

Predicting surface forms in complex inflectional paradigms through phonological constraints

Eric Rosen
Johns Hopkins University

1 Introduction

In this paper, I extend an analysis of inflectional paradigms in Rosen (2019) that derived surface forms of complex inflectional paradigms through blended gradient input forms. The previous analysis was able to predict surface forms in the complex system of Chiquihuitlán Mazatec (Jamieson, 1982; Ackerman & Malouf, 2013), in which exponents could vary among as many as eighteen descriptive inflectional classes that occurred in three cross-cutting dimensions, where there was a potential for a power set of 5775 inflectional class combinations. It was also shown, however, that the model predicted certain paradigm configurations to be impossible, namely, that if exponent x occurs for paradigm cell A and exponent y occurs for paradigm cell B, both for the same lexeme, then it should never occur that the two exponents are switched in those positions for a different lexeme. The following tables, taken from Rosen (2019) illustrate these predictions, with an example from the relatively simple paradigm for Russian noun inflection.

	Real		Impossible	
	Class 3	Class 4	Class 3	Class 4
Gen. Sg.	i	i	i	e
Dat. Sg.	e	i	e	i

Table 1: Crossing diagonals

We do in fact find languages with inflectional paradigms that exhibit this kind of pattern, two examples being Ngití (Finkel & Stump, 2007) and Kwerba (Malouf, 2013), as shown below in Tables 2 and 3.

I show that the exponents in these paradigms can be derived by extending the analysis in Rosen (2019) to include phonological constraints that encode relations between known forms of a paradigm and forms that a speaker wishes to predict. In the word-based morphological literature, the concept of ‘principal parts’ (Finkel & Stump, 2007:inter alia) is widely implemented as a way of encoding these relations, where a principal part is a form that is likely to be known and from which other related forms can be predicted, based on the form of the principal part. Whereas a conventional way of predicting a paradigm form x from a principal part p is through rules of the form “if $p = A$ then $x = B$ ”, here, the approach is through the concept of *syncretism*: the occurrence of the same exponent for more than one paradigm cell, such as the same exponent *-is* for both first and second person singular second conjugation verbs in French. Specifically, the constraint proposed here rewards an exponent for identity with a principal part form and has its weight relativized among morphosyntactic feature combinations (henceforth MFCs), that make up the columns of a paradigm. Intuitively, this varying of the weight of the constraint among paradigm columns captures the fact that we see, cross-linguistically, tendencies for certain pairs of MFCs to be more mutually syncretic than others – for example, genitive singular is syncretic with with nominative plural in some declensions of Latin but in almost no cases is it syncretic with accusative singular. By doing this, we are viewing inflectional paradigms as a kind of dynamic system, where, instead of purely deriving an inflectional form by assembling together a set of input components that are subjected to a set of rules or constraints, the shapes of different forms across a paradigm also depend on their relationships to each other. This latter view of inflectional systems is argued

* This research was in part generously supported by NSF Grant BCS1344269. Thanks also to Paul Smolensky, Matthew Goldrick, James Blevins, Géraldine Walther and four anonymous reviewers for helpful suggestions.

Conjugation	Semi-initial vowel	Infinitive	Imperative singular	Imperative plural	Perfective present	Perfective recent past	Perfective intermediate past	Perfective remote past	Perfective narrative past	Imperfective near future	Imperfective distant future	Imperfective past continuous	Imperfective past habitual	Imperfective past conditional	Subjunctive	Nominalized stem1	Nominalized stem2	Inf. form
MFC# in Malouf		7	1	0	12	13	10	14	11	3	2	5	6	4	15	8	9	
v1a	a/I/U	L	L	L	L	L	L	L	L	L	L	L	H	L	L	L	L	afà-ta 'to cry out'
v1b	O	L	L	L	L	L	L	L	L	H	H	L	H	L	H	L	L	obhì-ta 'to cultivate'
v2a.tr	a/I/U	M	M	M	L	L	L	L	M	M	M	M	H	L	M	M	L	ada-ta 'to cross'
v2a.itr	a/I/U	M	M	M	L	L	L	L	M	M	M	M	H	L	M	M	M	upo-ta 'to climb'
v2b.tr	O	M	M	M	L	L	L	L	L	M	M	L	H	L	M	M	L	òdò-ta 'to guard'
v2b.itr	O	M	M	M	L	L	L	L	L	M	M	L	H	L	M	M	M	òdzi-ta 'to cry'
v3.tr	a/I/U	M	LM	M	M	M	M	M	M	M	M	M	H	L	M	LM	M	ùdhò-ta 'to pour'
v3.itr	a/I/U	M	LM	M	M	M	M	M	L	L	L	H	L	L	M	M	M	ùdhò-ta 'to sleep'
v4.tr	a/I/U	H	LM	LM	LM	LM	LM	LM	H	H	H	H	H	L	H	LM	M	andí-tá 'to write'
v4.itr	a/I/U	H	LM	LM	LM	LM	LM	LM	H	H	H	H	H	L	H	LM	LM	akpé-tá 'to whistle'

