

UC Berkeley

Proceedings of the Annual Meeting of the Berkeley Linguistics Society

Title

The No Blur Principle Effects as an Emergent Property of Language Systems

Permalink

<https://escholarship.org/uc/item/8xk6t5wq>

Journal

Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 41(41)

ISSN

2377-1666

Authors

Ackerman, Farrell

Malouf, Robert

Publication Date

2015

DOI

10.20354/B4414110014

Peer reviewed

PROCEEDINGS OF THE FORTY-FIRST ANNUAL MEETING OF THE
BERKELEY LINGUISTICS SOCIETY

February 7-8, 2015

General Session

Special Session
Fieldwork Methodology

Editors

Anna E. Jurgensen
Hannah Sande
Spencer Lamoureux
Kenny Baclawski
Alison Zerbe

Berkeley Linguistics Society
Berkeley, CA, USA

Berkeley Linguistics Society
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA

All papers copyright © 2015 by the Berkeley Linguistics Society, Inc.

All rights reserved.

ISSN: 0363-2946

LCCN: 76-640143

Contents

Acknowledgments	v
Foreword	vii
<i>The No Blur Principle Effects as an Emergent Property of Language Systems</i>	
Farrell Ackerman, Robert Malouf.	1
<i>Intensification and sociolinguistic variation: a corpus study</i>	
Andrea Beltrama	15
<i>Tagalog Sluicing Revisited</i>	
Lena Borise	31
<i>Phonological Opacity in Pendau: a Local Constraint Conjunction Analysis</i>	
Yan Chen	49
<i>Proximal Demonstratives in Predicate NPs</i>	
Ryan B. Doran, Gregory Ward	61
<i>Syntax of generic null objects revisited</i>	
Vera Dvořák	71
<i>Non-canonical Noun Incorporation in Bzhedug Adyghe</i>	
Ksenia Ershova	99
<i>Perceptual distribution of merging phonemes</i>	
Valerie Freeman	121
<i>Second Position and “Floating” Clitics in Wakhi</i>	
Zuzanna Fuchs.	133
<i>Some causative alternations in K’iche’, and a unified syntactic derivation</i>	
John Gluckman	155
<i>The ‘Whole’ Story of Partitive Quantification</i>	
Kristen A. Greer	175
<i>A Field Method to Describe Spontaneous Motion Events in Japanese</i>	
Miyuki Ishibashi	197

<i>On the Derivation of Relative Clauses in Teotitlán del Valle Zapotec</i> Nick Kalivoda, Erik Zyman	219
<i>Gradability and Mimetic Verbs in Japanese: A Frame-Semantic Account</i> Naoki Kiyama, Kimi Akita.	245
<i>Exhaustivity, Predication and the Semantics of Movement</i> Peter Klecha, Martina Martinović.	267
<i>Reevaluating the Diphthong Mergers in Japono-Ryukyuan</i> Tyler Lau	287
<i>Pluractionality and the stative vs. eventive contrast in Ranmo</i> Jenny Lee	307
<i>Labial Harmonic Shift in Kazakh: Mapping the Pathways and Motivations for Decay</i> Adam G. McCollum	329
<i>Reference to situation content in Uyghur auxiliary ‘bolmaq’</i> Andrew McKenzie, Gülnar Eziz, Travis Major	353
<i>Case-Marking in Estonian Pseudopartitives</i> Mark Norris	371
<i>Discourse Coherence and Relativization in Korean</i> Sang-Hee Park.	397
<i>Negotiating Lexical Uncertainty and Speaker Expertise with Disjunction</i> Christopher Potts, Roger Levy	417
<i>Patterns of Misperception of Arabic Consonants</i> Chelsea Sanker.	447
<i>The Imperative Split and the Origin of Switch-Reference Markers in Nungon</i> Hannah Sarvasy	473
<i>Asymmetries in Long-Distance QR</i> Misako Tanaka.	493
<i>The cross-linguistic distribution of sign language parameters</i> Rachael Tatman	503

<i>Homophony and contrast neutralization in Southern Min tone sandhi circle</i>	
Tsz-Him Tsui 515
<i>Cultural Transmission of Self-Concept from Parent to Child in Chinese American Families</i>	
Aya Williams, Stephen Chen, Qing Zhou 533
<i>Fruits for Animals: Hunting Avoidance Speech Style in Murui</i>	
Katarzyna Izabela Wojtylak 545
<i>A Quest for Linguistic Authenticity: Cantonese and Putonghua in Postcolonial Hong Kong</i>	
Andrew D. Wong. 563

Acknowledgments

As the Executive Committee of the 41st Annual Meeting of the Berkeley Linguistics Society, we would like to express our gratitude to the conference participants, volunteers, session chairs, faculty, and staff members for their participation. It was your contributions that made the conference a success. We are especially grateful for the patience, hard work, and support of Paula Floro and Belén Flores, without whom BLS 41 would not have been possible. We would also like to thank the following departments and organizations of the University of California, Berkeley for their generous financial support:

Department of Linguistics
Student Opportunity Fund
Graduate Assembly
Department of Philosophy
Department of Psychology
Department of Ethnic Studies
French Department
Department of Spanish and Portuguese
Berkeley Language Center
Office of Academic Affairs Vice President

Foreword

This monograph contains a number of the talks given at the 41st Annual Meeting of the Berkeley Linguistics Society, held in Berkeley, California, February 7-8, 2015. The conference included a General Session and the Special Session *Fieldwork Methodology*. The 41st Annual Meeting was planned and run by the second-year graduate students of the Department of Linguistics at the University of California, Berkeley: Kenny Baclawski, Anna Jurgensen, Spencer Lamoureux, Hannah Sande, and Alison Zerbe.

