Patterns and discriminability in language analysis

Farrell Ackerman, Robert Malouf and James P. Blevins

Abstract

Recent developments in the Word and Pattern approach to complex morphology have argued that words and the patterned relations between words are primary objects of morphological analysis. The primacy of words has two part/whole dimensions: the nature of their internal structure and the nature of their external relations to one another. Words consist of constitutive parts and words themselves are parts of larger patterns of systemic relatedness. We argue that internal structure is essentially discriminative, rather than morphemic, i.e., what is crucial for morphological organization is the ability to discriminate (patterns of) words from one another and all types of internal distinctions suffice to facilitate the necessary discriminability to establish patterns of words. The value and the operation of a discriminative perspective on the internal structure of words is also evident in the analysis of an entirely different phenomenon. Greenberg's (1963) Universal 34 states that "No language has a trial number unless it has a dual. No language has a dual unless it has a plural." We present an associative model of the acquisition of grammatical number based on the Rescorla-Wagner learning theory Rescorla & Wagner (1972) that predicts this generalization. Number as a real-world category is inherently structured: higher numerosity sets are mentioned less frequently than lower numerosity sets, and higher numerosity sets always contain lower numerosity sets. Using simulations, we demonstrate that these facts, along with general principles of probabilistic learning, lead to the emergence of Greenberg's Number Hierarchy. The value of a discriminative perspective for language analysis (Ramscar & Yarlett 2007, Ramscar et al. 2010, 2013) becomes clear in both word-based morphology and its explanatory role addressing a typological conundrum.

1. Introduction

New developments in the Word and Pattern[1] approach to complex morphology have argued that words and the patterned relations between words are primary objects of

Word Structure 9.2 (2016): 132–155
DOI: 10.3366/word.2016.0091

www.euppublishing.com/loi/word

morphological analysis. The primacy of words has two part/whole dimensions: the nature of their internal structure and the nature of their external relations to one another. Words consist of constitutive parts and words themselves are parts of larger patterns of systemic relatedness. In the first part of this article we focus on internal structure as essentially discriminative, rather than as morphemic, i.e., what is crucial for morphological organization is the ability to discriminate (patterns of) words from one another and all types of internal distinctions facilitate the discriminability needed to establish patterns of words. The notion that discriminative organization and learning are relevant to language analysis (as argued at length by Arnon & Ramscar (2012), Ramscar & Yarlett (2007), Ramscar & Dye (2010), Ramscar et al. (2010, 2013)) is likely to be unfamiliar to morphologists trained up in the notion that morphology concerns the identification of meaningful bits and their composition in complex words. In order to provide a better sense of the nature of a discriminative perspective on language analysis, in the second part of the article we demonstrate how it is applicable to an entirely different linguistic issue, namely, the perennial conundrum represented by Greenberg's Universal 34: "No language has a trial number unless it has a dual. No language has a dual unless it has a plural." We conclude by suggesting that a discriminative perspective is important for understanding morphological organization and language organization more broadly.

2. Words: Parts and Wholes

It is standard in modern linguistics to conceive of morphology as the study of the internal structure of words, where this internal structure is construed morphemically. That is, words are linearly and hierarchically arranged combinations of meaningful bits called morphemes and the task of morphology is to identify these bits and their lawful concatenations. A traditional alternative to this perspective comes from the classical Word and Paradigm tradition, where words rather than sub-word units, are often cross-linguistically the smallest meaningful elements. This tradition focuses on morphological organization from the following perspective: How are the elements that make up words, both segmental and suprasegmental, organized into (the patterns that characterize different types of) words and how are words organized into larger patterned collections? The analytic effort here is not in identifying meaningful bits, though this obviously obtains in some languages in some instances, but in identifying the manner in which meanings as wholes are distinguishable from related words bearing different meanings.

The general intuition guiding the assumption of bi-uniqueness in the alignment of morphological information and encoding is, in some sense, useful. But the theoretical (and typological) consequences drawn from this are misplaced: the emphasis is generally on the pieces of words. In complex morphological systems, the unit over which bi-unqiueness can be defined often corresponds more closely to the word than to any of its constitutive parts. Accordingly, a morphological perspective based on discriminabilty of forms subsumes all straightforward (agglutinative and analytic) encodings, while also accounting for the many instances where there is deviation, but where related words are discriminably distinct. This reflects the fact that in morphology there is a pervasive tendency for discriminablity between forms, which can be achieved in an extraordinary variety of ways. Viewed from this perspective it becomes clear why the word, as a synthetic and/or periphrastic construct, (often) functions as a fundamental unit of morphological analysis. There are many ways to be discriminable, which exploit phonemic as well as sub-phonemic cues (Baayen et al. 2003, Kemps et al. 2005, Seyfarth et al. 2016). Even syncretism, combining identity of form with difference in morphosyntactic properties, can be useful in organization if it reliably relates the same wordforms with different patterns of associated morphosyntactic properties. This is not plausibly regarded a design feature of grammars in the sense of a factor that inheres in and guides the shapes of grammars. Rather, it constitutes an efficient and effective strategy that guides the emergence of learnable patterns (Baayen & Ramscar 2015). As discussed in the other articles in this volume, the second part/whole dimension of this approach emphasizes the way that discriminable internal structure determines the implicative relations that define patterns of words. Matthews (1991:204) emphasizes the relationship between these dimensions of morphological analysis in observing that the organization characteristic of paradigmatic systems morphological structure reflects patterns of part-whole relations within individual words.

> ...words are not merely wholes made up of parts, but are themselves construable as parts with respect to systems of forms in which they participate.

This is consistent with Kruszewski's (1883:68–69) cogent summary of his own views, as cited in Radwanska Williams (1993:85).

> ...we have shown that the acquisition and use of language would be impossible if language represented a mass of isolated words. Words are connected to each other directly: 1) through association by similarity and 2) through association by contiguity. From these arise nests or systems and series of words. The associations by similarity make possible creativity in language. Words which are connected to other words by weak ties of similarity or which are not similar to them at all, are very easily forgotten. They are subject to an unconscious process which makes them more similar to other words. Besides the direct or straight connections between words, there are also indirect or diagonal ones, by force of which our mind classifies words into the same divisions as the objects which correspond to them. The two laws of association play the same role in linguistics as they do in psychology. The correspondence of the world of words to the world of thoughts is the fundamental law of the development of language.

By hypothesis, the basis of the similarity between words need not be described in terms of their morphemic composition, but can be specified in terms of discriminability from one other words with respect to the patterning of any of their parts. Any potentially differentiating dimension of form (i.e., segmental or suprasegmental properties) may serve this function, with the words as wholes functioning as the bearers of distinctive meanings. The idea that the internal structure of words need not be morphemic is, according to Lass (1997:315) the result expected from the ordinary processes of language change, which (commonly) effectuate striking separations between individual bits of form and meanings. These basic notions can be illustrated by considering the following data from Hungarian, Mari and Dinka.

Understanding a complicated whole by reducing it into constitutive parts is a familiar and successful analytic strategy in the sciences. Within the field of linguistic morphology, Blevins (2016), following Hockett (1987), refers to this as DISASSEMBLY. For example, the complex Hungarian word in (1) can be disassembled into 4 parts, each bearing its own meaning.

(1) bátor-ság-om-ról
brave-ness-1SG.POSS-about
'about my bravery'

An analysis is straightforward: a simple concatenation of the disassembled parts produces a word that reflects both the form and the meaning of these parts. In the present instance, the meaning of the assembled whole can be construed as a composite of its meaningful pieces: an adjectival base meaning 'brave' is followed by a marker of nominalization and this is followed in turn by a possessive marker and a case marker. What this type of theory requires is, accordingly, an inventory of parts and a set of rules that arrange them into words. Morphemes, as the elements of the inventory, are bi-unique form-meaning pairings.

