5. The predictions of delta-rules simulations plotted against the performance of participants in the fribble category learning experiment (Ramscar et al., 2010). The control category was designed to check there were no learning differences between the two groups other than those predicted and comprised exemplars that all shared one, highly salient feature (all were blue). Because learning simply involved making a binary pairing between the colour blue and the category label, performance was on this category expected to be identical regardless of whether LF and FL training was given. Reproduced with permission from Ramscar et al. (2010) *Cognitive Science*, 34, 909–957 (Wiley-Blackwell).

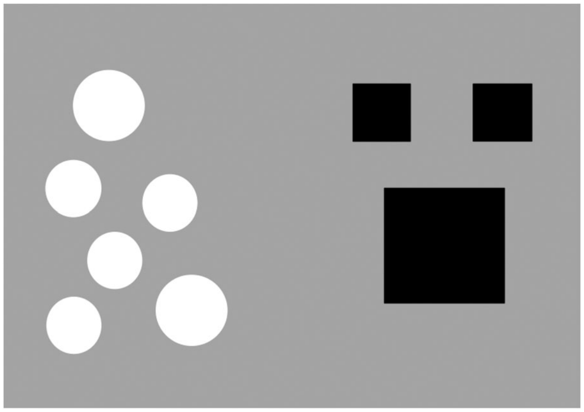

Figure 6. An illustration of the challenges involved in colour and number learning. This picture contains six circles and three squares; white circles and black squares; more circles than squares / less squares than circles; and some of the circles and squares are larger and some are smaller. Yet somehow children must learn the cues that discriminate the appropriate and inappropriate uses of these words in context.

4. A discriminative account of learning and processing in inflection systems

Returning to the central concern of this paper – the learning and processing of inflectional morphology – it is worth highlighting a consistent finding in the studies reported in the last section. The children in the colour and number word learning experiments learned to use colour and number words correctly in the appropriate contexts when they were trained using post-nominal constructions, and the adults in the Feature-Label learning condition in the categorisation experiments learned to classify the objects appropriately. Yet in none of these experiments did the simulations or the theoretical analysis predict that learners were extracting discrete "concepts" that corresponded to the "meanings" of colour or number words, or "fribble categories" in the categorisation experiment. Rather, in every case, what was learned in each simulation was a network of mappings that allowed it to use colour / number / category labels in context. There is little reason to presume that what the participants learned in the experiments differed in this regard.

With regards to the debate on the nature of inflectional processing, these findings are notable for their implications about the way that learning should be conceptualised in this task. Returning to the positions described earlier, it seems clear that regardless of whether researchers may have argued that inflection was best captured by a single- or dual-route model, etc., the one thing they have always agreed on is that what children learn in the course of morphological development are ways of composing and decomposing morphemes (Seidenberg & Plaut, 2014). To take, for example, a child learning English nouns: the dual-route and the single-route theories *both* assume that a child must learn a set of associations between forms and meanings, and then learn how to transform one pattern that can be formed from these associations into another pattern. The child must learn an association between the concept mouse and 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 plurality (or sets of objects but excluding multiple mouses, etc.) and a morpheme + s, etc. Then the child must learn the process by which plural forms are produced by transforming a singular form into a plural. Indeed, it was exactly this process that the Rumelhart and McClelland "pattern associator" model was designed to emulate.

Yet the discrete, bidirectional system of mapping between forms and meanings envisaged in this account is difficult to reconcile with the highly interconnected systems produced by the actual learning processes described in the previous section. Not only is it the case that error-driven learning models do not learn a set of binary pairings between each "meaning" and each form, it is also the case that the learning process is lossy. The discriminative nature of error driven learning means that the dimensions that are actually mapped to any given form will invariably be abstract – and indeed, often highly abstract – such that the patterns of mapping actually produced by learning will hardly correspond to a "meaning" at all (Ramscar et al., 2010). Further, and equally importantly, the neat picture of form-meaning mapping imagined in transformational accounts of morphological processing is inconsistent with research into human categorisation, which has shown that human performance on categorisation tasks is best accounted for by models that do not actually contain representations of pre-established (or even determinable) categories at all. Rather, consistent with the models described in the previous section, the models that best capture human performance in categorisation tasks treat it as a process in which a set of discrete outcomes such as labels, responses, etc., are discriminated in context from a more continuous system of inputs (e.g. Love et al., 2004; see Ramscar & Port, 2015 for a review).

In a similar vein, coding theory indicates that while it is possible for a relatively low dimensional system such as the form contrasts that comprise a language to be losslessly encoded in a high dimensional system like real world context, the converse is impossible (Ramscar et al., 2010). This is consistent with the work reviewed in the previous section, which showed that although people found learning to use context to discriminate the mappings to a set of forms a somewhat straightforward task, they consistently failed to learn the reverse mappings (see Apfelbaum & McMurray, 2017; Hoppe et al., 2020; Nixon, 2020; Ramscar et al., 2013c and Vujović et al., 2021 for replications of this basic pattern of findings in different experimental settings). Accordingly, while it makes sense to assume that children can learn to use context to encode particular linguistic forms in relation to the aspects of the world that are relevant to them (and hence their intended meanings, Ramscar, 2021), the idea that they can learn to use linguistic forms to encode meanings in the bidirectional way that linguists have traditionally imagined actually makes little sense from the perspective of what is known about learning (Ramscar et al., 2010; Ramscar & Port, 2016).

Further, if we allow that morphological systems are not sets of discrete mappings between "units of meaning" and lexical forms, some peculiarities in the framing of the debate reviewed earlier become apparent. Not the least of these is the assumption – which lies at the very heart of the Rumelhart and McClelland model – that the goal of learning to process inflectional morphology is that of mastering transformational rules that add a discrete English past tense morpheme + ed to a verb stem to generate a past tense form, or a discrete plural morpheme + s to a singular noun stem to generate a plural. As noted above, a consequence of this assumption, the training set of the Rumelhart and McClelland model comprises a list of uninflected stems that are transformed into past tense forms, as if the learning environment contained speakers producing repetitive bursts of present-past tense verb forms or singular-plural noun forms. Yet as noted earlier, not only is it unrealistic to assume that children learn inflection from adults who wander around chanting, *go-went, dog-dogs*, *talk-talked, mouse-mice*, etc. but it has long been accepted that children do not in fact learn like this. What children actually hear and learn from are sentences like, *shall we go walk the dog*? (Gleitman, 1965).

In other words, not only is the learning scenario assumed by Rumelhart and McClelland implausible, but critically, the theoretical account of inflection learning embodied in their model (and the many models that followed it) was inherently compositional, albeit that it was implemented using a discriminative learning algorithm. These considerations thus raise an obvious question. How exactly should the problem of inflection learning be conceptualised if one assumes that children learn language – and how to process inflectional patterns – discriminatively?