The original submissions of the papers in this volume were reviewed for style by Anna Jurgensen and Hannah Sande. Resubmitted papers were edited as necessary by Anna Jurgensen and Kenny Baclawski, and then compiled into the final monograph by Anna Jurgensen. The final monograph was reviewed by Spencer Lamoureux. The endeavor was supported by Alison Zerbe's management of the Berkeley Linguistic Society's funds for publications.

The BLS 41 Executive Committee
July 2015

The No Blur Principle Effects as an Emergent Property of Language Systems

FARRELL ACKERMAN¹, ROBERT MALOUF²

¹*University of California, San Diego*; ²*San Diego State University*

1 Introduction

In languages with multiple inflection classes, the formal variation associated with lexically conditioned allomorphy typically shows only a loose correlation with systematic phonological or semantic conditions (synchronically, at least) and often seems to serve no apparent communicative function. All natural languages show a certain degree of what Baerman *et al.* (2010:2) call “gratuitous” morphological complexity and Wurzel (1986:76) describes as “ballast” in the linguistic system. Many linguists have proposed universal constraints that place limits on the degree of additional complexity that inflection class systems can bring to a language (e.g., Carstairs 1987, Carstairs-McCarthy 1991, Plank 1991, Müller 2007).

In this paper, we compare two such proposals, the No Blur Principle (Carstairs-McCarthy 1994) and the Low Conditional Entropy Conjecture (Ackerman & Malouf 2013). The LCEC, it is claimed, reflects an emergent strategy of cross-linguistic morphological organization that permits language to solve the Paradigm Cell Filling Problem of Ackerman *et al.* (2009) and Ackerman & Malouf (2013). The PCFP is this: given exposure to a small number of surface word forms, speakers of morphologically complex languages must learn to reliably predict the full inflectional (or derivational) paradigms of novel words. Ackerman *et al.* (2009) argue that within the general framework of Word & Paradigm Morphology word relations in morphologically complex languages are structured in such a way as to facilitate the correct inferences involved in solving the PCFP. In particular, they posit that conditional entropy, the average uncertainty about the realization of one cell in the paradigm given another, is the relevant measure of morphological complexity.¹ Many languages have large paradigms with many inflectional classes, but, according to the LCEC a common property among them is that their organization will dependably reflect an average paradigm entropy that is fairly low.

In this paper, we aim to show that the NBP is not an independent and sufficient constraint on possible inflectional organization. Rather, it is a special case of the LCEC. While languages which obey the NBP will have low entropy, there are other, non-NBP strategies for maintaining low conditional entropy and the learnability benefits that accrue from it. Furthermore, the same evolutionary pressure(s) on language from which organization displaying low conditional entropy emerges, will often (but not always) lead to languages which fit the NBP.

In section 2 we define the NBP and in section 3 we describe conditional entropy and the LCEC. In section 4 we offer a simulation experiment designed to test the predictions of the NBP and LCEC. Finally in section 5 we present our conclusions.

¹See Bonami & Beniamine (2015) for an analysis in terms of joint entropies for the frequent empirical situation where the words occupying multiple cells all provide information that reduces the uncertainty in predicting an unknown form of a known lexeme.

2 No Blur Principle

The **No Blur Principle** (Carstairs-McCarthy 1994, 2010, 2014) states that in languages with multiple inflection classes, the realization of each cell of the paradigm for each class must either uniquely identify the inflection class or be the default realization for that cell. For example, in the (hypothetical) paradigm in (1), every form uniquely identifies the inflection class:

(1)	CLASS	NOM.SG	ACC.SG	NOM.PL	ACC.PL
	I	<i>-am</i>	<i>-as</i>	<i>-an</i>	<i>-ag</i>
	II	<i>-om</i>	<i>-os</i>	<i>-on</i>	<i>-og</i>
	III	<i>-im</i>	<i>-is</i>	<i>-in</i>	<i>-ig</i>
	IV	<i>-um</i>	<i>-us</i>	<i>-un</i>	<i>-ug</i>

Once speakers see, for example, an accusative plural wordform for some lexeme in *-ag*, then they know that this lexeme must be a member of class I. In the following paradigm, in contrast, not all forms are similarly diagnostic:

(2)	CLASS	NOM.SG	ACC.SG	NOM.PL	ACC.PL
	I	<i>-am</i>	<i>-əs</i>	<i>-ən</i>	<i>-əg</i>
	II	<i>-om</i>	<i>-os</i>	<i>-on</i>	<i>-og</i>
	III	<i>-əm</i>	<i>-is</i>	<i>-ən</i>	<i>-əg</i>
	IV	<i>-əm</i>	<i>-əs</i>	<i>-un</i>	<i>-ug</i>

Here, forms in $-\partial C$ fail to uniquely identify the class: a lexeme whose nominative singular ends in $-\partial m$ could be in class III or in class IV. However, the $-\partial C$ forms can be denominated as a default class, used whenever a paradigm cell has no diagnostic realization that uniquely identifies class membership. This result can be factored out in the following way:

(3)	CLASS	NOM.SG	ACC.SG	NOM.PL	ACC.PL
	I	<i>-am</i>	—	—	—
	II	<i>-om</i>	<i>-os</i>	<i>-on</i>	—
	III	—	<i>-is</i>	—	—
	IV	—	—	<i>-un</i>	<i>-ug</i>
	def.	<i>-əm</i>	<i>-əs</i>	<i>-ən</i>	<i>-əg</i>

In (3), every form which is specific to some inflection class (i.e., every non-default form) is diagnostic: a nominative singular in $-\partial m$ is an uninformative default, but one in $-\partial m$ unambiguously signals class II.

Finally, consider one more system:

(4)	CLASS	NOM.SG	ACC.SG	NOM.PL	ACC.PL
	I	<i>-ɔm</i>	<i>-əs</i>	—	—
	II	<i>-ɔm</i>	<i>-ɔs</i>	<i>-on</i>	—
	III	<i>-əm</i>	<i>-ɔs</i>	—	—
	IV	<i>-əm</i>	<i>-əs</i>	<i>-un</i>	<i>-ug</i>
	def.	—	—	<i>-ən</i>	<i>-əg</i>

In this language, forms in *-om* and *-os* are neither defaults nor diagnostic of class membership. There are two nominative singular and accusative singular forms that exhaustively partition the paradigm, each associated with two of the four classes. Thus the nominative singular and accusative singular cells are ‘blurred’. This is the type of cross-linguistic paradigm organization that the NBP rules out.