On this view, Hungarian morphology is taken as a model system for understanding cross-linguistic morphology, since it so clearly exemplifies the essential assumptions guiding morphemic theory. In fact, this basic analytic strategy appears so intuitive, given the instructive example of Hungarian, that it might seem commonsensical to extend it to languages in which the meanings of the parts and the composite meaning of the whole are less transparently related than they are in Hungarian. This is the issue Lounsbury (1953) raised when he wondered whether morphological theory should be predicated on theorizing a "fictive agglutinative analog". How much are all languages underlyingly like an idealized version of the structure that Hungarian exhibits on the surface? In this connection, a fundamental theoretical question that arises, for both Hungarian and other languages, concerns whether the instructions for reassembly of the pieces in terms of e.g. linear sequence or hierarchy, can adequately recapture the nature of Hungarian complex words and whether the participation of the parts in Hungarian, and in languages generally, can be characterized without diacritic instructions for reassembly posited in all such theories?

Hungarian provides a case where the pieces themselves can be construed as meaningful and where the cumulation of pieces exhaustively comprises the meaning of the word. In contrast, consideration of the first past tense realizations of *-em* conjugation verbs in Mari (Uralic) provides a simple demonstration of how the organization of elements can be as crucial as the elements themselves. This is illustrated with partial paradigms for the representative verb *kol* 'die'.

(2)

	1ST PAST AFFIRMATIVE	1ST PAST NEGATIVE
SG1	kolə̂-**š-ə̂m**	**š-ə̂m** kolə̂
SG2	kolə̂-**š-ə̂c**	**š-ə̂c** kolə̂
SG3	kolə̂-**š**	**ə̂š** kolə̂

Mari clearly contains segmentable elements, including markers associated with past tense and subject person/number. These properties are reflected in the tense values and person features of the glosses in (2). On the other hand, it is the location of these markers as a suffix or as an independent unit preceding the verbal stem that is associated with positive versus negative polarity of verbs, respectively.[2] The same pieces deployed in different morphological configurations convey different polarity values for verbs: they take on different functions in the word context in which they occur. So, it is the discriminable difference between words with the same markers in different positions that is associatedwith polarity values for e.g., 'die' in Mari.[3]

The marking of singular versus plural number contrasts in in the Agar dialect of Dinka (Eastern Nilotic) presents another, far more complex challenge of this kind. As analyzed in Andersen (2014), Dinka distinguishes the case and number of its mostly mono-syllabic and di-syllabic nouns in terms of word internal interactions among four parameters. These are: (i) vowel length, (ii) tone, (iii) voice quality of the vowel, and (iv) vowel quality alternation grade.[4] Andersen presents the following noun pairs to illustrate the evident independence of any specific collection of these parameters with determinate number values:

(3)

SG	PL	
dí̱t	djɛ̱̀ɛt	'bird'
kɔ̱̀ɔɔr	kà̱ar	**'elbow'**
rjɛ̱̀ɛm	rî̱m	'blood'
cjé̤ec	cí̤c	'bee'
lá̤j	là̤aj	'animal'
mà̱ac	mɛ̤̂ɛc	'fire'
dò̱m	dṳ̂um	'field'
tò̤oɲ	tô̤oɲ	'pot'
t̪wô̤ooɲ	t̪ṳ́uɲ	'ember'
ɰà̤am	ɰɔ̤̀ɔm	**'thigh'**

Describing one simple contrast – for example, the pairs of words for 'thigh' and 'elbow' – will suffice to clarify how singular and plural variants of different items display mirror image patterns of vowel length and tone. Whereas the singular for 'elbow' has triple length for its vowels and low tone on the first vowel, the same word internal pattern characterizes the plural of ‘thigh’. Similarly, the double length vowels and low tone for the singular of ‘thigh’ parallels the length and tone pattern exhibited by the plural of ‘elbow’.

An approach to word structure that focuses on the surface shapes of word pairs, rather than on the generation of individual words, is able to extract the hidden patterning in this profusion of forms. A crucial ingredient for understanding this system is the recognition that words are primary objects of morphological theory and that, consequently, contrasts between words disclose the patterned nature of morphological organization. Indeed, a fundamental distinction between “morpheme”-based and word-based approaches is the claim in the former approach that words are epiphenomenal and the hypothesis in the latter that they represent an important independent level of analysis. This basic view is foundational for the careful exploration of Georgian morphologyin Gurevich (2006:44):

> The meaning of the whole word licenses the exponents to be used, but there is no precondition that the meanings of the exponents have to combine to comprise the meaning of the whole. Compositionality may, indeed, emerge, but as a side product rather than a central principle, or perhaps as an effective learning strategy. The whole itself may contribute meaning to the meanings of the parts, or may override the meanings of the parts.

(Classes of) words are wholes that are distinguishable by means of the patterns associated with their pieces. From a word-based perspective, surface wordforms are best viewed as ‘recombinant gestalts’ (Bickel 1994) or configurations of recurrent partials (segmental or suprasegmental) that get distributed in principled ways among members of paradigms. This parallels what Oudeyer (2006:22) describes as the “systematic reuse” (we would suggest *systemic* reuse) of phonological distinctions:

> all languages have repertoires of gestures and combinations of gestures which are small in relation to the repertoires of syllables, and whose elements are systematically reused to make syllables.

The domain of morphology can, similarly, be seen as an instance of a complex adaptive system, which redeploys the same pieces in new ways for different purposes. From this perspective, it becomes clear how and why the analysis of morphology can benefit from methods developed in other fields which study such systems.

The adoption of the word as an independent and necessary unit of analysis in turn permits words to function as parts of paradigmatic systems or niches. In this respect, a consequence of permitting words to be contrasted with words is the possibility of discovering morphological organization in the systems of relations between words. This accords with Robins’ (1959:128) familiar observation:

> . . .words anchored, as it were, in the paradigms of which they form a part usually bear a consistent, relatively simple and statable grammatical function.

> The word is a more stable and solid focus of grammatical relations than the component morpheme by itself. Put another way, grammatical statements are abstractions, but they are more profitably abstracted from words as wholes than from individual morphemes.

In this connection Blevins (2016) comments on the consequences of focusing on pieces and their composition, while ignoring the pivotal role of the word in morphology. An exclusive focus on the reduction of the word into smaller pieces precludes a whole line of inquiry into morphology, namely, the examination of how surface patterns of words cohere into organized systems. In particular, the surface expressions of words do not simply motivate the need for operations that transpose covertly uniform representations into overtly variable wordforms. In addition, surface wordforms are highly informative units, exhibiting relations of similarity and difference that facilitate the organization of inflectional paradigms and morphological families.

More generally, the discriminative perspective adopted here makes clear the role that a morphemic approach plays in forcing a radical dissociation of syntagmatic and paradigmatic structure. On the assumption that ‘parts’ are individually meaningful, and ‘wholes’ merely sum the meanings of their parts, there is no straightforward way in which part-whole relations can be extended to words and paradigms (or other classes of forms). However, if parts serve a discriminative function, which allows them to identify larger structures, syntagmatic and paradigmatic structure will reflect implicational relations along two interlocking dimensions. The discriminability of a sub-meaningful morphotactic part will identify a larger, meaning-bearing wordform. Once discriminated, the wordform will in turn imply other forms in its paradigm as well as within the larger morphological system. So, the system consists of the aggregate types of discriminative strategies that serve to distinguish words from one another and the patterns of relations that arise when words are recognized as contrasting in this way.

Other articles in this volume focus on the emergent effects that result from the way that the discriminability of words supports implicative patterns of organization. In the remainder of this article we focus on how the notion of discriminativity (or discriminability) has been operationalized in learning theory and how it can be applied to motivate Greenberg’s Universal 34 concerning number encoding. Exploring the applicability of discriminativity in another domain of language seems to constitute a natural extension of Kruszewski’s observation that “similarity” operates both in language and in psychology.