The archisegment I is realized as [i] or [ĩ]; U, as [u] or [ũ]; and O, as [o] or [õ].

Table 2: “The tone of the root vowel in Ngiti conjugation (Kutsch Lojenga 1994:217ff)” from Finkel & Stump (2007)

Cl.	1.aug	1.dim	1.du	1.pl	2.aug	2.dim	2.du	2.pl	3.aug	3.dim	3.du	3.pl
1	a	a	ac	ec	a	a	ac	ac	a	a	ac	naN
2	a	naN	aN	eN	a	naN	aN	aN	a	naN	aN	naN
3	a	naN	aN	e	a	naN	aN	a	a	naN	aN	a
4	a	naN	aN	era	a	naN	aN	ara	a	naN	aN	ara

Table 3: Kwerba paradigm (from Malouf (2013))

for extensively by Ackerman et al. (2009b,a); Sims & Parker (2016); Ackerman & Malouf (2016); Blevins (2016, 2006:inter alia).

2 Deriving exponents through blended gradient inputs

In Rosen (2019), inflectional stems and exponents are derived in the framework of Gradient Symbolic Computation (Smolensky & Goldrick, 2016). Smolensky et al. (2020) describe this framework as follows:

“... a novel category of computation – a cognitive architecture that unifies symbolic and neural-network computation. Representations are symbol structures whose components are associated with continuously-varying activation values. Knowledge is represented through weighted constraints, specified by a Harmonic Grammar. This formalism is part of a larger research program in which computation derives outputs from gradient representations in phonology, syntax and semantics (Cho et al., 2017; Faust & Smolensky, 2017; Faust, 2017; Goldrick et al., 2016; Hsu, 2018; Müller, 2017; Rosen, 2016, 2018b,a, 2019; Smolensky et al., 2014; Smolensky & Goldrick, 2016; van Hell et al., 2016; Zimmermann, 2017b,a, forthcoming).”

This framework allows blended gradient inputs. In Rosen (2019), what we think of as an exponent can occur in two input locations: (a) as a ‘base input’ on a lexical base and (b) as a separate ‘inflectional input’ that represents some combination of morphosyntactic features. For example in the Ngiti paradigm shown above that will be analysed in §3, a lexeme in conjugation v1a contains an input blend of tones L and H (activated at 0.75 and 0.2 respectively as explained in §3) and the infinitive MFC has an input blend of tones H and M (activated at 0.1 and 0.3.)

	Lexical base	Inflectional affix
Input:	$\overbrace{\text{afa}}^{\text{root}} \{L, H\}$	INFIN. = $\{L, H, LM\} + /ta/$
Output:	$\overbrace{\text{a}^L \text{f} \text{a}}^{\text{root exponent}} \overbrace{\text{ta}}^{\text{affixal exponent}}$	

Table 4: Infinitive of lexeme, descriptive class v1a

In that account, the choice of an exponent that surfaces results from the possibility for two instances of the same phonological object such as a tonal or segmental feature to coalesce in the output. When there is a blend of input elements with gradient activations in both the lexical base and inflectional affix, as shown in Table 4, the element with the highest aggregate activation will surface in the output. In the GSC framework, an optimal output candidate is chosen based on the aggregate Harmony of that candidate with respect to weighted constraints. Relevant here are MAX and DEP Faithfulness constraints, where MAX contributes positive Harmony equal to the amount of the input activation of a feature that surfaces and DEP penalizes with negative Harmony equal to the amount of the deficit between full activation of 1.0 and the amount of input activation. If only MAX and DEP constraints are relevant, then the exponent with the highest aggregate activation will be the one that surfaces, as long as the result Harmony is above zero.