In this connection, Carstairs-McCarthy (1994) makes two claims, namely that (i) the NBP represents a universal property of language; and (ii) it follows from a fundamental learning bias. The first claim is an empirical one: either all languages follow the NBP or they do not. Evaluating this claim however is not entirely straightforward. Beyond the evidential issues that arise in evaluating any putative universal on the basis of a limited sample of languages (Cysouw 2005, Piantadosi & Gibson 2014), the interpretation of the NBP for a given language depends on the analysis of the data that one assumes. In particular, Carstairs-McCarthy argues that affixal inflection and stem alternations form two distinct systems, and that the NBP applies only to affixal morphology. Since many inflection class systems involve interactions between affixes and stems, evaluating the NBP requires understanding the nature of each of these two types of exponence and the consequences of their interdependencies for identifying the principles guiding paradigm structure.² At the present time, it appears that several languages do seem to follow the NBP, though the sample of languages that have been analyzed this way is somewhat small and not particularly diverse. In contrast, however the principle is interpreted, there do seem to be recalcitrant counter-examples from a wide distribution of languages that cannot be assimilated to the NBP.

Carstairs-McCarthy’s (1994) second claim is that the origin of the NBP as a universal is motivated by Clark’s (1987, 1990, 1993) **Principle of Contrast** for lexical learning. Clark argues that this principle is one of two pragmatic principles that guide vocabulary acquisition: this principle states that “Speakers take every difference in form to mark a difference in meaning” (Clark 1993:64). This operates in conjunction with her Conventionality Principle according to which learners reuse the lexical item used by adults to denote a particular semantic distinction. As a consequence, when a learner hears an unfamiliar term employed to convey a semantic distinction with an already established conventional word, the Principle of Contrast leads the learner to assume that the new form has a different meaning, however subtle, from the known one. Carstairs-McCarthy (2010) generalizes the link between the Principle of Contrast and the NBP by introducing the notion of **vocabular clarity**. Carstairs-McCarthy observes that children acquiring language in a bilingual environment seem to have no significant problem keeping the vocabularies of the two languages distinct. Children are (he argues) able to switch between alternate vocabularies as long as there are contextual cues that indicate which vocabulary should be active in a particular environment. He further points out that many languages have alternative vocabularies to be used in different social contexts (what he calls multivocbulism). In Guugu Yimidhirr (Pama-Nyungan), for example, a man would use the word *gadiilbaga* ‘axe’ with a brother-in-law, *guliirra* ‘axe’ with someone of the opposite sex, and *warrbi* ‘axe’ in informal speech (Carstairs-McCarthy

²Vigorous debate concerning these issues continues, with some pointing out languages which seem to violate the NBP (Blevins 2000, Stump 2005, Halle & Marantz 2000, Baerman 2012, Stump & Finkel 2013, Baerman 2014a,b) while others argue that, when analyzed properly, apparent counterexamples can be seen to fall into line with the NBP’s predictions (Cameron-Faulkner & Carstairs-McCarthy 2000, Enger 2007, 2013, Carstairs-McCarthy 2010, 2014).

2010:94). What seem to be synonyms really are not, as the social context would determine in any particular situation which of the three words for ‘axe’ should be used.

Inflection class systems, by hypothesis, represent a particularly abstract form of multi-vocabulism. In the paradigm in (1), there are four different affixes that mean ‘nominative singular’, but the choice of which affix to use in any context is determined by the inflection class of the noun to be inflected. The paradigms in (1)–(3) all show vocabular clarity: each affix is associated with a unique context, either a specific inflection class or the default. The paradigm in (4) however does not. This nominative singular $-om$ and the accusative singular $-os$ are each associated with two inflection classes and therefore two ‘meanings’, in violation of the Principle of Contrast and the NBP.

This second claim, that the NBP is a consequence of the Principle of Contrast/vocabular clarity, is more difficult to evaluate than the first. In order for this claim to be true, two things must hold: acquisition of inflection class systems must be actually governed by the Principle of Contrast³ and, that must then lead to the specific distribution of languages we find. We will focus in this paper on the second requirement. Suppose vocabular clarity does obtain. Would that actually lead to the NBP?

It is appealing to account for the NBP as a consequence of the way that language is learned or used, but any such account is faced with Epstein’s (1999, 2006) **Generativist’s Question**: “How could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity?” Or:

“To explain a macroscopic regularity x is to furnish a suitable microspecification that suffices to generate it. The core request is hardly outlandish: To explain a macro- x , please show how it could arise in a plausible society. Demonstrate how a set of recognizable — heterogeneous, autonomous, boundedly rational, locally interacting — agents could actually get there in reasonable time.” (Epstein 2006:51)

In section 4, we will present a computational simulation that attempts to answer Epstein’s challenge. But first, we consider an alternative constraint on morphological complexity, the Low Entropy Conjecture.

3 Low entropy conjecture

For speakers of languages with complex paradigms and multiple inflection classes, it is unlikely that every speaker will encounter and memorize every form of every lexeme. For example, in Tundra Nenets (Samoyedic branch of Uralic) each noun lexeme has 210 possible inflected forms for the morphosyntactic feature property combinations for 7 cases, 3 numbers, and 3 persons and numbers for possessors: $(7 \times 3) + (7 \times 3 \times 3 \times 3) = 210$ distinct wordforms. So, if a speaker needs to produce a wordform that they have not previously learned, what strategies can they use to infer the correct form? This task receives a general formulation as the Paradigm Cell Filling Problem (PCFP) in Ackerman *et al.* (2009):

³We leave to another forum the (in)appropriateness of the analogy between lexical learning and inflectional learning that forms the basis of the NBP

Paradigm Cell Filling Problem: Given exposure to an inflected wordform of a novel lexeme, what licenses reliable inferences about the other wordforms in its inflectional family?