3. Discriminative learning: A typological challenge

In many languages, the number of items in the referent of a noun phrase is obligatorily encoded by an inflectional feature on the head noun or other lexical category in a clause. For example, in English we have distinction between *the book*, which must refer to a single book, and *the books*, which must refer to more than one.

While many languages show the same singular vs. plural distinction that English does, this is not the only attested system. Another fairly common type of language distinguishes between the SINGULAR (exactly one), the DUAL (exactly two), and the PLURAL (more than two). For example, in Upper Sorbian (a Slavic language spoken in Germany), we find singulars like *hród* 'castle' and *dźěłam* '(I) work', duals like *hrodaj* 'two castles' and *dźěłamoj* '(we two) work', and plurals like *hrody* 'castles' and *dźěłamy* '(we) work' (Stone 1993, Corbett 2000). In Hmong Daw (a Hmong-Mien language of Laos and southern China), personal pronouns distinguish three persons and three grammatical numbers (Jaisser 1995:118):

	SING	*DUAL*	*PLURAL*
1ST	*kuv*	*wb*	*peb*
2ND	*koj*	*neb*	*nej*
3RD	*nws*	*nkawd*	*lawv*

Beyond the singular and dual, some languages even distinguish a TRIAL (= exactly three) number. This four-way distinction is found, for example, in the subject agreement prefixes for human referents in Larike, a Central Malayo-Polynesian language spoken on Ambon Island (Laidig 1993, Siewierska 2013):

	SING	*DUAL*	*TRIAL*	*PLURAL*
1.INCL	—	*itua-*	*itidi-*	*ite-*
1.EXCL	*au-*	*arua-*	*aridu-*	*ami-*
2	*a-/a-*	*irua-*	*iridu-*	*imi-*
3.HUM	*ma-/mei-*	*matua-*	*matidu-*	*mati-*
3.NONHUM	*i-*			*iri-*

While cross-linguistic typological studies have revealed that there is a certain amount of variation among the grammatical number systems of diverse languages, that variation has been found to fall within fairly strict limits. Greenberg's (1963) Universal 34 states that "No language has a trial number unless it has a dual. No language has a dual unless it has a plural." This Number Hierarchy can be expressed as a chain of implicational relations: Trial → Dual → Plural → Singular. Greenberg's typological studies over a small corpus of 34 languages encouraged him to postulate numerous universal generalizations, both as absolutes and as tendencies. In succeeding years many of his preliminary claims have had to be modified or abandoned. The Number Hierarchy, however, has proved to be remarkably robust.[5]

Many of the unattested language types in this domain are silly: no language distinguishes between a prime vs. a composite number of referents, or has a special suffix indicating exactly 47 items. The non-existence of such languages does not demand an explanation. However, some of the language types that we do not find seem a priori more likely. For example, while the two-way SING/NOT-SING (i.e., plural) distinction is very common, the superficially similar DUAL/NOT-DUAL is essentially unattested. Along the same lines, while dual marking is fairly common, trial is rare, quadral (= exactly four) is only marginally attested, and markers for specific numbers greater than four are never found (at least in spoken rather than signed languages). It is less obvious why no known languages use these conceptually plausible number marking systems.

3.1. Universals in linguistic theory

One of the fundamental goals of linguistic theory is to offer explanations for why certain patterns recur cross-linguistically and others do not. Debate around the status of and explanations for language universals, recently highlighted by Evans and Levinson's (2009) provocative argument against innatist language-specific representational bases for universals, has been focused on two alternative models.

One hypothesis regarding the source of implicational universals like Greenberg's Universal 34 is that they follow from universal and innate properties of the human language faculty. These properties are specific to language and do not follow from any more general properties of human cognition or culture. By this view, universals like the Number Hierarchy reveal something about the underlying structure of human language and constraints on possible surface organization (Silverstein 1976, Noyer 1992, Harbour2011,2014). Specifically, under this view Universal Grammar provides a range of possible mental representations for grammatical number systems which are able to capture attested systems but not the unattested ones.

As an alternative, one could attempt to derive universals like the Number Hierarchy from language usage or general complexity properties (e.g., Croft & Poole 2008, Miestamo 2009). This approach has proven to be quite fruitful for a range of putative universals. But so far, the Number Hierarchy has resisted language-external explanation. For example, it is not immediately obvious how a Singular-Plural system is any simpler than a Dual-NotDual system, yet the first is the most common type and the second is (virtually) unattested (but see Evans 2012). In fact, Seuren (2013) offers the Number Hierarchy as a best example of a universal which is very unlikely to submit to an explanation that does not depend on language-specific properties of Universal Grammar. This is not an argument against grounding the Number Hierarchy in properties of general cognition, merely an observation that no convincing explanation along those lines has yet been proposed. In what follows we develop a learning model sensitive to realistic recurrences of set size and plausible assumptions about the representation of numerosities.

3.2. Model

Children learn the concept of numerosity before they learn morphosyntactic expressions of number (Barner et al. 2007, Wood et al. 2009, Clark & Nikitina 2009) and well before they master number name meanings Slusser & Sarnecka 2011). Learning the distinction between singular and plural sets does not appear to be dependent on morphological marking (Li et al. 2009).

Clark & Nikitina (2009) consider the use of *two* as a general purpose plural marker by English-speaking two- and three-year-old children. They argue that this arises at least in part because *two* is the numeral larger than one that is used most often by adults in child-directed speech. Furthermore, their observations indicate that the trajectories of individual children in learning grammatical number can vary quite a bit and are sensitive to the frequency of forms in adult speech. This frequency-sensitivity suggests that a general associative-learning process is playing an important role in the development of grammatical number.

Ramscar et al. (2011) developed a model of children's acquisition of number names (*one, two, three*, etc.) based on the Rescorla-Wagner theory of associative learning, which as a side effect predicts the subitizing limit (Kaufman et al. 1949), a constraint on the human ability to recognize the number of items in a set without explicit counting.

Rescorla and Wagner's (1972) learning model, rooted in Pavlovian learning theory, seeks to explain the way that associative learning gradually builds connections between perceptual cues and specific outcomes over the course of many learning trials. Early work in this direction (Hull 1943, Bush & Mosteller 1955) showed that associative learning follows a negatively accelerated learning curve which can be described using a difference equation:

$$\Delta V_A^O = \alpha\left(\lambda - V_A^O\right)$$

ΔV_A^O is the change in the strength of the association between cue A and the outcome O after a learning trial in which A and O occur together, V_A^O is the previous weight of the association prior to the current trial, α is the learning rate, and λ is the maximum conditioned response. Essentially, each learning trial moves us $\alpha\%$ of the way to λ. Extinction of associations work the same way. After each learning trial in which A does not occur with O, the change in the association strength is given by:

$$\Delta V_A^O = \alpha\left(0 - V_A^O\right)$$

R-W's model extends this approach in a way that can deal with compound cues AX. When the cues A and X both occur with O on a learning trial, the update for the association weights is given by:

$$V_{AX}^O = V_A^O + V_X^O$$

$$\Delta V_A^O = \alpha_A\beta\left(\lambda - V_{AX}^O\right)$$

$$\Delta V_X^O = \alpha_X\beta\left(\lambda - V_{AX}^O\right)$$

ΔV_A^O is the change in the strength of the association between cue A and the outcome O after a learning trial in which A and O occur together, V_A^O is the previous weight of the association prior to the current trial, and λ is the maximum conditioned response. The learning rate depends on the salience of the cue (α_A and α_X) and the salience of