In the case of the predicted impossible pattern show above in table 1, the activation inequalities necessary to derive this pattern lead to a contradiction. The inequalities are shown in Appendix C.

It should be noted that this approach to inflectional morphology has in common with word-based approaches that it does not strictly separate stems from affixes. The input tones in table 4 occur on both what we refer to as the lexical base and on the inflectional affix. The tonal exponents shown in the Ngit paradigm are part of the root but are listed separately by Finkel & Stump (2007:43), who write:

“Close inspection of these examples reveals that there’s not really any morphological variation from one conjugation to the next in Ngit except with respect to the tone of the root-final vowel. (The members of each conjugation are also generally restricted with respect to the quality of their stem-initial vowel.) We can therefore abstract from the rest of the morphology of these forms as in Table 9 [= Table 2 here], whose horizontal axis lists the different morphosyntactic property sets that vary in their realization from one Ngit conjugation to the next and whose vertical axis lists the conjugations themselves.”

Given the influence of word-based morphology on the present approach, the constraint proposed here, although inspired by Output-Output Faithfulness (Benua, 1997), has some differences from the familiar version of O-O Faithfulness. First, the principal parts it refers to are not exactly the same as a base form, in that there can be more than one principal part for a paradigm¹ and known forms can vary from lexeme to lexeme (although that route is not followed here, but see footnote 4.) Secondly, because the present approach does not draw a clear line between stems and affixes, it does not predict that a constraint that measures identity to a principal part should be measured any differently for what some might regard as a stem versus and affix.²

¹ Finkel & Stump (2007) propose various combinations of three principal parts for Ngit

² An anonymous reviewer asks the question of whether the constraint would apply differently to stems than to affixes. They also comment: “Typically, Base-Faithfulness is given the same ranking for all inflected forms (at least, within some “sub-paradigm”— e.g., for present tense forms, within the plural, etc.), and is not relativized to individual cells (Base-Faith/1pl, Base-Faith/2pl, Base-Faith/3pl, etc.).” Given this comment, it is important to distinguish the approach here from Base-Faithfulness. Here, we are viewing principal parts as forms for which any form could be chosen for the sake of prediction and not as a single distinguished form as in the usual concept of a base form. Rather than measuring whether a form corresponds to some distinguished base, we are viewing a complex array of relations between different forms of a paradigm, where syncretism between forms will occur variably.

3 Gradiently weighted syncretism to a principal part

In spite of the ‘crossing diagonals’ pattern that is evident in the paradigms of Ngiti and Kwerba as shown above, the exponent for a given lexeme and MFC can be predicted from input forms if the choice of exponent is determined by the additive effect of two factors: (a) the aggregate activation of inputs as described above as they are measured by MAX and DEP constraints and (b) Harmonic rewards for syncretism of a candidate exponent to a form that is a principal part. In the case of the Ngiti paradigm, we find through a global beam search of MFCs as possible principal parts, that all exponents can be derived correctly if the two principal parts are nominalized-2 and the perfect recent past forms. The syncretism constraint operates over each of the principal part forms. For each of the two principal parts, for each MFC, there is a Harmonic reward³ if an exponent candidate for that MFC matches the exponent that occurs for the MFC of the principal part.⁴ The reward depends on the weight of the syncretism constraint, which varies according to the MFC in question. The reason for considering constraint weights that vary according to the MFC for which a candidate is being considered is that some MFCs will exhibit syncretism to a principal part MFC while others will not. A learning algorithm, discussed below, found the following weights for syncretism relative to one of the two principal parts for each of the sixteen MFCs in the Ngiti paradigm.

Table 5 shows the learned weights for the constraint relative to each MFC. Those with zero values are left blank. The actual principal parts are bolded.