In pedagogical practice, following Word-and-Paradigm models of morphology, an inflectional system is represented via a set of full exemplary paradigms reflecting the inflectional classes of a languages. Speakers may memorize complete paradigms for frequent lexemes, but for infrequent lexemes speakers must produce novel wordforms by analogy from known wordforms representative of patterns of forms constitutive paradigms. Morphological systems are structured by patterns of implicational relations among all wordforms within paradigms, which provide speakers with a means for carrying out these predictions on the basis of incomplete information. This is a pervasive type of example in language of a far broader task that humans constantly confront, namely, how to arrive at reasonable solutions to complex problems under evidently uncertain conditions.

In order to assess the strength of implicational relations among wordforms, we will use the information-theoretic notion **entropy** as the measure of uncertainty or predictability (Ackerman *et al.* 2009, Ackerman & Malouf 2013). Suppose we are given a random variable X which can take on one of a set of alternative values x_1, x_2, \dots, x_n with corresponding probability $p(x_1), p(x_2), \dots, p(x_n)$. Then, the amount of uncertainty in X , or, alternatively, the degree of information conveyed on learning the value of X , is the entropy $H(X)$:

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i)$$

The entropy $H(X)$ is the weighted average of the **surprisal** $-\log_2 p(x_i)$ for each possible outcome x_i . The surprisal is a measure of the amount of information expressed by a particular outcome measured in bits, where 1 bit is the information content of a choice between two equally probable outcomes. Outcomes which are less probable (and therefore harder to predict) have higher surprisal. Specifically, surprisal is 0 bits for outcomes which always occur ($p(x) = 1$) and approaches ∞ for very unlikely events (as $p(x)$ approaches 0). The more choices there are in a given domain and the more evenly distributed the probability of each particular alternative, the greater the uncertainty or surprise there is (on average) that a particular choice among competitors will be made and, hence, the greater the entropy. Conversely, choices with only a few possible outcomes or with one or two highly probable outcomes among many unlikely exceptions have a low entropy.

To solve a particular instance of the PCFP, a speaker needs to generate an unknown form on the basis of (at least) one other known form. To quantify the predictability of one form given the other, we can measure the size of the surprise associated with these forms using **conditional entropy** $H(Y|X)$, the uncertainty in the value of Y given that we already know the value of X :

$$\begin{aligned} H(Y|X) &= H(X, Y) - H(X) \\ &= \sum_{x \in X} \sum_{y \in Y} P(x, y) \log_2 P(y|x) \end{aligned}$$

The smaller $H(Y|X)$ is, the more predictable Y is on the basis of X , i.e., the less surprised one is that Y is selected given knowledge of X . In the case where X completely determines

Y , the conditional entropy $H(Y|X)$ is 0 bits: given the value of X , there is no question remaining as to what the value of Y is. On the other hand, if X gives us no information about Y at all, the conditional entropy $H(Y|X)$ is equal to $H(Y)$: given the value of X , we are just as uncertain about the value of Y as we would be without knowing X at all.

For any paradigm, the difficulty of the PCFP is the average conditional entropy across all possible pairs of wordforms:

$$H(P) = \sum_{c_1} \sum_{c_2} H(c_1|c_2)$$

If the PCFP as formulated here accurately describes something about the way language is structured, then in order for speakers to successfully produce forms in a morphologically complex language, the average conditional entropy must be low. Ackerman & Malouf (2013) computed the average conditional entropy for paradigms in a small sample of paradigms taken from typologically and genetically diverse languages, finding that the average conditional entropy ranged from 0 bits to 0.75 bits. The sample included languages with impressively complex-looking morphological systems that have been raised as violations of the NBP, such as Nuer and Chiquihuitlán Mazatec (a language with at least 109 different verbal conjugations). Despite the large range in the apparent complexities of these languages, as measured by the number of paradigm cells, allomorphs, and inflection classes, the average conditional entropy fell within a narrow range. The **Low Entropy Conjecture** is the prediction that all languages will show low average conditional entropy along the lines discovered for these languages. Ackerman & Malouf (2013) refer to this measure as calculating the **integrative complexity** or I-complexity of a morphological system.

With this as background, the NBP can be construed as a particular strategy for achieving low conditional entropy. If all cells in a paradigm indicate class membership, as in (1), then the average conditional entropy is 0 bits: every form predicts every other form perfectly. Moving away from that ideal increases the entropy, but adding default realizations for cells only increases the entropy by a small amount. Class-unique forms are highly predictive but hard to predict, while default forms are easily predicted but not very predictive. Together they yield low average entropy. As paradigms become more ‘blurred’, the average conditional entropy goes up as well. A paradigm in which every cell is blurred:

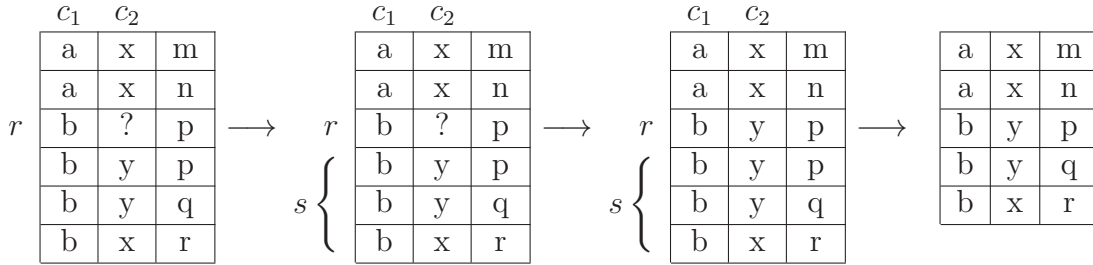
(5)	CLASS	NOM.SG	ACC.SG	NOM.PL	ACC.PL
	I	<i>-am</i>	<i>-as</i>	<i>-on</i>	<i>-ag</i>
	II	<i>-am</i>	<i>-os</i>	<i>-an</i>	<i>-og</i>
	III	<i>-om</i>	<i>-as</i>	<i>-on</i>	<i>-ag</i>
	IV	<i>-om</i>	<i>-os</i>	<i>-an</i>	<i>-og</i>

also has relatively high conditional entropy. No forms uniquely identify inflection class membership and no forms are any better represented than any others.