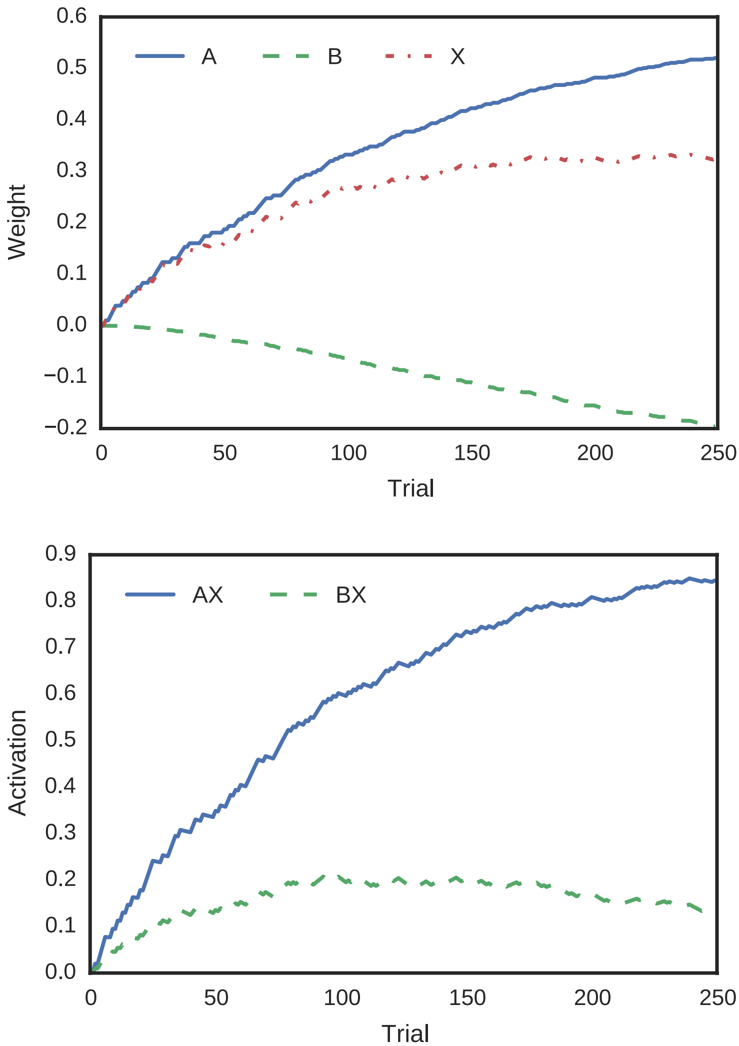

Figure 1. Discriminative learning: when the (hypothetical) compound cue AX occurs with O but the cue BX does not, the model learns that A and not X is a predictor of O.

O (β), and the learning step size depends on the previous association weight of the compound V^O_{AX}.

One consequence of this model is that a cue which always occurs when O does may not end up with a strong association if some other cue is a better predictor of the outcome. Suppose the compound cue AX consistently occurs with O, while the cue BX consistently does not. We expect the sum V^O_{AX} to approach λ and V^O_{BX} to approach 0, and as shown in Figure 1 that is what happens.

Looking at the total strength of the prediction in the lower graph, V_{BX}^{O} initially grows with V_{AX}^{O}, then begins to fall towards 0. Looking at the individual cue weights in the upper graph, we see that initially V_{A}^{O} and V_{X}^{O} are roughly equal and that X is competing with A and B to explain the (non-)occurrence of O. After about 50 trials, though, A wins the competition, and the learner has been able to discriminate among the cues that are present with O to find the ones which are the best predictors (Ramscar et al. 2010, Baayen et al. 2011). As R-W observe:

> "Organisms only learn when events violate their expectations. Certain expectations are built up about the events following a stimulus complex; expectations initiated by the complex and its component stimuli are then only modified when consequent events disagree with the composite expectation." (Rescorla & Wagner 1972:75)

In this model, learning is driven by prediction error: V_{AX}^{O} is the learner's expectation prior to a trial, and $(\lambda - V_{AX}^{O})$ or $(0 - V_{AX}^{O})$ is how much that expectation is violated on a positive or negative learning trial.

While it is not without problems, the R-W model has been enormously influential in the development of animal associative learning models. R-W learning has also been successfully applied to model human causal reasoning (Lober & Shanks 2000, Danks 2007), is closely related to both connectionist (Gluck & Bower 1988, Shanks 1991, Van Overwalle & Van Rooy 1998) and information theoretic Gallistel (2002) models, and is formally equivalent to the perceptron (Dawson 2008). In the domain of language, R-W learning has been applied to a range of problems, including modeling acquisition of plural forms (Ramscar & Yarlett 2007, Ramscar et al. 2013) and number names (Ramscar et al. 2011), word recognition (Baayen et al. 2011).

3.3. Input model

In order to construct a simulation of grammatical number acquisition based on R-W learning theory, we first need a representation of the input cues that a learner would be exposed to. The ability to subitize, or recognize the numerosity of small sets, is either innate or developed very early and can be taken as a semantic primitive (Wynn 1992, Piantadosi et al. 2012). Following Ramscar et al. (2011), we include numerosity of the set and all subsets as cues:

$$1 = [1]$$
$$2 = [1, 2]$$
$$3 = [1, 2, 3] \ldots$$

For example, a learner who encounters a set of one item is only exposed to the cue '1'—a single-item set.[6] A learner who encounters a set of two items is exposed to the cue '2'—a double-item set—and also to the cue '1', since encountering a double-item set entails encountering a single-item (sub)set. A learner who encounters a set of three

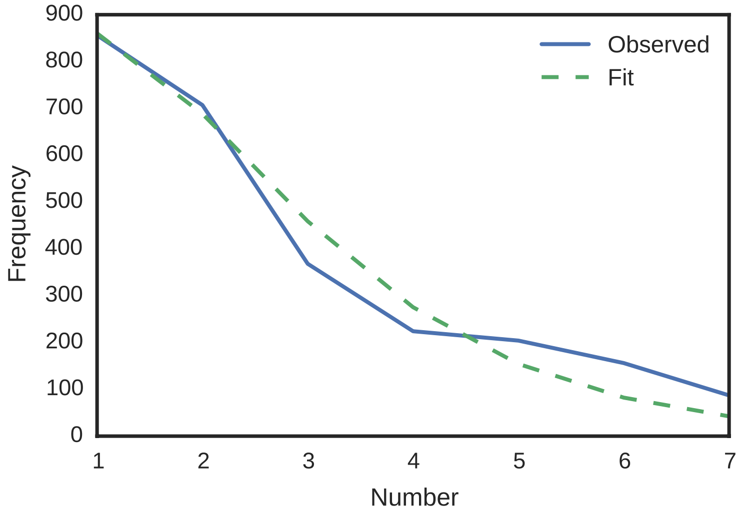

Figure 2. Frequency of number mentions in COCA (Davies 2008–, Ramscar et al. 2011) and fitted values.

items is exposed to triple-item, double-item, and single-item sets, etc. The learner's task is to determine which sets (the cues) best predict the use of each grammatical number marker (the outcomes).

R-W learning is also potentially very sensitive to statistical properties of the input. Thus, we also need an accurate probabilistic model of the frequencies at which learners would be exposed to various inputs. Specifically, we need to know how often children encounter talk about sets of a given numerosity while learning grammatical number.

In their simulation of number name acquisition, Ramscar et al. (2011) report counts of number names from 1 to 7 used as prenominal modifiers in the Corpus of Contemporary American English (COCA) and the Corpus del Español. The distribution of English counts is shown in Figure 2.

Using the VGAM package in R (Yee 2010), we found that a zero-truncated negative binomial distribution (a mixture of Poissons with no zero counts and gamma-distributed rates to account for overdispersion) provided a good fit to the counts in both languages. Therefore, we generated random inputs for our simulations by sampling from a zero-truncated negative binomial fit to the English counts (size = 3, prob = 0.6).