Infinitive	0.15
Imperative singular	0.2
Imperative plural	0.55
Perfective present	0.55
Perfective recent past	0.55
Perfective intermediate past	0.55
Perfective remote past	0.55
Perfective narrative past	0.15
Imperfective near future	0.05
Imperfective distant future	0.05
Imperfective past continuous	0.15
Imperfective past habitual	0.05
Imperfective past conditional	0.05
Subjunctive	0.05
Nominalized stem 1	0.05
Nominalized stem 2	0.15

Table 5: Learned weights of syncretism to a principal part for each MFC

The tableau in (1) on page 5 shows an example of how these input forms and constraints work for deriving the correct exponent tone for the imperative past MFC for lexeme *udhɔ* ‘pour’ given by Finkel & Stump (2007) in class v3.tr. as shown above in Table 2. For this lexeme class, both principal parts surface with a M tone, so any candidate with a M tone will reap an extra reward of 0.05 for each of the two principal parts, since 0.05 is the learned weight of the Syncretism constraint for that MFC. The learning algorithm worked successfully for many different weights of MAX and DEP, but a low weight of DEP was chosen so that some tonal feature would always surface. It is assumed here that there is a low weight for a UNIFORMITY constraint that would penalize outputs with more than one input correspondent and such a constraint is omitted from the tableaux. For simplicity of exposition, following the comment by Finkel & Stump (2007) on page 3 we only show tonal features as candidates.

The following are the constraints in the analysis:

- DEP: penalizes material with no input correspondent. Harmony penalty is the activation deficit between an output and its corresponding input times the weight of the constraint.
- MAX: rewards a candidate where underlying elements have output correspondents. Reward is based on the amount of each underlying that surfaces times the weight of the constraint.

³ An anonymous reviewer asks why this constraint is formulated positively rather than negatively and observes correctly that it would not make a qualitative difference if it were measured negatively. Because syncretism is a departure from canonicity, as described by Corbett (2007), it makes more sense to reward it when it helps with predicting forms rather than to penalize forms that are canonical by being distinct from a principal part.

⁴ The model assumes that a speaker will know the correct form for each of the principal parts but needs to predict the other forms. That a speaker would know the correct form of some principal part for every lexeme may be an idealization. (See for example, Malouf (2018).) A further step with this analysis would be to allow the possibility of different known forms for different lexemes.

- **SYNCRETISM-TO-PRINCIPAL-PART**: rewards a candidate whose output correspondent matches a principal part to the amount of the constraint weight.

(1)

$\{0.1 \cdot L, 0.2 \cdot M, 0.3 \cdot LM\}$ $\{0.1 \cdot L, 0.35 \cdot H, 0.1 \cdot M\}$ /udhɔ/ ‘pour’ + IMP.PST.	MAX	DEP	SYNC. TO PR. PART		Harmony
	1.0	-0.1	0.05		
			to perf.rec.pst.	to nom.stem.2	
(a) L	0.2	-0.08			0.12
(b) H	0.35	-0.065			0.285
(c) M $\text{ḥ} \text{ḥ}$	0.3	-0.07	0.05	0.05	0.33
(d) LM	0.3	-0.07			0.23

In this example, candidate (c) earns extra Harmony twice for matching the M tone in both principal parts, each of which carries a M tone for this lexeme class. This candidate has a lower aggregate activation from inputs than candidate (b) but the reward from the Syncretism constraint gives it optimal Harmony.

On the other hand, a reward for syncretism to a principal part does not always result in the identity of an exponent to a principal part. In class v.2a.tr., both principal parts have a L tone, and the Syncretism weight for imperfective past continuous is 0.15; nevertheless, that form surfaces with a M tone, because of high input activation on a M tone for the lexeme, as shown in (2).

(2)

$\{0.05 \cdot L, 0.1 \cdot H, 0.6 \cdot M\}$ $\{0.3 \cdot L, 0.15 \cdot H, 0.05 \cdot M\}$ /ada/ ‘cross’ + IMP.PST.CONT.	MAX	DEP	SYNC. TO PR. PART		Harmony
	1.0	-0.1	0.15		
			to perf.rec.pst.	to nom.stem.2	
(a) L	0.35	-0.065	0.15	0.15	0.585
(b) H	0.25	-0.075			0.175
(c) M $\text{ḥ} \text{ḥ}$	0.65	-0.035			0.615
(d) LM	0.0	-0.1			-0.1

4 Learning input activations and constraint weights

As mentioned above, the weights of MAX and DEP were found not to be crucial for determining the correct exponent for each cell in the paradigm and were pre-set at $w_{Max} = 1.0$ and $w_{Dep} = -0.1$. A error-driven learning algorithm was used to find constraint weights for the SYNCRETISM constraint and input activations for the tonal features for each lexeme class and each MFC. The algorithm is similar to Pater & Boersma (2013) but also learns activations as well as constraint weights. The algorithm is modeled after EDGAR in Smolensky et al. (2020) and applies in the same way to the learning simulation that was carried out here. (See Smolensky et al. (2020) for precise details.)

