Inflectional paradigms as interacting systems

Eric Rosen
Johns Hopkins University
`erosen27@jh.edu`

Abstract

In the framework of Gradient Symbolic Computation, (Smolensky and Goldrick, 2016), we present a model that predicts correct forms in complex inflectional paradigms through a single underlying form for a lexeme along with underlying forms for certain morphosyntactic combinations. Output-Output Correspondence constraints (Burzio, 1996; Benua, 1997; Burzio, 1999) capture interdependencies between forms in different paradigm cells. Our model avoids complex sets of rules as well as the need to index lexemes to inflectional classes. Instead, the ways that an exponent can vary across lexemes derive from a lexeme's underlying representation, which can contain partially-activated blends of segments. This approach takes advantage of a blurring of the distinction between stems and affixes and evaluates MAX Faithfulness constraints across a whole paradigm rather than separately for each word form. We present a neural-based gradient ascent algorithm for learning weights and activations that correctly predict output forms, by optimizing the Harmony of a whole paradigm.

1 Introduction

This paper proposes a new approach to the *cell-filling problem* in inflectional paradigms (Ackerman and Malouf, 2013) through a combination of symbolic representations and neural networks. We adopt a paradigmatic view of inflection that belongs to a word-based model of morphology, (Ackerman et al., 2009a,b; Blevins, 2016, 2006) where inflectional forms are derived through whole-word relations and comparisons rather than by constructing a form from component parts. Our approach differs from other word-based approaches in that it derives word forms from a single underlying form of a lexeme but also differs from morpheme-based models (e.g., Halle and Marantz (1994)) by viewing paradigms as fundamental linguistic objects in which correspondence relations between its members determine the shapes of surface forms.

We look specifically at noun inflection in Estonian, which as argued by Blevins (2008), poses serious problems for morpheme-based models that try to segment words into feature-representing morpheme constituents and assemble each word form separately according to some derivational schema. We shall not offer a detailed argument in favour of word-based morphology but refer the reader to the arguments presented in the references above. The new contribution of the present work is that it dispenses with inflectional classes and rather than deriving forms through a complex set of general rules that can be overridden by more specific rules, which themselves can have exceptions, it places the burden of deriving forms in a paradigm more heavily on a grammar that remains constant for all lexemes and on underlying forms for lexemes that can capture the way its exponents surface throughout a paradigm. A reviewer asks whether our proposed partial underlying representations are performing the same function as inflectional classes. Even if Estonian inflectional patterns fell into a small number of discrete, regular inflectional classes, which is far from being the case,[1] there are fundamental differences between the present model and one that indexes lexemes to classes. The essence of inflectional classes is that affixes vary according to the class but here the UR for each affix is the same

---

[1] As Blevins (2008), writes, "the declensional system is not organized into sets of concentric classes and types, but forms networks of interdependent patterns. At one extreme are highly general patterns, which predict the variation in form inventories and paradigm structure that defines traditional declension classes. At the other extreme are idiosyncratic patterns, which characterize small subclasses or even individual items. Between these extremes lie patterns that characterize subtypes or cut across classes."

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 136-147.
Held on-line February 14-19, 2021

throughout and theme vowels are encoded in the UR of the lexeme.[2] Given the complexity and irregularity of the Estonian inflectional system, we propose that gradiently-activated underlying representations can capture this complexity better than an inflectional class system that has many subclasses and many exceptions within those subclasses. (See table 2.) A similar approach to inflection by Rosen (2019) is unable to handle inflectional patterns that are 'flip-flopped' in which, as shown in table 1 for lexeme α, feature or segment x occurs in paradigm cell A and feature or segment y occurs in paradigm cell B, but for lexeme β the exact reverse occurs.

Cell	Lexeme α	Lexeme β
A	x	y
B	y	x

Table 1: A flip-flop pattern

This kind of pattern occurs in many languages – e.g., Ngiti (Finkel and Stump, 2007, Table 9) – and abundantly in Estonian in what Blevins (2008) refers to a 'strengthening' and 'weakening' gradation. We show in §6 that it is crucially the calculation of MAX faithfulness at the paradigm level, along with Output-Output Correspondence constraints between cells that accounts here for this kind of patterning.

The difficulties of deriving inflectional forms in Estonian through the concatenation of a single underlying stem with an affix that represents a morphosyntactic feature are compounded by (a) the abundance of stem alternants that do not correspond to a natural class of morphosyntactic feature combinations (illustrated below in tables 8 and 12) and (b) the occurrence of four different theme vowels (a, e, i, u) which vary independently of inflectional patterns (Blevins, 2006, 551) and do not occur at all in some forms of some lexemes. Instead of a 'constructivist' approach, we view a paradigm as a single, interacting system in which the parts have relationships to each other.