3.4. Results

For the first set of simulations, we ran 250 iterations of R-W learning with randomly generated inputs as the cues and the correct grammatical number for each input as the outcome. The trajectories for singular and plural outcomes are shown in Figure 3.

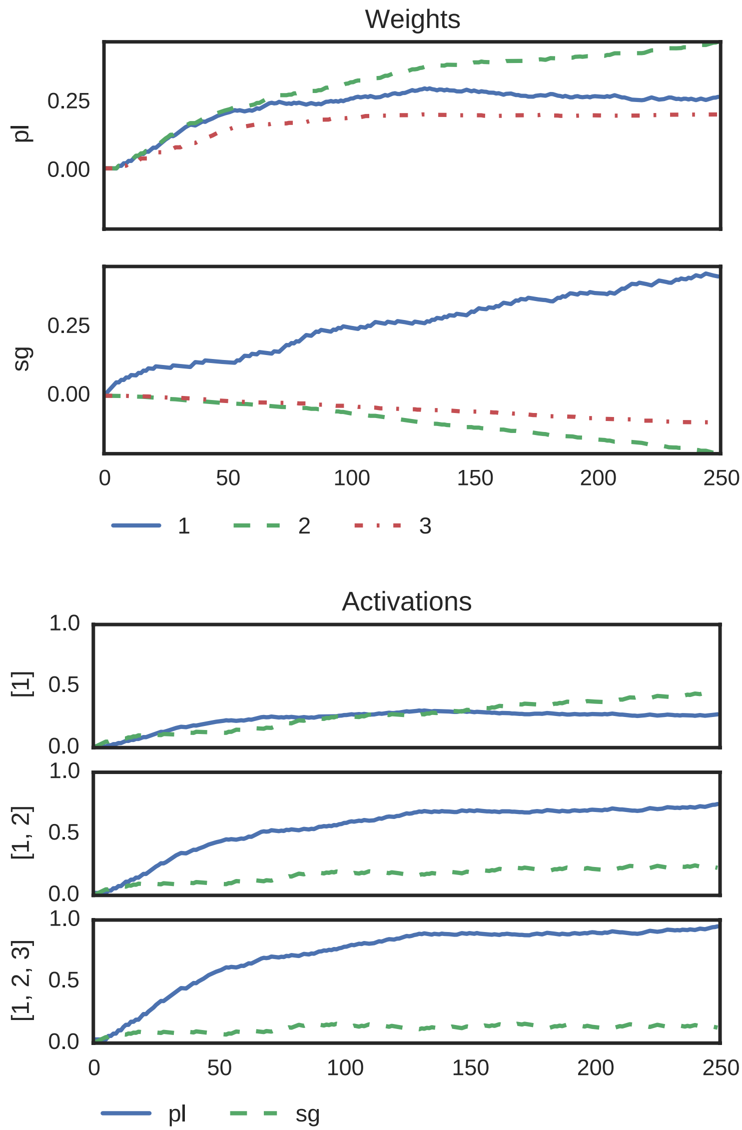

Figure 3. Simulated learning of a SINGULAR/PLURAL distinction.

The WEIGHTS are the strength of the association between individual cues 1, 2, or 3 and outcomes SG or PL: using the notation introduced above, V_1^{SG}, V_2^{SG}, V_3^{SG}, V_1^{PL}, V_2^{PL}, V_3^{PL}. The ASSOCIATIONS in the lower graph are the sum of the weights of the cues that are present in each context. For example, the bottom panel shows the change in $V_1^{SG} + V_2^{SG} + V_3^{SG}$ versus $V_1^{PL} + V_2^{PL} + V_3^{PL}$ as learning progresses. The prediction generated by the model for a context is the outcome with the largest total activation given the cues that are present.

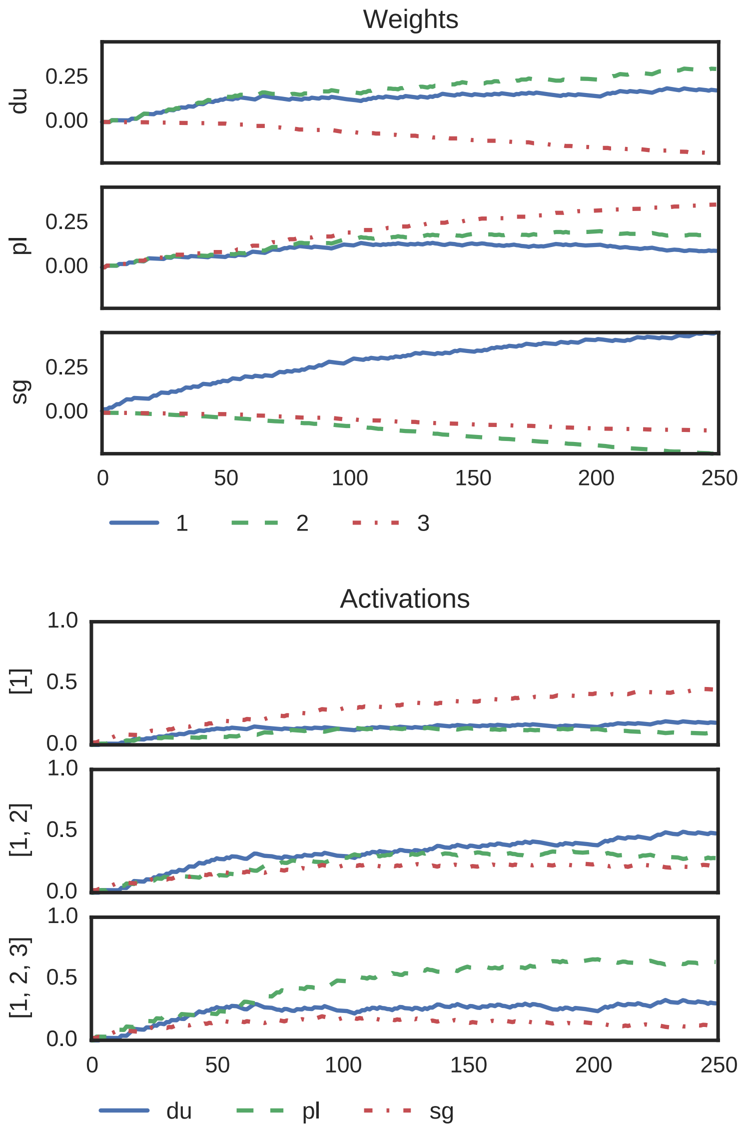

Figure 4. Simulated learning of SINGULAR/DUAL/PLURAL.

For this experiment, we see that learning what the singular means is easy: 1 is the only cue that has a positive association with the outcome SG. Plural is a bit more difficult, though: a single-item subset is always present in sets of higher numerosity. The learner therefore always encounters a cue of 1 when they encounter any set and any number-marker outcome. But, after about 50 iterations, the model has succeeded at discriminating between 2 and 1 and has identified the former as a good predictor of the outcome PL. The total activations in the lower graph show that the model is correctly distinguishing singular and plural sets after about 150 learning trials.

We repeated the simulation for a singular/dual/plural system, in Figure 4.

Learning proceeds in much the same way as in the previous case, though in this system there are two discrimination tasks to be solved: the learner must learn that the cue 1 is not associated with DU nor is it associated with PL. Additionally, it must learn that 2 is not associated with PL, even though all plural sets include 2 as a member. As before, the learner is able to correctly label singular, dual, and plural sets after about 150 iterations.

The next simulation considers the task of learning a dual/non-dual distinction, a grammatical number system type that is not found among human languages. The results are in Figure 5.