Given that the way that children actually encounter morphological variations is by coming across their use in normal discourse, and given that normal discourse contains few contexts that offer evidence for transformation – i.e. children neither encounter nor learn from contexts in which they hear adults reciting, *break-broke, car-cars*, etc. – then a straightforward way of recasting the task of learning inflection is as follows: the problems children face is that of learning what it is about the environment that warrants the use of particular forms in particular contexts during normal discourse. This is the approach taken in a series of models proposed by Ramscar and Yarlett (2007), Ramscar and Dye (2009) and Ramscar et al. (2013b) that sought to use delta-rule learning to explain the patterns of development typically associate with children learning English singular and plural noun morphology.

As with the past tense, children's irregular plural production seems to follow a "U-shaped" developmental trajectory, such that children who have produced "mice" in one context will still produce over regularised forms like "mouses" in others. However, the English plural system appears to be difficult to learn even as compared to the past tense system, which likely reflects differences in the input. Children encounter more inflected than uninflected verb forms, and more irregular than regular forms (by token), whereas singular forms dominate the distribution of noun forms, and most plural types and tokens are regular. Accordingly, the period during which children over-regularise noun plurals is far more protracted than is the case for verbs (Graves & Koziol, 1971; Ramscar & Yarlett, 2007), a factor that makes plural learning a particularly fruitful domain for intervention studies.

The analysis and results discussed so far indicate that lexical learning involves discriminating the cues to the use of forms in ordinary discourse contexts, and that patterns of morphological variation reflect similar semantic and contextual factors (Ramscar, 2002; Ramscar & Dye, 2011). Accordingly, the models of morphological development proposed by Ramscar and Yarlett (2007), Ramscar and Dye (2009) and Ramscar et al. (2013b) assume that children encounter morphological variations in context, and are faced with the task of learning to discriminate the cues present in these contexts that are informative about the use of the different morphological variants that occur in them. From this perspective, children do not learn to transform present tense forms into past tenses or singulars into plurals. Rather they must learn to discriminate the sets of contextual cues that are associated with the use of different form contrasts. It is thus worth emphasising that these models employ essentially the same learning rule as the Rumelhart and McClelland (1986) model. Where they differ is in the way that they represent the learning task, and the assumptions that they make about the goal of linguistic knowledge acquisition itself (see Broeker & Ramscar, 2020, for further discussion of the influence of these representational assumptions on the performance of learning models).

The representations of the learning task and the environment in the Ramscar and Yarlett (2007), Ramscar and Dye (2009) and Ramscar et al. (2013b) models assume that initially any kind of stuff in the world is potentially informative about any lexical contrast. Given this assumption, children learning the cues to the singular and plural forms of nouns will be faced with the task of discriminating the more specific cue dimensions in the objects associated with a given form from the other, less specific dimensions those objects will also inevitably comprise. That is, for example, children must learn that *mousiness* is a better cue to the form "mice" than *stuff*. Similarly, learning to discriminate singular from plural forms will involve learning the dimensions of numerosity that best discriminate singular and plural forms (i.e. that multiple mouse objects best predicts "mice").

Figure 7(A) highlights some of the potentially informative dimensions that covary with the irregular plural form "mice" and are thus potential contextual cues to this form. Critically, however, although these dimensions obviously all co-occur with "mice" at the same rate (because they are dimensions of mice themselves), their covariance with other nouns will necessarily differ, and this will result in cue competition. Because generic cues like stuff will be reinforced whenever "mice" is encountered, this will cause learners to expect "mice" to occur whenever stuff is present. This will lead to prediction-errors in contexts where stuff occurs, but "mice" is not heard, which will cause the value of these generic cues to wane over time, and multiple mouse-items to be learned as the best cue to "mice".

The reinforcement and unlearning of the various environmental cues to any given form will be determined by the way that the cues and forms are distributed in the learning environment. Figure 7(B) shows how the potential cues to *mice* overlap relative to a set of idealised cues representing the various contrasting properties of different singular and plural forms in learning, i.e. irregular forms, regular singulars, and the final contrast + *s* that is a common feature of regular plurals. Although the plural forms classed as "regular" in English can differ slightly (by employing different sibilant allomorphs of their final sibilant), broadly speaking they are similar in that they all end in a final sibilant that discriminates plural from singular forms. By contrast, the irregular singular and plural forms differ from one another – and are thus discriminable – in various ways. It is thus important to note that in these models, regular plural forms such as "rats" are not assumed to comprise a stem that is inflected for plurality by adding + s. Rather, *rat* and *rats* are assumed to be different word forms, and it is assumed that children must learn to discriminate one from the other, and the contextual cues appropriate to the usage of each. At the same time, because the forms *rat* and *rats* are only discriminable on the basis of their final sibilant, and because they appear in highly similar contexts, the specific "ratty" cues to them are more difficult to discriminate than they would be if their forms did not overlap quite so much. Meanwhile, the fact that the same final sibilant discriminates the plural and singular forms of a great many words means that it is far more difficult to discriminate the specific semantic cues to the final sibilant in *rats, cats and bats* than it might otherwise be.

Given the distribution of cues and forms – in which regulars are far more frequent than irregulars – a learner will initially come to expect plural forms that end in sibilants whenever sets of objects are described. This over-general expectation causes interference when irregular plurals are produced, and this leads to the production of over-regularised forms. However, further exposure to the same distribution will serve to eliminate this interference. This is because the same generic cues that lead to over-regularisation must inevitably also produce expectations for irregular forms (*mice*) in contexts where regular forms will be used. The prediction errors that result from this will cause the unlearning of these generic dimensions as cues to *mice*, increasing the relative strength of more specific cues, and reducing the likelihood of future over-regularisation.

Behaviourally, of course, we know English-speaking children typically do go through a period of saying "mouses" in childhood, and as they grow older, they eventually learn to produce only the adult form, "mice". The Ramscar et al. models 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. This in turn leads to an important question – how will the process described above proceed when children themselves produce language?