4 Simulation

Our hypothesis is that speakers’ need to solve the PCFP serves as a strong evolutionary pressure on language, guiding development of morphological systems in the direction predicted by the LCEC. As a side effect, many languages will also reflect the NBP. To test this, we constructed a simple agent-based computational simulation, adapting the **iterated learning**

Figure 1: Predicting cell c_2 from cell c_1 for class r by analogy with s



model (Esper 1966, Kirby 2002, Kirby & Hurford 2002). In this framework, an artificial language is constructed and successive generations of simulated language learners acquire the language and pass it on to their offspring. This type of simulation is useful for exploring the effect of the **learning bottleneck**: during acquisition, children do not have direct access to their parents’ grammars and have to infer its properties from limited linguistic input. This makes it a good choice for testing the consequences of the PCFP.

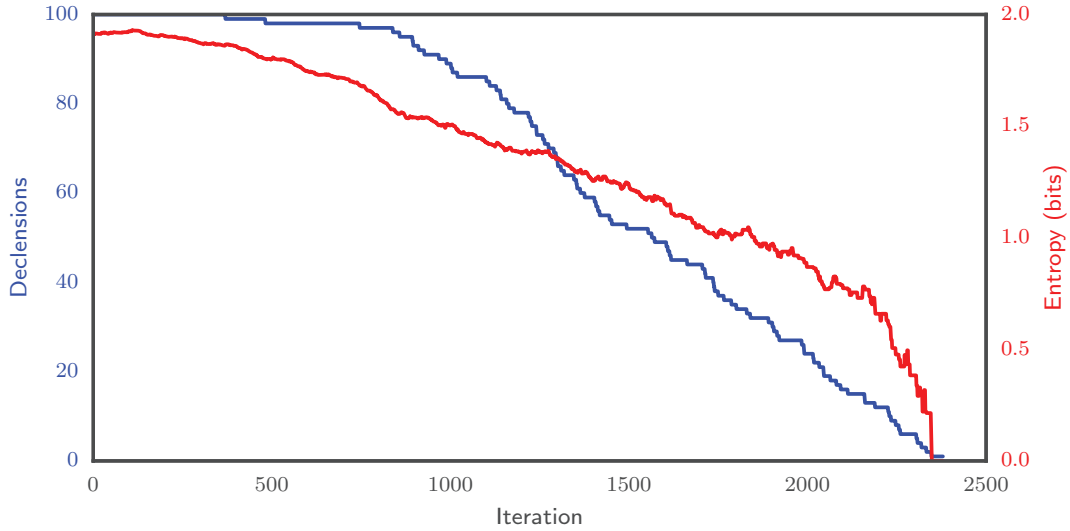
For this simulation, we start with a simple random language with eight paradigm cells and three allomorphs per cell. We generate an initial lexicon of 100 words by selecting randomly from the space of $3^8 = 6,561$ possible inflection classes. The result is a highly unrealistic language – virtually every word is in its own inflection class – but it perhaps reflects what might arise after a sound change obscures a previously predictable morphological system. By any measure, a random language of this type will show very high I-complexity. No paradigm cell will be a good predictor of inflection class, so the system is maximally blurred and also has high average conditional entropy.

For a language of this type, there will be no general way of solving the PCFP. Every lexeme will have to be memorized, as prediction is foreclosed owing to the way the system is organized. To investigate how a ‘difficult’ language like this might evolve, we simulate what might happen if a speaker were forced to predict an unknown form of a lexeme based on analogy with other known forms, as in Figure 1. Specifically, we select an inflection class r , a known predictor cell c_1 , and an unknown predicted cell c_2 . We then collect the inflection classes s with the same realization as r for cell c_1 : these classes form the basis for the analogy. The realization for c_2 which is most common among the classes s is predicted as the realization of c_2 for r . This predicted realization now becomes the new correct realization for c_2 in r , and the inflection class inventory is reconfigured accordingly — if r is now the same as some other class, the two classes are merged into one. The simulation continues, repeating this process until the system stabilizes and 25 iterations pass without any changes.⁴

An example showing the evolutionary pathway followed by one simulated language is given in Figure 2. The language starts with 100 lexical items in 100 different inflection classes, and over the course of the simulation the class differences are completely leveled out and only one class remains. That the number of distinct classes is reduced over time is not surprising: since the starting state is unlearnable (with respect to the PCFP) our simulated speakers are not able to learn it, and the ‘mistakes’ they make will inevitably cause the

⁴Source code for the simulation is available at <http://github.com/rmalouf/bls15>.

Figure 2: The evolutionary course of one simulated language



language to evolve into one that is simple enough to be learnable.⁵ What is interesting is the nature of the simplifications that happen and the types of inflectional systems that these simplifications lead to.

In each simulation run, the simplifications that occur and the final state that is reached depend heavily on many random choices. To explore the distribution of final states — which are common, which are unusual, and which are unattested — we can repeat the simulation many times and aggregate their results. In many cases, as in Figure 2, the final state has only a single inflection class, reflecting a complete leveling of the class differences in the starting state. By necessity these languages obey the NBP, as with only one class any form is trivially diagnostic of class membership, and similarly the average conditional entropy is 0 bits. Since we are primarily interested in the systems with more than one inflection class, these fully reduced systems will be ignored in the discussion that follows.