5 Kwerba paradigm

The simpler Kwerba paradigm, repeated in Table 6, can be generated with just one principal part, where any of six out of the twelve MFCs can function as such. If the 3rd.pl. MFC is chosen as the principal part for Kwerba, the following table shows the weights for Syncretism to that MFC. Finkel & Stump (2007) add a note that ‘N’ represents a nasal that is homorganic with the following consonant.

Cl.	1.aug	1.dim	1.du	1.pl	2.aug	2.dim	2.du	2.pl	3.aug	3.dim	3.du	3.pl
1	a	a	ac	ec	a	a	ac	ac	a	a	ac	naN
2	a	naN	aN	eN	a	naN	aN	aN	a	naN	aN	naN
3	a	naN	aN	e	a	naN	aN	a	a	naN	aN	a
4	a	naN	aN	era	a	naN	aN	ara	a	naN	aN	ara

Table 6: Kwerba paradigm (from Malouf (2013))

1.aug	1.dim	1.du	1.pl	2.aug	2.dim	2.du	2.pl	3.aug	3.dim	3.du	3.pl
	0.1				0.1		0.2		0.1		0.6

Table 7: Weights for Syncretism to 3rd.pl. for Kwerba

The following tableau shows how the exponent for MFC 2nd.pl. in class 3 benefits from Syncretism to base form *a* for 3rd.pl. for that lexeme. The learned input for a lexeme of that class has a blend $\{0.05 \cdot a, 0.2 \cdot naN, 0.15 \cdot aN, 0.2 \cdot e\}$ and the MFC 2nd.pl. has a blend $\{0.5 \cdot a, 0.25 \cdot ac, 0.1 \cdot aN, 0.05 \cdot ara\}$. Candidates (b), (d), (e) and (g) all have a higher aggregate activation as shown by the values in the MAX column, but candidate (a) wins because of Syncretism to the base form *a*.

	$\{0.05 \cdot a, 0.2 \cdot naN, 0.15 \cdot aN, 0.2 \cdot e\}$	$\{0.5 \cdot a, 0.25 \cdot ac, 0.1 \cdot aN, 0.05 \cdot ara\}$					
	lexeme 3	+	2nd.PL.	MAX	DEP	SYNC	Harmony
				1.0	-0.1	0.2	
(a)	a			0.1	-0.09	0.2	0.21
(b)	ac			0.25	-0.075		0.175
(c)	ec				-0.1		-0.1
(d)	naN			0.2	-0.08		0.12
(e)	aN			0.25	-0.075		0.175
(f)	eN				-0.1		-0.1
(g)	e			0.2	-0.08		0.12
(h)	era				-0.1		-0.1
(i)	ara			0.05	-0.095		-0.045

6 Discussion

The present approach to the “cell-filling problem” in inflectional paradigms (Ackerman & Malouf, 2013), occupies a middle ground between rule-based approaches (e.g., Baerman (2012) whose rules may require exceptions, and neural-network approaches (e.g., Malouf (2018)) which can handle large amounts of complex data but for which it can be difficult to see exactly where a speaker’s linguistic knowledge is represented within vast arrays of connection weights and distributed representations of linguistic elements in a relatively high-dimensional vector space. The Gradient Symbolic Computation framework provides an interface between symbol structures that are familiar in conventional linguistic theory and a neural basis of cognition in which gradiently-valued elements play a part. In this account, both words and MFCs that determine the content of paradigm cells are represented underlyingly by gradient blends of phonological material and the way they surface also depends on syncretic tendencies across the paradigm for word forms to identify with principal part forms for the same lexeme. A further research step in this framework would be to find ways to encode implicational relations between paradigm cells that goes beyond a static set of principal parts and captures the relations between any pair of cells in a paradigm.