When it comes to learning to generate forms, two scenarios might be envisaged. In the first, the act of producing a word form would serve to reinforce itself, such that when the competing influences present in a child's

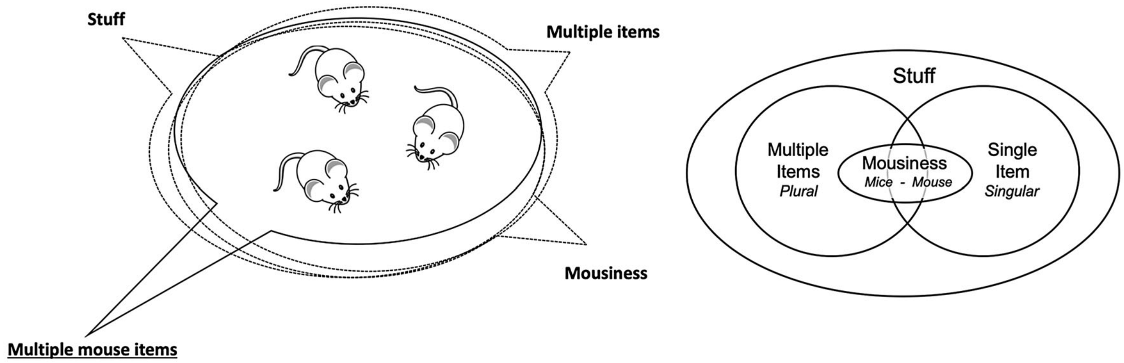

Figure 7. (A) Some of the semantic / contextual dimensions that will be reinforced by a child's exposure to the word "mice" in context. (B) A more abstract representation of the relative specificity of the four dimensions as cues to the forms that comprise singular and plural nouns. While less specific cues will receive positive reinforcement early in learning, because of their ubiquity they will produce more error than the uniquely informative cue multiple mouse items. Cue competition will thus cause the influence of the less specific cues to wane as children's experience grows.

underlying model of the world leads to their saying "mouses", the form "mouses" is reinforced as a response. In the second scenario, production would be driven by the child's model of their intended behaviour (their intention to simply repeat what they have observed), and it is the factors that lead the child to produce the behaviours that are reinforced in the underlying model, not whatever noisy behaviour emerges from them.

This latter perspective (which is often called *model-based learning*, Dayan & Berridge, 2014) assumes that a child who erroneously produces "mouses" was in fact trying to say "mice", simply because "mice" was the form they had heard previously in multiple mouse contexts. Accordingly, because "mice" was the form they had stored in their memory as part of their model of the world, what gets updated in learning is the child's model – and the representations that led to the child's intended behaviour – rather than the behaviour that results from a particular instance of a child "running their model". (In other words, the models assume that learning in children proceeds in much the same way as it proceeds in a pair of trainee ballroom dancers, in whom practice leads the performance of the right steps, and not merely their continually treading on each other's toes.)

Using a computational simulation to demonstrate how learning a model of the world in this way can lead to behaviours that wax and wane as internal representations develop, Ramscar and Yarlett (2007) showed how an error-driven model learning from the distribution of plural forms in English predicted that at an appropriate stage in learning, the elicitation of over-regularised forms from children would *reduce* likelihood that they would over-regularise in the future. In a series of experiments, Ramscar & Yarlett then showed that children exhibited the behaviour predicted by the model: When seven-year-old children repeatedly produced the same plurals across blocks of trials, their rates of over-regularisation went down in later blocks. This behavioural change was observed even when children were given positive feedback on the incorrect forms they produced, lending further support to the idea that learning reinforces children's models of the world, and not simply their behaviour.

Ramscar and Dye (2009) and Ramscar et al. (2013b) then presented a series of models that indicated that when the challenges facing a child learning noun forms are explicitly set up in the way shown in Figures 1 and 2, the distribution of forms and semantics in English invariably leads to what has been described as "U-shaped" performance in plural production. These models predicted that children's mastery of correct irregular forms would be preceded by a phase in which both correct and incorrect irregular plurals were produced. Notably, they further predicted that the ultimate elimination of the interference that gives rise to over-regularisation would be driven by the error caused by the inappropriate expectation of irregular forms (e.g. *mice*) when the semantics of regular forms are present in a lexical context. In other words, the models showed how the same non-discriminative semantic dimension that causes children to expect a sibilant final form in an irregular context – leading to over-regularisation – also causes them to expect irregular forms in regular contexts, gradually causing the non-discriminative semantic cues to be unlearned as cues to irregulars, and thus reducing over-regularisation.

Further, unlike in the original Rumelhart and McClelland model, this U-shaped pattern of learning did not rely on manipulating the input in a controlled simulation. To demonstrate this, Ramscar et al. (2013b) trained a version of their model on a set of plural and singular noun forms taken from a corpus of child directed speech. Even when the various plural and singular noun forms were presented to the model in the same sequence in which they occurred chronologically, Ramscar et al. (2013b) et al. found that the model's production of irregular forms followed a U-shaped pattern of development.

Importantly, and as is hopefully clear from the foregoing, the error-driven nature of the learning that led to the U-shaped developmental pattern in these models makes a clear prediction. If we were to engage children at an appropriate stage of development in a task that required them to invoke the semantics of regular forms in its performance, then this ought to result in learning that will *reduce* their production of over-regularised irregular forms, both in the absence of feedback, *and* in the absence of their being exposed to any further irregulars. To test this, Ramscar et al. first pretested children on a task that elicited both regular and irregular plural and singular forms. One group of children then performed a control task (colour naming), while the other experimental group performed a memory task involving the same set of regular plural forms that had been used in the elicitation task (along with some other regular lures). The children were then post tested using the same elicitation task. Consistent with the detailed predictions of the models, the colour task had no effect on children's over-regularisation rates. However, younger children in the experimental condition showed a small but significant increase in over-regularisation. By contrast, the same exposure to the semantics of regular plural forms brought about a large and significant *decrease* in over-regularisation in the older children in this condition. That is, if one assumes that children use error driven learning to discriminate the contextual cues to forms, the patterns observed in children's over-regularisation and their retreat from it were exactly as one would predict given the way that children learn, the forms that they learn, and the way that these forms are distributed across contexts.

Further, and perhaps most importantly, given that it has long been observed that English speaking children typically go through a phase of over-regularizing irregular nouns and verbs, and given that it is widely accepted that they eventually stop doing this in the absence of explicit feedback (U-shaped learning), it follows that while the predictions and findings reported above are somewhat unintuitive they ought not to come as a complete surprise. What other plausible explanation for this pattern of behaviour is there aside from learning? Even if the workings of model-based, error-driven learning processes are themselves somewhat unintuitive, it should not be a surprise that the developmental patterns exhibited by children reflect the way that the models of the world – and in this case, language – that they learn actually develop; and the mechanisms that enable them to learn these models actually work.