To gather a representative sample of languages, the complete simulation described above was repeated until 500 languages with more than one inflection class were collected. The distribution of inflection class inventory sizes is shown in Figure 3. Most systems are fairly small (the median number of classes is 12 and 69% have fewer than 20 classes), but there are handful of languages with very large inventories of inflection classes (the largest has 88 classes). The distribution of numbers of blurred cells is shown in Figure 4. The majority of simulated languages (283 in all) have no blurred cells and another 50 have only one blurred cell. In all, 56.7% of the simulation runs lead to a final state that obeys the NBP. But, that leaves 43.3% of the sample as languages that do not obey the NBP, and 36 languages in which all eight paradigm cells are blurred. Finally, Figure 5 shows the distribution of

⁵This effect is reminiscent of the Weak Anthropic Principle (“We must be prepared to take account of the fact that our location in the universe is necessarily privileged to the extent of being compatible with our existence as observers”, Carter 1974) and Stein’s Law (“If something cannot go on forever, it will stop”, Stein 1989).

Figure 3: Number of declensions

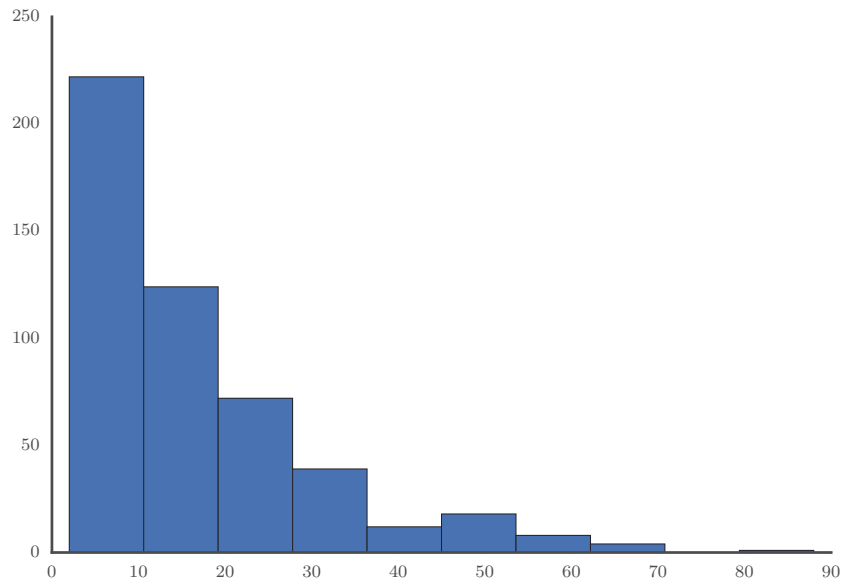


Figure 4: Number of blurred cells (out of 8)

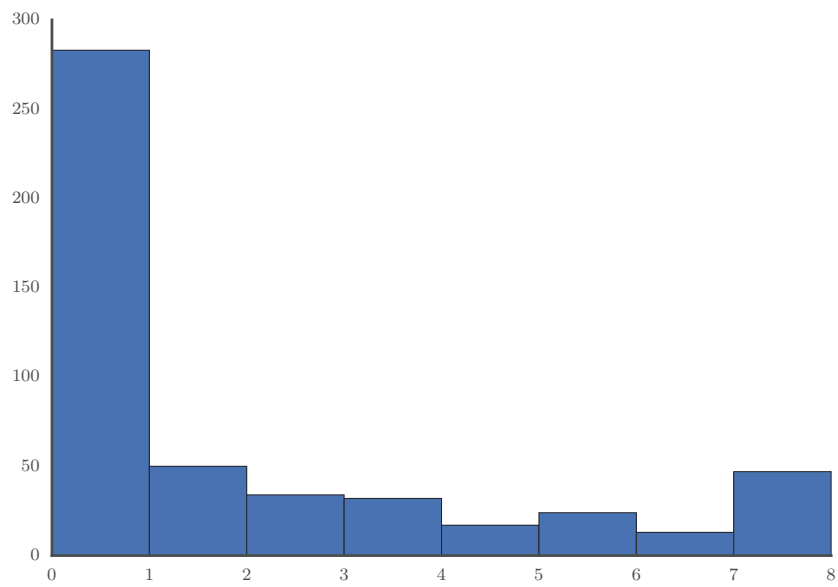
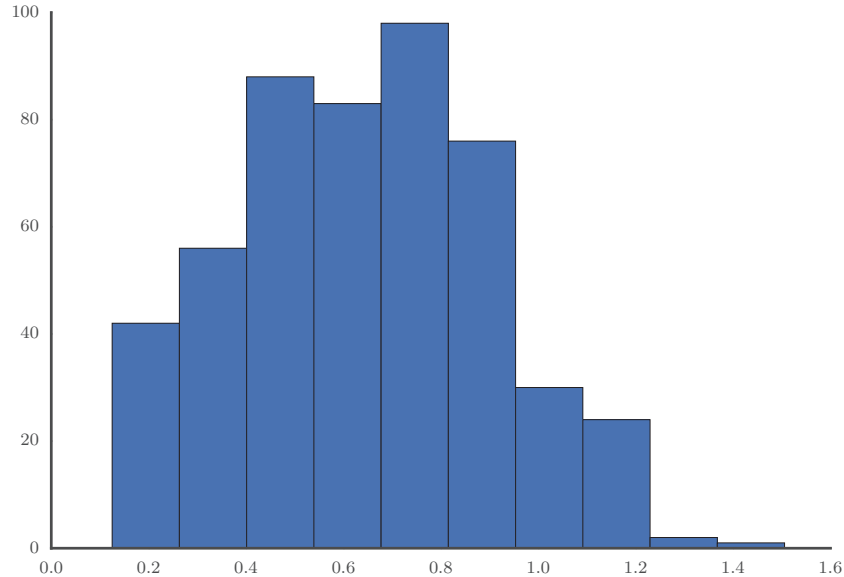