Appendices

A Learned activations for Ngiti

Class ↓	L	H	M	LM
v1a	0.75	0.2		
v1b	0.35	0.4		
v2a.tr.	0.05	0.1	0.6	
v2a.itr.	0.05	0.05	0.4	
v2b.tr.	0.2	0.05	0.4	
v2b.itr.	0.2	0.05	0.3	
v3.tr.	0.1		0.2	0.3
v3.itr.	0.35		0.15	0.1
v4.tr.	0.05	0.5	0.05	0.25
v4.itr.		0.4		0.05
Case ↓	L	H	M	LM
imp.pl.			0.35	0.2
imp.sg.	0.05		0.15	0.4
imp.dst.fut.	0.1	0.35	0.1	
imp.nr.fut.	0.1	0.35	0.1	
imp.pst.cnd.	0.55			
imp.pst.cont.	0.3	0.15	0.05	
imp.pst.hab.		0.6		
inf.		0.1	0.3	
nom.1			0.35	0.35
nom.2	0.1		0.15	
prf.ind.pst	0.2		0.1	0.05
prf.nar.pst.	0.3	0.15	0.05	
prf.pres.	0.2		0.1	0.05
prf.rec.pst.	0.2		0.1	0.05
prf.rem.pst.	0.2		0.1	0.05
subj.	0.1	0.35	0.1	

B Learned activations for Kwerba

Class ↓	a	ac	ec	naN	aN	eN	e	era	ara
1	0.65	0.65	0.45	0.4					
2				0.1	0.2	0.25			
3	0.05			0.2	0.15		0.2		
4				0.1	0.15			0.2	0.1
Case ↓	a	ac	ec	naN	aN	eN	e	era	ara
1-aug.	0.4								
1-dim.	0.2			0.2					
1-dual		0.25			0.15				
1-pl.			0.35			0.2	0.2	0.2	
2-aug.	0.4								
2-dim.	0.2			0.2					
2-dual		0.25			0.15				
2-pl.	0.05	0.25			0.1				0.05
3-aug.	0.4								
3-dim.	0.2			0.2					
3-dual		0.25			0.15				
3-pl.	0.05			0.3					0.05

C Inequalities that show a contradiction

Here, ι represents the activation of i and ϵ of e for a given (subscripted) feature value or descriptive inflectional class.

		Impossible	
		Class 3	Class 4
Gen. Sg.	i		e
Dat. Sg.	e		i

Table 8: Crossing diagonals (repeated from Table 1 on page 1)

$$\iota_{g.sg.} + \iota_3 > \epsilon_{g.sg.} + \epsilon_3 \quad (4)$$

$$\epsilon_{g.sg.} + \epsilon_4 > \iota_{g.sg.} + \iota_4 \quad (5)$$

$$\iota_{dat.sg.} + \iota_4 > \epsilon_{dat.sg.} + \epsilon_4 \quad (6)$$

$$\epsilon_{dat.sg.} + \epsilon_3 > \iota_{dat.sg.} + \iota_3 \quad (7)$$

$$(4) + (5) : \iota_3 + \epsilon_4 > \epsilon_3 + \iota_4 \quad (8)$$

$$(6) + (7) : \iota_4 + \epsilon_3 > \epsilon_4 + \iota_3 \quad (9)$$

(9) contradicts (8)