5. Why do languages have regular and irregular forms? A discriminative learning perspective on morphological structure

Perhaps because linguists have long tended to assume that language processing is rule-governed, and that the existence of regular morphological patterns are evidence for rules, it has also been assumed that regularity is somehow a desirable or even normative goal for morphological systems. This view that supposes that irregular paradigms represent deviations from the uniform patterns that systems (or their speakers) seek to maintain, and from this perspective, the existence of phenomena like suppletion, in which an inflected stem-change produces an unpredictable or seemingly unrelated allomorph (e.g. "mouse" / "mice") can seem somewhat puzzling (Blevins, 2016; Blevins et al., 2017; these supposed anomalies are explained away as being the non-productive leftovers from previous systems that were in themselves regular, Pinker, 1999).

By contrast, from the perspective of the discriminative account of morphological learning and processing described here, the difference between overtly suppletive forms (such as "mouse" / "mice") and more regular forms (such as "rat" / "rats") is simply one of degree. "Mouse" and "mice" are simply more discriminable than their regular counterparts, this in turn serves to accelerate the rate at which a learner's representation of the specific form/meaning contrasts they communicate becomes discriminated from other forms communicating similar meanings. The underlying logic of the models described above is that all learning serves to increase the level of suppletion in a system of meaning that is mapped onto a system of form contrasts, such that over time, the degree of suppletion in even regular plurals like rats will increase as a result of cue competition (Ramscar et al., 2018). However, the degree that forms and meanings overlap will tend to modulate the rate at which this occurs, with high degrees of form similarity decelerating this process, and low degrees accelerating it (see e.g. Tomaschek et al., 2021; Wedel et al., 2013).

From this perspective, suppletive irregular forms are not categorically different types, as the dual-route theory described earlier maintains, but rather they are extreme instances of the system of contrasts that all linguistic communication relies on. Given this, it would appear that from a discriminative perspective, it is not the existence of irregular forms that is a puzzle, but rather what stands in need of explanation is the regularity that is seen ubiquitously in the morphological systems of the world's languages (Blevins et al., 2017). Once again, learning suggests an answer to this question. Although it is common to describe children as language learners – such that, by implication, one might assume that adults are not – the nature of linguistic distributions guarantees that language learning continues throughout the lifespan, such that in the modern world at least, no child ever learns the communicative code of a natural language in its entirety (Ramscar, 2021; Ramscar et al., 2014). It is from this perspective that the tendency of languages to organise morphological forms systematically, such that they form neighbourhoods, begins to make sense. In the models of plural morphology described in the previous section, the distribution of the regular plural forms inhibited the unlearning of the generic meaning cues to the regular sibilant feature, and as a consequence of this, the "correct" forms of unattested regular plurals were predicted (i.e. produced) as a virtuous result of exactly the same process of distributional learning that caused children to over-regularise irregular plurals. This meant that the set of input-output relationships learned by these models effectively served to implicitly encode the forms of regular noun plurals even before they were encountered. This finding offers an alternative perspective on the coexistence of regular and irregular patterns in morphology. Given that a side effect of regular patterns is that they reduce the rate at which the discrimination of the more specific cues to individual forms occurs during learning, it follows that the organisation of *morphological forms* into neighbourhoods will serve to allow learners to *implicitly* learn the forms of large numbers of less frequent lexical items (Blevins et al., 2017; Ramscar et al., 2018; see also Marzi, Ferro & Pirrelli, 2019).

Thus not only can a discriminative learning perspective offer an explanation for the various patterns of behaviour observed as children master inflection systems, it can provide an account of why these systems comprise both regular and irregular forms in the first place. Regular and irregular patterns in morphology represent a socially evolved trade-off that balances the opposing communicative pressures of discriminability and learnability described above. The existence of frequent, well-discriminated irregular forms serves to enhance the marking and learning of important communicative contrasts. Meanwhile, it is the fact that forms are not fully discriminated from one another in regular neighbourhoods that makes them productive. An important benefit this brings is that it provides a means for filling the inevitable gaps that will necessarily occur in any individual language learner's experience (Blevins et al., 2017; Ramscar et al., 2018).

6. Can learning – and its development – help explain why human communicative capacities appear to be unique?

The analyses and results described so far in this review suggest that Miller (1967) may have been premature in concluding that linguistic knowledge cannot by acquired as a form of conditioning. Rather, following on from Rumelhart and McClelland (1986), they indicate that in learning to communicate children simply learn a complex set of conditioned relationships between contexts and linguistic forms. Further, they suggest that a primary mechanism underlying the acquisition of these relationships is error-driven learning. However, if communicative conventions—and language—are simply the product of learning, this raises a question noted at the outset of this review: why is linguistic communication the sole preserve of humans? Why, given that they share the same environment, do babies learn language but not puppies? One obvious source of answers to these questions comes from considering the differences between the way that the brains of humans and other animals develop, and the impact of these developmental differences on learning.

Like many other primates, humans are born with an immature brain, which experiences a massive proliferation of synaptic connections (synaptogenesis), followed by an extended period of pruning (synaptic elimination) after birth. However, the rate of these developments is markedly different in humans as compared to other primates. Postnatal brain development in monkeys happens at largely the same rate in all cortical areas (Rakic et al., 1986), taking around 4 years in rhesus monkeys (*Macaca mulatta*, Malkova et al. (2006)). Human cortical development is far more uneven. In the auditory cortex synaptogenesis peaks after three months, and synapse elimination ends by age 12 (Huttenlocher & Dabholkar, 1997). By contrast, in prefrontal cortex, the last function region to reach maturity, the synaptic overproduction-elimination process continues into the third decade of life (Petanjek et al., 2011).

In adults, the prefrontal cortex supports mechanisms that guide attention and select responses in context (Yeung et al., 2004), functioning as a "dynamic filter" for goal directed behaviour (Shimamura, 2000). Young children's lack of this functionality can be seen in their performance on selection tasks, such as guessing which hand a piece of candy is in when the hands are biased 25:75. Children younger than 5 years old simply fixate on the high-probability hand, choosing it at a rate that overmatches its probability. After age 5 they then begin to switch between the hands and to probability match, a strategy that continues that into adulthood (Derks & Paclisanu, 1967). Given that probability matching reduces the amount of candy won, this is one situation where children's inability to think flexibly is an advantage.

Another area in which cognitive inflexibility may – by definition – be advantageous is convention learning (Ramscar & Gitcho, 2007; Thompson-Schill, Ramscar, & Chrysikou, 2009). Human communication relies on the acquisition of structured, conventionalised knowledge, and it seems likely that conventionalised knowledge will be *far* more likely to be acquired if the learning process is itself conventionalised. This is exactly what is likely to happen when learners are unable to filter their attention to the input. Consistent with this suggestion, it is notable that in contrast to children, adults struggle to master the linguistic conventions of new languages (Johnson & Newport, 1989). From this perspective, this reflects the fact that increases in learners' ability to selectively respond or attend to the world will decrease the overlap in what they learn. In this regard, it is further notable that despite the many claims made about the effects of maturation on learning "syntax" (and the many claims about the relative triviality of inflectional morphology when it comes to explaining "language" *tout court*), studies have shown that when it comes to learning English as a second language, mastery of inflectional morphology – learning to say "mice" as opposed to "mouses" – is the aspect of grammar that is most negatively affected by age of acquisition (Johnson & Newport, 1989).