Parallel to the singular/plural system, in this system the model needs to learn that 2 is a good cue for the outcome DU but 1 is not, even though both cue 1 and cue 2 are always present with outcome DU. Unlike the singular/plural system, however, the model must also cope with the fact that 2 sometimes occurs with NONDU, for tripleton and higher sets. Furthermore, the model needs to learn that 3 is a good cue for NONDU. This is challenging, because 1 quickly becomes a strong cue for NONDU in this system, and 3 never occurs without 1.

After 250 iterations, the model has only begun to discriminate 1 and 2 (in the topmost graph), and the activation for DU never rises above NONDU for sets of numerosity 2. Note that this is not to say that the dual/non-dual distinction is unlearnable in general, only that it was not learned in 250 iterations. This suggests that some measure of relative degree of learnability is a determinant for the observed distributions of number marking.

The results in Figures 3–5 are traces of individual simulations. The actual learning trajectory and the final weights in the run depend on the particular randomly-generated training examples that are provided in each learning trial. To see how learning progresses in a population, we performed another set of simulations. In these experiments, we combined 100 learners, each following its own trajectory. After each learning trial, we calculated the fraction of the population which had mastered the appropriate grammatical number system: e.g, for the singular/dual/plural system, a 'correct' learner would have to assign the highest total activation to SG for a singleton set, DU for a set with two items, and PL for higher numerosities. As shown in Figure 6, the singular/plural and singular/dual/plural systems are learned quickly and reliably by all members of the population.

The same is not true for the dual/non-dual system. Some members of the population learn it quickly, but after 250 iterations less than 40% of the population can use the system correctly. The dual/non-dual system offers a much more difficult discrimination task to the learner than do either the singular/plural or the singular/dual/plural systems.

To test what features of the system lead to this contrast in learning difficulty, we performed two more sets of population-based simulations. In one, we

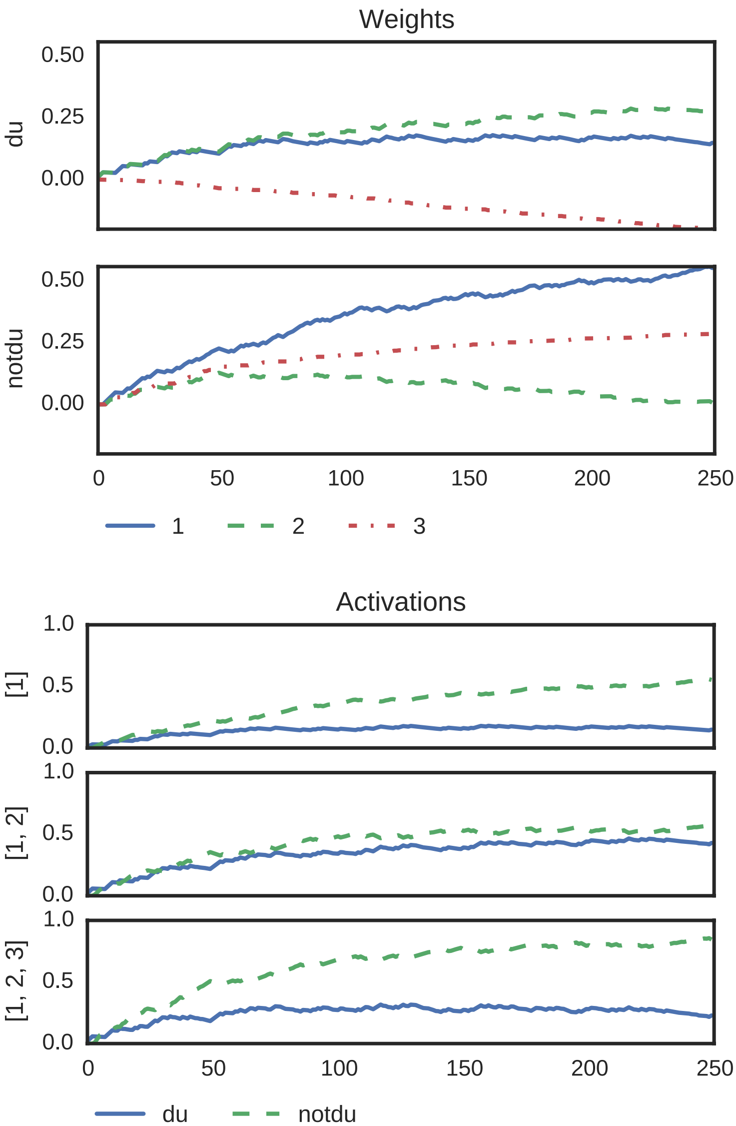

Figure 5. Simulated learning of a NON-DUAL/DUAL distinction.

removed subset cues from the inputs, including only the numerosity of the set as a whole:

$$1 = [1]$$
$$2 = [2]$$
$$3 = [3] \ldots$$

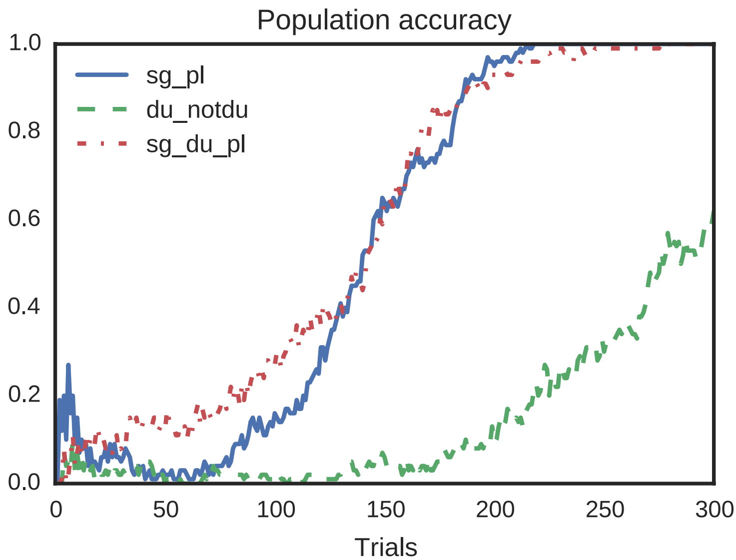

Figure 6. Fraction of a simulated population of learners that have mastered each system after *k* trials.

Given this input representation, the singular/plural distinction is learned very quickly, but both the singular/dual/plural and the dual/non-dual systems are learned very slowly (and at similar rates).

In the second set of simulations, we altered the input probability distribution to make sets of numerosity two more frequent than singletons (in contradiction to the real observed frequencies in Figure 2). In this case, all three systems are learned easily and quickly, and there is no distinction between the attested and unattested systems.

3.5. Discussion

In general, these results suggest that the implicational organization of number distinctions identified in Universal 34 may emerge from the interaction of a reliably common and skewed distribution of set sizes with particular assumptions concerning the representation of numerosities. The simulations replicate the expectations based on attested grammatical marking strategies: SG versus PL is learned both early and easily, while further distinguishing the DU takes longer, but is attainable. In contrast, the differentiation of DU from NON-DU proves a difficult discrimination to make. Given the success of the simulations in achieving empirically attested patterns without positing language specific representations, it was crucial to explore the contribution of two basic assumptions of the model, namely, our representation of numerosity and our assumption concerning the role of exposure to the frequencies of particular numerosities. When we expunged subset cues from the representations, singular/dual/plural patterned with dual/non-dual. This suggests the facilitating role of subset organization for the emergence of singular/plural and singular/dual/plural versus dual/non-dual. When we altered the frequencies for experienced numerosities so that duals occurred more often than singulars, the difference between all three patterns of number organization, i.e., singular/plural, singular/dual/plural, and dual/non-dual disappeared: all were learned equivalently well. This suggests that, as hypothesized, real world experience has an important shaping influence on the organization of grammatical number systems.