Figure 5: Average conditional entropy (in bits)



average conditional entropies for the simulated paradigms. The mean value is 0.64 bits. In each case, the simulated values fall broadly within the range of values that are observed for real language systems. Many languages obey the NBP, but a minority do not. Most languages with more than one inflection class have a dozen or so classes, but a few have many more than that. Instructively, the average conditional entropy of most paradigms is slightly more than 0.5 bits, but a few are more than 1 bit.⁶

The simulation shown in Figure 1 is a model of the PCFP, as the learner at each iteration must generate an unseen form from a seen form by analogy. Crucially, however, the simulation does not include anything like the Principle of Contrast or a requirement for vocabular clarity. On the other hand, it does reflect a general human inferencing capacity which applies in its own particular way to the Principle of Contrast as a pragmatically based assumption. The basic result, as evident in Figure 3, is that the majority of languages produced by this method do in fact obey the NBP. In this way the NBP can be appropriately interpreted as an **emergent** property of these morphological systems. It is not explicitly enforced by the model at any point, but its effect arises through the dynamics of language transmission. In addition, it becomes clear that in the absence of NBP effects there are still many ways to organize a complex system with acceptably low conditional entropy for solving the PCFP.

One advantage of simulation experiments is that it is easy to vary the details of the model in small ways and observe what consequences this has for the result. For example, in the simulation sketched in Figure 1, the realization chosen for the unknown cell is always the one

⁶Presumably these latter would be lowered if more veridical conditions commonly found in real languages were added to the simulations.

that is most common among the compatible classes s . This can be interpreted as a preference for a default realization (though not for vocabular clarity). We can remove that and repeat the simulation with an alternate random selection strategy, where the simulated language learner chooses a realization for an unknown cell randomly from among the alternatives in s . This is a small change in the model, but it leads to a large difference in the outcome. In this case, the languages undergo no simplification at all. The final states are just as complex (in number of declensions and conditional entropy) as the initial ‘impossible’ state, and in all of them every cell of the paradigm is blurred.

5 Conclusions

Starting with some assumptions about the PCFP, we have traced the evolution of (simulated) languages as they are transmitted through successive generations of speakers and learners until a stable system is reached. Beginning with a language which displays very high paradigm entropy and radically violates the NBP, much simpler systems develop quite quickly. These systems are simple in that they display multiple distinct inflectional classes like extant morphological systems. While neither the NBP nor the Principle of Choice is part of the learning model, many of simulated evolutionary trajectories lead to languages which accord with the NBP. However, many of them do not, achieving low entropy via different organizational strategies. Thus, the dynamics of language learning and use imposed by the PCFP coupled with language operating as a complex adaptive system, provide an indirect explanation for the NBP as an emergent typological regularity found across many but not all languages. In this way the NBP is construable as one type of effect attributable to a general process responsible for morphological organization.

In evaluating the NBP’s empirical status in the face of apparent counterexamples, Carstairs-McCarthy concludes:

“either (a) the Principle is simply wrong, and apparent compliance with it in some languages is purely accidental; or (b) the Principle deserves a place in an overall theory of how inflectional morphology operates, but its effects are obscured or overridden in some circumstances that are not yet well understood.” (2014:61)

In this paper, we have offered a third answer. We propose that the NBP is not a design feature of language and need not be part of morphological theory. But, it is also not accidental that it describes an organization attested in many languages. Rather, as a special case of the LCEC, it is expected that some or even most languages will fit the NBP.

The NBP is not in fact a linguistic universal but rather an emergent property in those languages for which it holds true. This answers Epstein’s challenge: we have shown how transmission of language through the learning bottleneck can lead to observed typological patterns. Recent work approaching language as a **complex adaptive system** (e.g., Blevins & Wedel 2009, Beckner *et al.* 2009) has developed alternate evolutionary explanations for typological universals or tendencies found across unrelated languages. This research emphasizes the role of patterns of language use and language change in the development of cross-linguistic regularities, rather than placing the burden of explanation on synchronic cognitive factors such as an innate Universal Grammar (Evans & Levinson 2009). This

systems approach to language (see also Ramscar & Yarlett 2007, Milin *et al.* 2009, Blevins in press, among others) links linguistic study to larger trends in the biological and social sciences (e.g., Epstein 2006, Miller & Page 2007, Capra & Luisi 2014).

References

- ACKERMAN, FARRELL, JAMES P. BLEVINS, & ROBERT MALOUF. 2009. Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In *Analogy in Grammar: Form and Acquisition*, ed. by James P. Blevins & Juliette Blevins, 54–82. Oxford: Oxford University Press.
- ACKERMAN, FARRELL, & ROBERT MALOUF. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89.429–464.
- BAERMAN, MATTHEW. 2012. Paradigmatic chaos in Nuer. *Language* 88.467–494.
- BAERMAN, MATTHEW. 2014a. Covert systematicity in a distributionally complex system. *Journal of Linguistics* 50.1–47.
- BAERMAN, MATTHEW, 2014b. Floating morphological paradigms in Seri. Paper presented at the Sixteenth International Morphology Meeting, Budapest.
- BAERMAN, MATTHEW, DUNSTAN BROWN, & GREVILLE CORBETT, 2010. Morphological complexity: a typological perspective. Online: http://www.morphology.surrey.ac.uk/Papers/Morphological_complexity.pdf.
- BECKNER, CLAY, RICHARD BLYTHE, JOAN L. BYBEE, MORTEN H. CHRISIANSEN, WILLIAM CROFT, NICK C. ELLIS, JOHN HOLLAND, JINYUN KE, DIANE LARSEN-FREEMAN, & TOM SCHOENEMANN. 2009. Language is a complex adaptive system: Position paper. *Language Learning* 59.1–26.
- BLEVINS, JAMES P. 2000. Inflection classes and economy. In *Explorations in Nominal Inflection*, ed. by Gereon Müller, Lutz Gunkel, & Gisela Zifonun, 41–85. Berlin: Mouton de Gruyter.
- BLEVINS, JAMES P. in press. *Word and Paradigm Morphology*. Oxford: Oxford University Press.
- BLEVINS, JULIETTE, & ANDREW WEDEL. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26.143–183.
- BONAMI, OLIVIER, & SARAH BENIAMINE. 2015. Implicative structure and joint predictiveness. In *Word Structure and Word Usage: Proceedings of the NetWordS Final Conference*, ed. by Vito Pirelli, Claudia Marzi, & Marcello Ferro.
- CAMERON-FAULKNER, THEA, & ANDREW CARSTAIRS-MCCARTHY. 2000. Stem alternants as morphological signata. *Natural Language and Linguistic Theory* 18.813–835.