To return to the model-based approach to inflection learning described above, the reason why children's production of erroneous forms like "mouses" did not negatively affect their learning of the conventionalised form "mice" in those models was because the models assumed that when it came to learning, children ignore their actual behaviour, and instead reinforce their representations of the forms that they intend to produce. Given the functional role prefrontal areas play in goal-directed and stimulus-driven attention, short-term memory and response selection, (Asplund et al., 2010; Nee & Jonides, 2011; Shimamura, 2000), this assumption was based on the idea that young children lack the ability to do otherwise. By contrast, because the human brain becomes ever better at performing tasks such as self-directed attention, holding information in short-term memory and selecting amongst competing responses as its prefrontal cortex matures, it becomes ever harder for learners to ignore the fact that the forms like "mouses" are actually emerging from their mouths. Consistent with this proposal, the development of the various functions supported by prefrontal cortex emerge and strengthen progressively between roughly the third and twentieth year in neurotypical humans (Diamond, 2002; Ramscar & Gitcho, 2007; Thompson-Schill et al., 2009; Tsujimoto, 2008). This period corresponds exactly with a linear decline in learners' ability to master English inflectional morphology. By contrast, the diminished chances of an individual mastering a system of irregular nouns and verbs do not change once adulthood is reached (Johnson & Newport, 1989; see Harmon & Kapatsinski, 2021, for a discussion of the ways in which top-down influences could change adult learning in these circumstances).

In other words, it seems that the less children are able to direct their attention in learning, the more their learning will approximate the basic model-based error-driven learning process described earlier. The representations that children learn will be more straightforwardly shaped by their immediate physical, social, and linguistic environments than those of adults, such that children's learning will be more conventional (i.e. the expectations that they learn about events will tend to converge; see also Finn et al., 2014; Friederici et al., 2013; Hudson Kam & Newport, 2005, 2009). Given that adults were once children, these considerations can thus also explain why other animals—which appear to be able to selectively attend to their environments and filter their responses from almost the moment they are born—fail to learn much by way of complex, conventionalised social and linguistic behaviour. They may even allow us to begin to describe more precisely what the biological endowments that enable humans to acquire language actually are.

7. The past tense debate redux

This review has argued that empirically, the main ideas behind the Rumelhart and McClelland (1986) past tense model – that inflectional patterns are learned, and that rule-like behaviour emerges out of the distribution of the forms in a language – are largely correct (see also Seidenberg & Plaut, 2014). At the same time, however, it has argued that much of the theory that lay behind Rumelhart and McClelland's model was wrong: its conception of the information that was actually relevant to inflection learning; its conception of the actual learning task; and its conception of how human communication actually works.

The first of these shortcomings is the least important. It is in the nature of models that they cannot be exhaustive in the information they capture, and at some level every model of a cognitive process is literally false (Box, 1979; Seidenberg & Plaut, 2014). Although processing in the Rumelhart and McClelland model relied entirely on discrete phonetic information, there is in principle no reason why information about the semantic / contextual structure of the task could not have been included in it. In the same vein, it would be reasonable to describe the level of phonetic detail in the Ramscar et al. plural models as impoverished at best – as configured, these models cannot account for irregular productivity in WUG/RICK tasks – simply because they were designed to highlight the role of the contextual factors that influence inflectional learning and its U-shaped development, not to simulate every detail of language learning.

The other shortcomings of the Rumelhart and McClelland model are, however, more serious. The last two problems described above are intimately related to nature of learning networks, which can be understood as systems that learn to represent the statistical structures underlying human cognitive capacities, such as our ability to learn and use inflectional rules (Seidenberg & Plaut, 2014). Section 4 described how the idea that children use these systems to extract the statistical structure relevant to processing inflections in context is hard to reconcile with the transformational model of inflectional processing implemented in the Rumelhart and McClelland model. Yet this transformational model was accepted in an almost unquestioning manner on all sides of the past tense debate (such that even where connectionist models did include semantic information, it was inevitably employed to help model a process in which root forms were transformed into past tense forms, see e.g. Joanisse & Seidenberg, 1999).

If one accepts that children use some kind of *learning network* to extract the statistical structures that predict the appropriate use of language in context, then an analysis of morphological processing that involves retrieving a form-meaning pair representing a stem and then transforming it into an inflected form makes little sense. Rather – in a somewhat ironic parallel of Rumelhart and McClelland's point that the distribution of forms could make the idea of rules redundant – it seems that if one accepts that the task faced by children is that of learning to use forms in context, the idea that they must learn form - meaning "pairings" (that must retrieved and transformed) becomes similarly redundant (Ramscar, 2021; Ramscar & Port, 2015). That is, if one assumes – as it seems learning models should assume – that children learn to discriminate the various statistical structures (environmental dimensions, form distributions, etc.) that serve as the cues to forms in their appropriate contexts (which is something akin to the process that many linguists call "pragmatics"), it follows that in doing so, they will simultaneously be learning the cues appropriate to any given form's use. Accordingly, given that children will be using distributional information to learn how to anticipate and use forms in context then from the perspective of describing how language is processed psychologically, the idea that children actually learn word meanings (what many linguists call "semantics") is entirely redundant. Rather, a detailed model of how children learn to use words will make the idea of "word meanings" redundant in descriptions of language processing in the same way that Rumelhart and McClelland sought to make the idea of "rules" redundant in these descriptions. (Such a model would put flesh and bones on Wittgenstein's (1953) famous comment that a word's meaning is its use.)

Critically, when seen from this perspective, it becomes clear that although Rumelhart and McClelland's past tense model embraced a distributional analysis of forms, its analysis of the roles of context and meaning – evidenced by its embracing a view of the inflection process that revolved around static inventories of form meaning pairings and transformations – was still firmly rooted in a non-distributional, somewhat generative view of how language works. That is, the Rumelhart and McClelland model embodied a very different analysis of form as compared to function; and on examination it seems clear that these analyses are irreconcilable. Moreover, and perhaps because the Rumelhart and McClelland model could be thought of as unfinished at a theoretical level (Seidenberg & Plaut, 2014), the somewhat half-baked view of inflection that it implemented persists today. As learning models have become ever more powerful, researchers still seek to employ their latest iterations in order to somehow make a version of the inflectional process implemented in the Rumelhart and McClelland model fit the empirical facts (Gorman et al., 2019; Kirov & Cotterell, 2018), albeit that the results to date are far from satisfactory (Corkery et al., 2019; Perconti & Plebe, 2020).