Table 9: Contradictory inequalities

References

- Ackerman, Farrell & Robert Malouf (2013). Morphological Organization: the Low Conditional Entropy Conjecture. *Language* 89:3, 429–464.
- Ackerman, Farrell & Robert Malouf (2016). *Cambridge Handbook of Morphology*, Cambridge University Press, Cambridge, chap. Implicative Relations in Word-Based Morphological Systems, 272–296.
- Ackerman, Farrell, James Blevins & Robert Malouf (2009a). *Analogy in Grammar: Form and Acquisition*, Oxford University Press, chap. Implicative Relations in Word-Based Morphological Systems, 54–82.
- Ackerman, Farrell, James Blevins & Robert Malouf (2009b). *The Oxford Handbook of Morphological Theory*, Oxford University Press, chap. Word and Paradigm Morphology, 54–82.
- Baerman, Matthew (2012). Paradigmatic chaos in nuer. *Language* 88:3, 467–494.
- Benua, Laura (1997). *Transderivational Identity: Phonological Relations between Words*. Doctoral dissertation, University of Massachusetts.
- Blevins, James (2006). Word-based morphology. *Journal of Linguistics* 42:3, 531–573.
- Blevins, James (2016). *Word and Paradigm Morphology*. Oxford University Press.
- Cho, Pyeong Whan, Matthew Goldrick & Paul Smolensky (2017). Incremental parsing in a continuous dynamical system: sentence processing in Gradient Symbolic Computation: Supplementary Materials 3: processing. *Linguistics Vanguard* 3.
- Corbett, Greville (2007). Canonical typology, suppletion and possible words. *Language* 83:1, 8–42.
- Faust, Noam (2017). How much for that vowel? Talk, Strength in Grammar Workshop, University of Leipzig.
- Faust, Noam & Paul Smolensky (2017). Activity as an alternative to autosegmental association. *Manchester Phonology Meeting*, Manchester, United Kingdom.
- Finkel, Raphael & Gregory Stump (2007). Principal parts and morphological typology. *Morphology* 17:1, 39–75.
- Goldrick, Matthew, Michael Putnam & Lara Schwarz (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition* 19:5, 857–876. ROA 1441.
- van Hell, Janet G., Clara Cohen & Sarah Grey (2016). Testing tolerance for lexically-specific factors in gradient symbolic computation. *Bilingualism: Language and Cognition* 19:5, 897–899.
- Hsu, Brian (2018). Scalar constraints and gradient symbolic representations generate exceptional prosodification effects without exceptional prosody. Handout, West Coast Conference on Formal Linguistics 36.
- Jamieson, Carol (1982). Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verbs. *Language* 48, 139–167.
- Malouf, Robert (2013). Resources for instrumented item-and-pattern morphology. [Github.com/rmalouf/morphology](https://github.com/rmalouf/morphology).
- Malouf, Robert (2018). Generating morphological paradigms with a recurrent neural network. *San Diego Linguistic Papers* 6.
- Müller, Gereon (2017). Gradient symbolic representations in syntax. Handout, IGRA 02: Syntax II.
- Pater, Joe & Paul Boersma (2013). Convergence properties of a gradual learning algorithm for harmonic grammar.
- Rosen, Eric (2016). Predicting the unpredictable: Capturing the apparent semi-regularity of rendaku voicing in Japanese through gradient symbolic computation. *Proceedings of the Berkeley Linguistics Society*, vol. 42. ROA 1299.
- Rosen, Eric (2018a). Evidence for gradient input features from sino-japanese compound accent. *Supplemental proceedings of the Annual Meeting on Phonology*, vol. 6.
- Rosen, Eric (2018b). Predicting semi-regular patterns in morphologically complex words. *Linguistics Vanguard* 4:1. ROA 1339.
- Rosen, Eric (2019). Learning complex inflectional paradigms through blended gradient inputs. *Proceedings of the Society for Computation in Linguistics*, vol. 2.
- Sims, Andrea D. & Jeff Parker (2016). How inflection class systems work: On the informativity of implicative structure. *Word Structure* 9:2, 215 – 239.
- Smolensky, Paul & Matt Goldrick (2016). Gradient Symbolic Representations in Grammar: The case of French Liaison. Rutgers Optimality Archive 1552, Rutgers University.
- Smolensky, Paul, Matthew Goldrick & Donald Mathis (2014). Optimization and Quantization in Gradient Symbol Systems: A Framework for Integrating the Continuous and the Discrete in Cognition. *Cognitive Science* 38, 1102–1138.
- Smolensky, Paul, Eric Rosen & Matthew Goldrick (2020). Learning a gradient model of french liaison. *Proceedings of the Annual Meeting on Phonology*.
- Zimmermann, Eva (2017a). Being (slightly) stronger: Lexical stress in Moses Columbian Salish. Handout, Strength in Grammar Workshop, University of Leipzig.
- Zimmermann, Eva (2017b). Gradient symbols and gradient markedness: A case study from Mixtec tones. Handout: Manchester Phonology Meeting.
- Zimmermann, Eva (forthcoming). Gradient Symbolic Representations in the output: A case study from Moses Columbian Salishan stress. *Proceedings of NELS*, vol. 48.