While our model explains aspects of the intriguing restrictions on the morphosyntax of cross-linguistic number marking, it only defines the broad constraints on number organization in natural language. It ignores the specific strategies that particular languages employ within the space of options permitted by Universal 34. Further research is required to understand the dynamics which motivate, sustain and alter the system of discriminations observed in specific languages. It is in this interdependence between general constraints on the organization of linguistic phenomena and the particularities of their encodings that an understanding of natural language is most likely to be found.

4. Conclusions

This article has focused on the notion that the patterns facilitated by discriminability constitute an important factor for the explanation of grammar organization in two domains. The first is the central domain of this volume, namely, Word and Pattern perspectives on morphology. An essential observation on this view is the analytic utility of words in understanding morphological organization. Central to this hypothesis is the claim that words are viewed internally as composed of parts that serve to discriminate them from one another: the words themselves often being the smallest units of analysis. From an external perspective, words are organized in ways that permit the development of implicative patterns facilitated by the discriminable structures of the words that comprise these patterns. Thus, the notion of discriminability is something that requires a better understanding in the morphological literature. It is, essentially, a construct that is more familiar from decades of experiments in learning theory, where it has been instructively employed in numerous behavioral domains, than it is in linguistics, where its value has been obscured by the analytic assumptions of morpheme-oriented approaches to morphology. We have accordingly proposed a novel deployment in the solution of a typological problem. This provides a useful illustration of how the basic idea of discriminability provides a resource for word structure and patterned word organization which has well-developed analogues in other domains. In particular, we presented a view of grammatical number learning based on associative learning (Rescorla & Wagner 1972, Ramscar & Yarlett 2007, Baayen et al. 2011). We showed that this model provides a possible language-external explanation for aspects of the Number Hierarchy. In particular, number as a real-world category is inherently structured in two ways: sets of higher numerosity are mentioned less frequently than sets of lower numerosity, and sets of higher numerosity always contain sets of lower numerosity. Using corpus-based computational simulations (Ramscar et al. 2011), we demonstrate that these facts, in interaction with general principles of probabilistic learning, plausibly lead to languages which violate the Number Hierarchy being much more difficult to learn than languages which follow it, which in turn motivates the emergence of the Number Hierarchy as an implicational universal. This shows that even fairly abstract properties of grammatical systems, when viewed from a developmental perspective, can be seen to have a physical or cognitive origin – and explanation – external to language. This presents morphology with a challenging and exciting task, specifically, to explore and extend this type of approach directly to morphological phenomena, as is presently being pursued in Seyfarth & Myslín (2014) and Arnon & Ramscar (2012), Ramscar & Dye (2010), Ramscar et al. (2013) and Dye et al. (2016).

Acknowledgments

Our analysis in terms of a discriminative learning model for explaining Universal 34 was originally developed by Ackerman and Malouf and Scott Seyfarth and appears here in essentially the same form it appeared in the Proceedings of the 37th Annual Meeting of the Cognitive Science Society, Pasadena California. The proposal is contextualized within the implicative approach articulated in Ackerman & Malouf (2016) with permission from Cambridge University Press. We thank David Barner, Michael Ramscar, and the participants at the 2014 workshop on "Corpus resources for quantitative and psycholinguistic analysis" held at Eszterházy College, Eger, Hungary for helpful criticisms.

NOTES

1. We use here the term Word and Pattern rather than the more familiar and traditional term Word and Paradigm, following J.P. Blevins' (2016) cogent arguments that the former better reflects the general nature of morphological organization: it is less relevant that words participate in conventionally construed full paradigms than that they participate in patterns of implicative association at any size. See the contribution in this volume by Bonami and Beniamine for further discussion.
2. Following Ackerman & Stump (2004), we assume that the periphrastic expressions for negative polarity fill cells in the morphological paradigm of the lexeme 'die'. See Bonami (2015) for a formal treatment of the relation between multiword morphological expressions and their syntactic realization.
3. This is reminiscent of the general contrast between the discriminative function of prefixes and suffixes discussed in Ramscar (2013).
4. See Baerman (2012) for a detailed analysis of similar data in Nuer that recognizes the importance of the ideas presented in this chapter for the organization of the Nuer system.
5. It is not quite exceptionless, though, and the actual typological situation is somewhat more complicated (Croft 1990, Corbett 2000, Evans 2012).

6. The learner is also exposed to the cues associated with the item's other properties such as shape and size, which are ignored here for simplicity.

References

Ackerman, F. & Malouf, R. 2016. Implicative relations on word-based morphological systems. In Hippisley, A. & Stump, G. (eds.), *The Cambridge handbook of morphology*, Cambridge: Cambridge University Press. 297–328.

Ackerman, F. & Stump, G. 2004. Paradigms and periphrasis: A study in realization-based lexicalism. In Sadler, L. & Spencer, A. (eds.), *Projecting morphology*, Stanford: CSLI Publications. 111–157.

Andersen, T. 2014. Number in Dinka. In Storch, A. & Dimmendaal, G. J. (eds.), *Number - Constructions and semantics: Case studies from Africa, Amazonia, India and Oceania*, Amsterdam: Benjamins. vol. 151 of *Studies in Language Companion Series*, 221–264.

Arnon, I. & Ramscar, M. 2012. Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*. 122: 292–305.

Baayen, R. H., McQueen, J. M., Dijkstra, T. & Schreuder, R. 2003. Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R. H. & Schreuder, R. (eds.), *Morphological structure in language processing*, Berlin: Mouton de Gruyter. 355–370.

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: 438–482.

Baayen, R. H. & Ramscar, M. 2015. Abstraction, storage and naive discriminative learning. In Dabrowska, E. & Divak, D. (eds.), *Handbook of cognitive linguistics*, Berlin: De Gruyter Mouton. 99–120.

Baerman, M. 2012. Paradigmatic chaos in Nuer. *Language* 88: 467–494.

Barner, D., Thalwitz, D., Wood, J. & Carey, S. 2007. On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science* 10: 365–373.

Bickel, B. 1994. In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language* 19: 73–127.

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

Bonami, O. 2015. Periphrasis as collocation. *Morphology* 25: 63–110.

Bush, R. R. & Mosteller, F. 1955. *Stochastic models of learning*. New York: John Wiley and Son.

Clark, E. V. & Nikitina, T. V. 2009. One vs. more than one: Antecedents to plural marking in early language acquisition. *Linguistics* 47: 103–139.

Corbett, G. G. 2000. *Number*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Croft, W. 1990. *Typology and universals.* Cambridge: Cambridge University Press.

Croft, W. & Poole, K. T. 2008. Inferring universals from grammatical variation: multidimensional scaling for typological analysis. *Theoretical Linguistics* 34: 1–37.

Danks, D. 2007. Causal learning from observations and manipulations. In *Thinking with data*, New York: Lawrence Erlbaum Associates. 359–388.

Davies, M. 2008–. The corpus of contemporary American English: 450 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

Dawson, M. R. W. 2008. Connectionism and classical conditioning. *Comparative Cognition and Behavior Reviews* 3: 1–115.

Dye, M., Milin, P., Futrell, R. & Ramscar, M. 2016. A functional theory of gender paradigms. In Kiefer, F., Blevins, J. P. & Bartos, H. (eds.), *Perspectives on morphological structure: Data and analyses*, Leiden: Brill. (in press).

Evans, N. 2012. Even more diverse than we had thought: The multiplicity of Trans-Fly languages. In Evans, N. & Klamer, M. (eds.), *Melanesian languages on the edge of Asia: Challenges for the 21st century*, University of Hawai'i Press.

Evans, N. & Levinson, S. C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–492.

Gallistel, C. R. 2002. Frequency, contingency and the information processing theory of conditioning. In Sedlmeier, P. & Betsch, T. (eds.), *Etc. Frequency processing and cognition*, Oxford: Oxford University Press. 153–171.