By contrast, this review has suggested that the reason why these models do not fit the empirical facts is because they are modelling the *wrong* process; and the reason that it is the wrong process is because it confuses what are two very different levels of analysis. The evidence reviewed here indicates that generating a past tense is not a transformational process. Rather, the process of "generating a past tense" is not really so different from that of "generating" any other word in context. From this perspective, learning a language involves learning the relationships between distribution of competing forms (at different levels of abstraction) across a distribution of contexts, and this review has described the successful predictions that models based on this approach to language learning make about the development of inflectional processing.

Further support for these ideas in relation to morphological processing comes from studies that have copied and adapted this approach in studies involving written word forms (where orthography also provides an addition source of contextual cues to forms). These have shown that treating reading as one of discriminating the appropriate cues to competing written forms in the contexts that they occur can allow models to fit, for example, the complex patterns of reading time data associated with words that occur in different morphological patterns and neighbourhoods (Baayen et al., 2011), response times in reading aloud (Hendrix et al., 2019), and response times in lexical decision tasks (Milin et al., 2017). Additional support for the idea of conceptualising learning as a discriminative process that occurs in context comes from studies of how children learn spoken forms themselves, which have shown how this analysis can successfully offer insight into the way that children extract acoustic features such as vowel / consonant pairs from the speech stream itself (Nixon & Tomaschek, 2021), as well as explaining why this task is harder for older learners (Nixon, 2020). This approach has also been successfully applied to modelling learning in developmental language disorder (DLD; Freudenthal et al., 2021), showing how the marked cross-linguistic differences the have been observed in children's acquisition of verb inflection can arise out the way that the specific properties of an input language affect children with DLD's ability to discriminate predictive relations between separated elements in sequences.

7.1. Why this debate needs to move beyond its past

The past tense debate went wrong because the Rumelhart and McClelland model – and the many other connectionist models that followed it – did not follow through on its analyses. The original past tense model sought to show how rules could emerge out of the statistical, distributional structure of language. Yet it (and most of those many other models) embodied a number of traditional assumptions that are difficult to reconcile with the error driven learning mechanisms the model itself was based on: that morphology involves transformation; that form-meaning mappings are bi-directional; and that words have discrete meanings. Perhaps unsurprisingly, the empirical evidence indicates that these assumptions are also incompatible with the way that learners extract the statistical conventions that govern language use from the distribution of context and form relationships they are exposed to. In a similar vein, is it far from clear that the traditional communicative process envisaged in transformational models – in which meanings are "encoded" and "decoded" by building up (or breaking down) smaller elements into larger wholes – is at all consistent with the dynamic, predictive nature of learning mechanisms, or the kind of communicative processes that they can support.

Describing an alternative, learning based conception of communication in detail is beyond the scope of this review (see Ramscar, 2021, for an outline). However, as others have noted, it is currently the case that the success of the many language engineering models that followed in the steps of Rumelhart and McClelland have outpaced the development of linguistic theories that are in tune with or can even explain the statistical, distributional processes they implement (for different perspectives on this, see e.g. Futrell et al., 2019; Lake & Baroni, 2018; Linzen et al., 2016; Papadimitriou et al., 2021; Perconti & Plebe, 2020). From the perspective described here, this disjoint is simply a sign of unfinished business. The past tense debate may have been one of the most important theoretical disputes in the history of cognitive science, but it was not so much resolved as it fizzled out. In the light of the foregoing, the reasons for this may appear obvious: the debate revolved around on one hand, a high level view of language that could not adequately describe the detailed facts of language learning and processing, and on the other, a learning approach that only went half way. The solutions might also appear obvious, even if detailed answers may not be. Making sense of human communication from the point of view of error driven learning will involve carrying on the revolution begun by Rumelhart and McClelland to its conclusion. And it will involve applying a coordinated statistical, distributional approach based on learning to every aspect of language, not just the acoustic patterns that are predicted when a given context suggests that a given signal should have an "inflected" form (or not).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Deutsche Forschungsgemeinschaft: [grant number RU 2718, Research Unit "Modal and Amodal Cognition: Functions and Interactions"].

ORCID

Michael Ramscar http://orcid.org/0000-0003-1680-1112

References

Alegre, M. A., & Gordon, P. (1996). Red rats exposes recursion in children's word formation. *Cognition*, *60*(1), 65. https://doi.org/10.1016/0010-0277(95)00703-2

Apfelbaum, K. S., & McMurray, B. (2017). Learning during processing: Word learning doesn't wait for word recognition to finish. *Cognitive Science*, *41*(S4), 706. https://doi.org/10.1111/cogs.12401

Asplund, C. L., Todd, J. J., Snyder, A. P., & Marois, R. (2010). A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nature Neuroscience*, *13*(4), 507. https://doi.org/10.1038/nn.2509

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438. https://doi.org/10.1037/a0023851

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*(2-3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford University Press.

Blevins, J. P., Milin, P., & Ramscar, M. (2017). The zipfian paradigm cell filling problem. In F. Kiefer, J. P. Blevins, & H. Bartos (Eds.), *Perspectives on morphological organization* (pp. 139–158). Brill.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 115. https://doi.org/10.1037/0096-1523.2.1.115

Box, G. E. (1979). Robustness in the strategy of scientific model building. In R. L. Launer, & G. N. Wilkinson (Eds.), *Robustness in statistics* (pp. 201–236). Academic Press. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Bröker, F., & Ramscar, M. (2020). Representing absence of evidence: Why algorithms and representations matter in models of language and cognition. *Language, Cognition and Neuroscience*, 1–24.

Brown, R. (1973). Development of the first language in the human species. *American Psychologist*, *28*(2), 97. https://doi.org/10.1037/h0034209

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 11–53). Wiley.

Chomsky, N. (1985). *Knowledge of language: Its nature, origin, and use*. Praeger.

Corkery, M., Matusevych, Y., & Goldwater, S. (2019, July 28-August 2). *Are we there yet? Encoder-decoder neural networks as cognitive models of English past tense inflection*. Proceedings of the 57th Annual Meeting of the Association for Computational linguistics, pp. 3868–3877.

Daw, N. D., Courville, A. C., & Dayan, P. (2008). Semi-rational models: The case of trial order. In N. Chater, & M. Oaksford (Eds.), *The probabilistic mind* (pp. 473–492). Oxford University Press.