- CAPRA, FRITJOF, & PIER LUIGI LUISI. 2014. *The Systems View of Life*. Cambridge: Cambridge University Press.
- CARSTAIRS, ANDREW. 1987. *Allomorphy in Inflexion*. London: Croom Helm.
- CARSTAIRS-MCCARTHY, ANDREW. 1991. Inflection classes: Two questions with one answer. In *Paradigms: The Economy of Inflection*, ed. by Frans Plank, 213–53. Berlin: Mouton de Gruyter.
- CARSTAIRS-MCCARTHY, ANDREW. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70.737–88.
- CARSTAIRS-MCCARTHY, ANDREW. 2010. *The Evolution of Morphology*. Oxford: Oxford University Press.
- CARSTAIRS-MCCARTHY, ANDREW. 2014. Affixes and stem alternants in Latvian nouns: implications for inflectional theory. *Baltic Linguistics* 5.59–80.
- CARTER, BRANDON. 1974. Large number coincidences and the anthropic principle in cosmology. In *Confrontation of Cosmological Theories with Observational Data*, ed. by M. S. Longair. Dordrecht: Reidel.
- CLARK, EVE V. 1987. The principle of contrast: A constraint on language acquisition. In *Mechanisms of Language Acquisition*, ed. by Brian MacWhinney, 1–33. Hillsdale, NJ: Lawrence Erlbaum Associates.
- CLARK, EVE V. 1990. On the Pragmatics of Contrast. *Journal of Child Language* 17.417–431.
- CLARK, EVE V. 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- CYSOUW, MICHAEL. 2005. Quantitative methods in typology. In *Quantitative Linguistics: An International Handbook*, ed. by Gabriel Altmann, Reinhard Köhler, & R. Piotrowski, 554–578. Berlin: Mouton de Gruyter.
- ENGER, HANS-OLAV. 2007. The No Blur Principle meets Norwegian dialects. *Studia Linguistica* 61.278–309.
- ENGER, HANS-OLAV. 2013. Vocubular clarity meets Faroese noun declensions. *Folia Linguistica* 47.345–374.
- EPSTEIN, JOSHUA M. 1999. Agent-based computational models and generative social science. *Complexity* 4.
- EPSTEIN, JOSHUA M. 2006. *Generative Social Science*. Princeton, NJ: Princeton University Press.
- ESPER, ERWIN A. 1966. Social transmission of an artificial language. *Language* 42.575–580.

- EVANS, NICHOLAS, & STEPHEN C. LEVINSON. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–492.
- HALLE, MORRIS, & ALEC MARANTZ. 2000. Clarifying ‘blur’: Paradigms, defaults, and inflectional classes. In *Inflectional Identity*, ed. by Asaf Bachrach & Andrew Nevins, 55–73. Oxford: Oxford University Press.
- KIRBY, SIMON. 2002. Learning, bottlenecks, and the evolution of recursive syntax. In *Linguistic Evolution Through Language Acquisition*, ed. by T. Briscoe. Cambridge: Cambridge University Press.
- KIRBY, SIMON, & JAMES R. HURFORD. 2002. The emergence of linguistic structure: An overview of the iterated learning model. In *Simulating the evolution of language*, ed. by A. Cangelosi & D. Parisi, 121–148. London: Springer.
- MILIN, PETAR, VICTOR KUPERMAN, ALEKSANDAR KOSTIĆ, & R. HARALD BAAYEN. 2009. Paradigms bit by bit: an information theoretic approach to the processing of paradigmatic structure in inflection and derivation. In *Analogy in Grammar: Form and Acquisition*, ed. by James P. Blevins & Juliette Blevins, 214–252. Oxford University Press.
- MILLER, JOHN, & SCOTT PAGE. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.
- MÜLLER, GEREON, 2007. Morphological underspecification derives paradigm economy effects. Presented at workshop on Theoretical and Computational Perspectives on Underspecification, Stuttgart.
- PIANTADOSI, S. T., & E. GIBSON. 2014. Quantitative standards for absolute linguistic universals. *Cognitive Science* 38.736–756.
- PLANK, FRANS. 1991. Rasmus Rask’s dilemma. In *Paradigms: The Economy of Inflection*, ed. by Frans Plank, 161–196. Mouton de Gruyter.
- RAMSCAR, MICHAEL, & DANIEL YARLETT. 2007. Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science* 31.927–960.
- STEIN, HERBERT. 1989. Problems and not-problems of the American economy. *The AEI Economist* June 1989.1–8.
- STUMP, GREGORY. 2005. Some criticisms of Carstairs-McCarthy’s conclusions. In *Yearbook of Morphology 2005*, ed. by Geert Booij & Jaap van Marle, 283–303. Dordrecht: Foris.
- STUMP, GREGORY, & RAFAEL FINKEL. 2013. *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press.
- WURZEL, WOLFGANG U. 1986. Die wiederholte Klassifikation von Substantiven. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39.76–96.