Gluck, M. A. & Bower, G. H. 1988. From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General* 117: 227–247.

Greenberg, J. H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. H. ed., *Universals of grammar*, Cambridge: MIT Press. 73–113.

Gurevich, O. 2006. *Construction morphology: The Georgian case*. Ph.D. dissertation, University of California, Berkeley.

Harbour, D. 2011. Valence and atomic number. *Linguistic Inquiry* 42: 561–595.

Harbour, D. 2014. Paucity, abundance, and the theory of number. *Language* 90: 185–229.

Hockett, C. F. 1987. *Refurbishing our foundations*. Amsterdam: John Benjamins.

Hull, C. L. 1943. *Principles of behavior*. New York: Appleton-Century-Crofts.

Jaisser, A. 1995. *Hmong for beginners*. Berkeley: Center for South and Southeast Asia Studies, University of California at Berkeley.

Joos, M. (ed.) 1957. *Readings in linguistics I*. New York: American Council of Learned Societies.

Kaufman, E., Lord, M., Reese, T. & Volkmann, J. 1949. The discrimination of visual number. *American Journal of Psychology* 62: 498–525.

Kemps, J. J. K., Rachèl, Ernestus, M., Schreuder, R. & Baayen, R. H. 2005. Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition* 33: 430–446.

Kruszewski, M. 1883. Outline of linguistic science. In *Writings in general linguistics*, Amsterdam: Benjamins. 37–153. Edited with an introduction by E.F.K. Koerner 1995.

Laidig, W. D. 1993. Insights from Larike possessive constructions. *Oceanic Linguistics* 32: 312–351.

Lass, R. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Li, P., Ogura, T., Barner, D., Yang, S.-Y. & S. Carey 2009. Does the conceptual distinction between singular and plural sets depend on language? *Developmental Psychology* 45: 1644–1653.

Lober, K. & Shanks, D. R. 2000. Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review* 107: 195–212.

Lounsbury, F. 1953. *Oneida verb morphology*. Yale University Publications in Anthropology 48. New Haven: Yale University Press. Chapter 1 reprinted in Joos (1957), 379–385.

Matthews, P. H. 1991. *Morphology*. Cambridge: Cambridge Univesity Press.

Miestamo, M. 2009. Implicational hierarchies and grammatical complexity. In Sampson, G. B., Gil, D. & Trudgill, P. eds., *Language complexity as an evolving variable*, Oxford: Oxford University Press. 80–97.

Noyer, R. 1992. *Features, positions and affixes in autonomous morphological structure*. Ph.D. dissertation, Massachusetts Institute of Technology.

Oudeyer, P.-Y. 2006. *Self-organization in the evolution of speech*. Oxford: Oxford University Press.

Piantadosi, S., Tenenbaum, J. & Goodman, N. 2012. Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition* 123: 199–217.

Ramscar, M. 2013. Suffixing, prefixing, and the functional order of regularities in meaningful strings. *Psihologija* 46: 377–396.

Ramscar, M. & Dye, M. 2010. Learning language from the input: Why innate constraints can't explain noun compounding. *Cognitive Psychology* 62: 1–40.

Ramscar, M., Dye, M. & McCauley, S. 2013. Error and expectation in language learning: The curious absence of 'mouses' in adult speech. *Langauge* 84: 329–347.

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: e22501.

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: 927–960.

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: 909–957.

Rescorla, R. A. & Wagner, A. R. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. & Prokasy, W. (eds.), *Classical conditioning II*, Appleton-Century-Crofts. 62–99.

Robins, R. H. 1959. In defense of WP. *Transactions of the Philological Society* 116–144.

Seuren, P. A. M. 2013. *From Whorf to Montague: Explorations in the theory of language*. Oxford: Oxford University Press.

Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F. & Malouf, R. 2016. Acoustic differences in morphologically-distinct homophones. Submitted.

Seyfarth, S. & Myslín, M. 2014. Discriminative learning predicts human recognition of english blend sources. In *Proceedings of the 36th annual conference of the cognitive science society*. 2064–2069.

Shanks, D. R. 1991. Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17: 433–443.

Siewierska, A. 2013. Third person zero of verbal person marking, Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/103.

Silverstein, M. 1976. Hierarchy of features and ergativity. In Dixon, R. M. W. (ed.), *Grammatical categories in australian languages*, Canberra: Australian National University. 112–171.

Slusser, E. B. & Sarnecka, B. W. 2011. Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology* 110: 38–51.

Stone, G. 1993. Sorbian. In Comrie, B. & Corbett, G. G. (eds.), *The Slavonic languages*, London: Routledge. 593–685.

Van Overwalle, F. & Van Rooy, D. 1998. A connectionist approach to causal attribution. In Read, S. J. & Miller, L. C. (eds.), *Connectionist and PDP models of social reasoning and social behavior*, New York: Erlbaum.

Wood, J., Kouider, S. & Carey, S. 2009. Acquisition of singular-plural morphology. *Developmental Psychology* 45: 202–206.

Wynn, K. 1992. Children's acquisition of the number words and the counting system. *Cognitive Psychology* 24: 220–251.

Yee, T. W. 2010. The VGAM package for categorical data analysis. *Journal of Statistical Software* 32: 1–34. http://www.jstatsoft.org/v32/i10/.

Authors' addresses: *(Farrell Ackerman)*
Department of Linguistics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0108
USA
Email: fackerman@ucsd.edu

(Robert Malouf)
Department of Linguistics and Asian/Middle Eastern Languages
San Diego State University
5500 Campanile Drive
San Diego, CA 92182-7727
USA
Email: rmalouf@mail.sdsu.edu

(James P. Blevins)
Homerton College
University of Cambridge
Hills Road
Cambridge CB2 8PH
UK
Email: jpb39@cam.ac.uk

Your short guide to the EUP Journals Blog http://euppublishingblog.com/

A forum for discussions relating to Edinburgh University Press Journals

THE UNIVERSITY OF EDINBURGH
EDINBURGH University Press

1. The primary goal of the EUP Journals Blog

To aid discovery of authors, articles, research, multimedia and reviews published in Journals, and as a consequence contribute to increasing traffic, usage and citations of journal content.

2. Audience

Blog posts are written for an educated, popular and academic audience within EUP Journals' publishing fields.

3. Content criteria - your ideas for posts

We prioritize posts that will feature highly in search rankings, that are shareable and that will drive readers to your article on the EUP site.

4. Word count, style, and formatting

- Flexible length, however typical posts range 70-600 words.
- Related images and media files are encouraged.
- No heavy restrictions to the style or format of the post, but it should best reflect the content and topic discussed.

5. Linking policy

- Links to external blogs and websites that are related to the author, subject matter and to EUP publishing fields are encouraged, e.g.to related blog posts

6. Submit your post

Submit to ruth.allison@eup.ed.ac.uk

If you'd like to be a regular contributor, then we can set you up as an author so you can create, edit, publish, and delete your *own* posts, as well as upload files and images.

7. Republishing/repurposing

Posts may be re-used and re-purposed on other websites and blogs, but a minimum 2 week waiting period is suggested, and an acknowledgement and link to the original post on the EUP blog is requested.

8. Items to accompany post

- A short biography (ideally 25 words or less, but up to 40 words)
- A photo/headshot image of the author(s) if possible.
- Any relevant, thematic images or accompanying media (podcasts, video, graphics and photographs), provided copyright and permission to republish has been obtained.
- Files should be high resolution and a maximum of 1GB
- Permitted file types: *jpg, jpeg, png, gif, pdf, doc, ppt, odt, pptx, docx, pps, ppsx, xls, xlsx, key, mp3, m4a, wav, ogg, zip, ogv, mp4, m4v, mov, wmv, avi, mpg, 3gp, 3g2*.