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 473-492. https://doi.org/10.3758/s13415-014-0277-8

Derks, P. L., & Paclisanu, M. I. (1967). Simple strategies in binary prediction by children and adults. *Journal of Experimental Psychology*, *73*(2), 278. https://doi.org/10.1037/h0024137

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. Stuss, & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). Oxford Press.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*(1), 1–24. https://doi.org/10.1093/applin/ami038

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. https://doi.org/10.1371/journal.pone.0101806

Freudenthal, D., Ramscar, M., Leonard, L. B., & Pine, J. M. (2021). Simulating the acquisition of verb inflection in typically developing children and children with developmental language disorder in English and Spanish. *Cognitive Science*, *45*(3), e12945. https://doi.org/10.1111/cogs.12945

Friederici, A. D., Mueller, J. L., Sehm, B., & Ragert, P. (2013). Language learning without control: The role of the PFC. *Journal of Cognitive Neuroscience*, *25*(5), 814–821. https://doi.org/10.1162/jocn_a_00350

Futrell, R., Wilcox, E., Morita, T., Qian, P., Ballesteros, M., & Levy, R. (2019). Neural language models as psycholinguistic subjects: Representations of syntactic state. *arXiv preprint* arXiv:1903.03260.

Gleitman, L. R. (1965). Coordinating conjunctions in English. *Language*, *41*(2), 260–293. https://doi.org/10.2307/411878

Gordon, P. (1985). Level-ordering in lexical development. *Cognition*, *21*(2), 73–93. https://doi.org/10.1016/0010-0277(85)90046-0

Gorman, K., McCarthy, A. D., Cotterell, R., Vylomova, E., Silfverberg, M., & Markowska, M. (2019, November 3–4). *Weird inflects but OK: Making sense of morphological generation errors*. Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL), pp. 140–151.

Graves, M., & Koziol, S. M., Jr. (1971). Noun plural development in primary grade children. *Child Development*, *42*(4), 1165–1173. https://doi.org/10.2307/1127801

Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, *41*(4), 313–360. https://doi.org/10.1006/cogp.2000.0737

Harmon, Z., & Kapatsinski, V. (2021). A theory of repetition and retrieval in language production. *Psychological Review*, *128*(6), 1112. https://doi.org/10.1037/rev0000305

Harris, C. L. (1992, July 29-August 1). *Understanding English past tense formation: The shared meaning hypothesis*. Fourteenth Annual Conference of the Cognitive Science Society, Erlbaum.

Haskell, T. R., MacDonald, M. C., & Seidenberg, M. S. (2003). Language learning and innateness: Some implications of compounds research. *Cognitive Psychology*, *47*(2), 119. https://doi.org/10.1016/S0010-0285(03)00007-0

Hendrix, P., Ramscar, M., & Baayen, H. (2019). NDRA: A single route model of response times in the reading aloud task based on discriminative learning. *PloS One*, *14*(7), e0218802. https://doi.org/10.1371/journal.pone.0218802

Hoppe, D. B., Hendriks, P., Ramscar, M., & van Rij, J. (2021). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*.

Hoppe, D. B., van Rij, J., Hendriks, P., & Ramscar, M. (2020). Order matters! influences of linear order on linguistic category learning. *Cognitive Science*, *44*(11), e12910. https://doi.org/10.1111/cogs.12910

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*(2), 151–195. https://doi.org/10.1080/15475441.2005.9684215

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. https://doi.org/10.1016/j.cogpsych.2009.01.001

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*(2), 167–178. https://doi.org/10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences*, *96*(13), 7592. https://doi.org/10.1073/pnas.96.13.7592

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*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kim, J. J., Marcus, G. F., Pinker, S., Hollander, M., & Coppola, M. (1994). Sensitivity of children's inflection to morphological structure. *Journal of Child Language*, *21*(1), 173–209. https://doi.org/10.1017/S0305000900008710

Kim, J. J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science*, *15*(2), 173–218. https://doi.org/10.1207/s15516709cog1502_1

Kiparsky, P. (1982). Lexical phonology and morphology. In I. S. Yang (Ed.), *Linguistics in the morning calm* (pp. 3–91). Hansin.

Kirov, C., & Cotterell, R. (2018). Recurrent neural networks in linguistic theory: Revisiting Pinker and Prince (1988) and the past tense debate. *Transactions of the Association for Computational Linguistics*, *6*, 651–665. https://doi.org/10.1162/tacl_a_00247

Lake, B., & Baroni, M. (2018, July 10-15). *Generalization without systematicity: On the compositional skills of sequence-to-sequence Recurrent networks*. Proceedings of the 35th International Conference on Machine Learning, pp. 2873–2882, Stockholmsmässan, Stockholm, Sweden, PMLR.

Legate, J. A., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, *14*(3), 315–344. https://doi.org/10.1080/10489220701471081

Lignos, C., & Yang, C. (2018). Morphology and language acquisition. In A. Hippisley, & G. Stump (Eds.), *The Cambridge handbook of morphology* (pp. 765–791). Cambridge University Press.

Linzen, T., Dupoux, E., & Goldberg, Y. (2016). Assessing the ability of LSTMs to learn syntax-sensitive dependencies. *Transactions of the Association for Computational Linguistics*, *4*, 521–535. https://doi.org/10.1162/tacl_a_00115

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309. https://doi.org/10.1037/0033-295X.111.2.309

MacDonald, S., & Ramscar, M. (2001, August 1-4). *Testing the distributional hypothesis: The influence of context on judgements of semantic similarity*. Proceedings of the 23rd Annual Conference of the Cognitive Science Society, Edinburgh.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, *40*(1-2), 121–157. https://doi.org/10.1016/0010-0277(91)90048-9

Malkova, L., Heuer, E., & Saunders, R. C. (2006). Longitudinal magnetic resonance imaging study of rhesus monkey brain development. *European Journal of Neuroscience*, *24*(11), 3204. https://doi.org/10.1111/j.1460-9568.2006.05175.x

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, J. T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), 1–165. https://doi.org/10.2307/1166115

Marslen-Wilson, W. D., & Tyler, L. K. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 823–836. https://doi.org/10.1098/rstb.2007.2091

Marzi, C., Ferro, M., & Pirrelli, V. (2019). A processing-oriented investigation of inflectional complexity. *Frontiers in Communication*, *4*, 48.

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PloS one*, *12*(2), e0171935. https://doi.org/10.1371/journal.pone.0171935

Miller, G. A.. (1967). The psycholinguists: On the new scientists of language. In G. A. Miller (Ed.), *The psychology of communication* (pp. 70–92). Penguin Books.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla–Wagner model. *Psychological Bulletin*, *117*(3), 363–386. https://doi.org/10.1037/0033-2909.117.3.363

Nee, D. E., & Jonides, J. (2011). Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *Neuroimage*, *54*(2), 1540–1548. https://doi.org/10.1016/j.neuroimage.2010.09.002

Ng, A. Y., & Jordan, M. I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes. *Advances in Neural Information Processing Systems*, *14*, 841–848.

Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081. https://doi.org/10.1016/j.cognition.2019.104081

Nixon, J. S., & Tomaschek, F. (2021). Prediction and error in early infant speech learning: A speech acquisition model. *Cognition*, *212*, 104697. https://doi.org/10.1016/j.cognition.2021.104697

O'doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337.

Papadimitriou, I., Chi, E. A., Futrell, R., & Mahowald, K. (2021). Deep subjecthood: Higher-order grammatical features in multilingual BERT. *arXiv preprint* arXiv:2101.11043.

Perconti, P., & Plebe, A. (2020). Deep learning and cognitive science. *Cognition*, *203*, 104365. https://doi.org/10.1016/j.cognition.2020.104365

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M. R., Uylings, H. B., Rakic, P., & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences*, *108*(32), 13281. https://doi.org/10.1073/pnas.1105108108

Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535.

Pinker, S. (1994). *The language instinct*. Morrow.

Pinker, S. (1997). Words and rules in the human brain. *Nature*, *387*(6633), 547–548. https://doi.org/10.1038/42347

Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books.

Pinker, S. (2001). Four decades of rules and associations, or whatever happened to the past tense debate. In Emmanuel Dupoux (Ed.), *Language, the brain, and cognitive development: Papers in honor of Jacques Mehler* (pp. 157–179). MIT Press.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*(1-2), 73–193. https://doi.org/10.1016/0010-0277(88)90032-7

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*(1), 21. https://doi.org/10.1016/0010-0277(93)90057-3

Prasada, S., & Pinker, S. (1993). Generalizations of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*(1), 1–56. https://doi.org/10.1080/01690969308406948

Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., & Goldman-Rakic, P. S. (1986). Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science*, *232*(4747), 232–235. https://doi.org/10.1126/science.3952506

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, *45*(1), 45–94. https://doi.org/10.1016/S0010-0285(02)00001-4

Ramscar, M. (2021). How children learn to communicate discriminatively. *Journal of Child Language*, *48*(5), 984–1022. https://doi.org/10.1017/S0305000921000544

Ramscar, M., & Dye, M. (2009, 29 July-1 August). *Error and expectation in language learning: An inquiry into the many curious incidences of "mouses" in adult speech*. Proceedings of the 31st Meeting of the Cognitive Science Society, Amsterdam.

Ramscar, M., & Dye, M. (2011). Learning language from the input: Why innate constraints can't explain noun compounding. *Cognitive Psychology*, *62*(1), 1–40. https://doi.org/10.1016/j.cogpsych.2010.10.001

Ramscar, M., Dye, M., Blevins, J., & Baayen, H. (2018). Morphological development. In A. Bar On, & D. Ravit (Eds.), *Handbook of communication disorders* (pp. 181–202). Mouton de Gruyter.

Ramscar, M., Dye, M., Gustafson, J. W., & Klein, J. (2013c). Dual routes to cognitive flexibility: Learning and response-conflict resolution in the dimensional change card sort task. *Child Development*, *84*(4), 1308–1323. https://doi.org/10.1111/cdev.12044

Ramscar, M., Dye, M., & Hübner, M. (2013). When the fly flied and when the fly flew: How semantics affect the processing of inflected verbs. *Language and Cognitive Processes*, *28*(4), 468–497. https://doi.org/10.1080/01690965.2011.649041

Ramscar, M., Dye, M., & McCauley, S. M. (2013b). Error and expectation in language learning: The curious absence of "mouses" in adult speech. *Language*, *89*(4), 760–793. https://doi.org/10.1353/lan.2013.0068.

Ramscar, M., Dye, M., Popick, H. M., & 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. https://doi.org/10.1371/journal.pone.0022501

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends In Cognitive Sciences*, *11*(7), 274–279. https://doi.org/10.1016/j.tics.2007.05.007

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*(1), 5–42. https://doi.org/10.1111/tops.12078

Ramscar, M., & Port, R. F. (2015). Categorization (without categories). In E. Dabrowska, & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 75–99). De Gruyter.

Ramscar, M., & Port, R. F. (2016). How spoken languages work in the absence of an inventory of discrete units. *Language Sciences*, *53*, 58–74. https://doi.org/10.1016/j.langsci.2015.08.002

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. https://doi.org/10.1080/03640210701703576

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*(6), 909–957. https://doi.org/10.1111/j.1551-6709.2009.01092.x

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. Implicit rules or parallel distributed processing? In R. McClelland (Ed.)the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 2, pp. 195–248). MIT Press.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*(3), 668. https://doi.org/10.1037/0012-1649.35.3.668

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115.

Seidenberg, M. S., & Plaut, D. C. (2014). Quasiregularity and its discontents: The legacy of the past tense debate. *Cognitive Science*, *38*(6), 1190–1228. https://doi.org/10.1111/cogs.12147

Selkirk, E. O. (1982). *The syntax of words*. MIT Press.

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, *28*(2), 207–218. https://doi.org/10.3758/BF03331979

Siegel, D. (1979). *Topics in English morphology*. Garland.

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*(3), 314–321. https://doi.org/10.3758/BF03210755

Stone, G. O. (1986). An analysis of the delta rule and the learning of statistical associations. In J.A. Feldman, P Hayes, & D.E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 444–459). Press.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*(2), 135. https://doi.org/10.1037/0033-295X.88.2.135

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.

Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. H. (2021). Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, *57*(1), 123–161. https://doi.org/10.1017/S0022226719000203

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *The Neuroscientist*, *14*(4), 345–358. https://doi.org/10.1177/1073858408316002

Vujović, M., Ramscar, M., & Wonnacott, E. (2021). Language learning as uncertainty reduction: The role of prediction error in linguistic generalization and item-learning. *Journal of Memory and Language*, *119*, 104231. https://doi.org/10.1016/j.jml.2021.104231

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179–186. https://doi.org/10.1016/j.cognition.2013.03.002

Widrow, B., & Hoff, M. E. (1960). *Adaptive switching circuits*. Tech. Rep. Stanford Univ Ca Stanford Electronics Labs.

Williams, E. (1981). On the notions “lexically related” and “head of a word”. *Linguistic Inquiry*, *12*(2), 245–274. https://www.jstor.org/stable/4178218

Wittgenstein, L. (1953). *Philosophical investigations*. Blackwell.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931. https://doi.org/10.1037/0033-295X.111.4.931