For example, we see correspondences between the genitive singular and nominative plural, where the latter differs from the former by the occurrence of the final consonant d. The two cells are also consistent on whether they have a Q3 syllable or not. Prediction of inflectional forms through implicative relations is discussed by Finkel and Stump (2007); Ackerman et al. (2009a); Sims and Parker (2016); Ackerman and Malouf (2016), inter alia. Here, we depart from an analysis that is based solely on implicational rules because of the complexity that such rules would have to take. For example, even though the partitive singular shows stem correspondences with the genitive plural, there still exist all the string differences between them shown in table 2. The most general rules are given in boldface. These are overridden by more specific rules in italics, and with exceptions to them in small caps.

part. sg. form	gen. pl.	Example
_a	**_ade**	maja 'house'
_na	*_nte*	seina 'wall'
except: LINNA → LINNADE 'city'		
_ra	*_rte*	koera 'dog'
except: JÄÄRA → JÄÄRADE 'ram'		
_t	**_te**	katust 'roof'
_ot	*_ode*	fotot 'photo'
except: RAADIOT → RAADIOTE 'radio'		
_nt	*_nde*	akent 'window'
except: KAANT → KAANTE 'cover'		
_at	*_ade*	teemat 'theme'
except: TÄNAVAT → TÄNAVATE 'street'		
_et	*_emete*	taset 'level'
except: VAHET → VAHEDE 'difference'		
_rget	*_rkmete*	märget 'note'
except: PÖRGET → PÖRGETE 'bounce'		
_lt	*_lde*	sammalt 'moss'
except: KEELT → KEELTE 'language'		
_rt	*_rde*	tütart 'daughter'
except: KOORT → KOORTE 'cream'		
_met	*_mnete*	seemet 'seed'
except: ILMET → ILMETE 'look'		
_det	*_dmete*	seadet 'device'
except: TEADET → TEADETE 'message'		
_u	**_ude**	toru 'tube'
_ikku	*_ike*	tähestik 'alphabet'
_i	**_ide**	armi 'scar'
_ssi	*_ste*	poissi 'boy'
except: PÜSSI → PÜSSIDE 'bush'		
_e	**_ede**	jöge 'river'
_d	**_de**	`maad 'land'
_rd	*_rede*	merd 'sea'
_ld	*_lede*	tuld 'fire'

Table 2: A myriad of string difference patterns

2 An alternative to implicational rules

Instead of a system of general rules that are preempted by more specific rules that in turn have exceptions to them, we propose that a whole

---

[2] Blevins (2008) observes that "[t]he choice of theme vowel is a lexical property of nominals and is not predictable from the phonological form or declension class of an item".

paradigm is generated by a single underlying form, a minimal set of affix-like material that underlies some of the paradigm cells and a set of Correspondence and Faithfulness constraints that determine what form occurs in which cell.

3 Estonian data

The data for this study consist of a set of noun forms annotated for case and number collected from RGCL (2018), with missing case/number combinations for the grammatical cases filled in where possible from lookups on Eesti Keele Institut (2020). Superheavy Q3 syllables were marked from lookups above, resulting in about 2500 usable lexemes for training and testing a model.

4 The GSC framework

The Gradient Symbolic Computation framework (Smolensky et al., 2014; Smolensky and Goldrick, 2016) is well suited to viewing a paradigm as an interacting system. This type of harmonic grammar consists of constraints that are familiar from Optimality Theory (Prince and Smolensky, 1993) but which are weighted rather than categorical and it allows linguistic objects to have partial activation values. Within this model, we can view a paradigm as a single system whose surface forms seek to achieve maximum Harmony with respect to the constraints of the system and the input values of both a lexeme and certain morphosyntactic combinations that are realized in the different cells of the paradigm. We can illustrate with a specific example of one Estonian lexeme, *põhjus* 'reason, cause', whose paradigm is shown below. In this paper we only consider the three cases nominative, genitive and partitive in the singular and plural, given that the forms of the other cases are completely predictable by reference to the genitive forms of each number (Blevins, 2008). Q3 syllables are bolded.

	Sg.	Pl.
Nom.	**põh**jus	**põh**jused
Gen.	**põh**juse	**põh**juste
Part.	**põh**just	**põh**jusi

Table 3: Paradigm for *põhjus* 'reason'

Throughout the language, the genitive singular robustly ends in a vowel, the nominative plural in -d and the genitive plural in -e, usually preceded by /t/ or/d/. Otherwise, the exponents in each cell can vary from lexeme to lexeme. We can consider that the morphosyntactic combinations associated with those three cells are underlyingly an unspecified vowel, /d/ and /{t/d}e/ and that the rest of the affix-like material that appears, such as the t in the partitive singular and genitive plural and the e in the genitive singular and nominative plural are part of the underlying representation of the whole lexeme, given that they can vary across lexemes. The underlying representation of the lexeme is then /põhjus{e,t,i}/, where {e,t,i} is a blend of the three segments that can occur at the end of the stem and before the -d and the -$\{t, d\}$e that always occur in the nominative and genitive plural. We also find that the partitive plural typically has more affixal material than the nominative singular and for reasons to be given below, we propose that it have an underlying form of pure activation ϕ.[3]

lexeme input: /**põh**jus{e,t,i}/

	Sg.	Pl.
Nom.		/d/
Gen.	/V/	/{t,d}e/
Part.		ϕ

Table 4: Input forms of morphosyntactic combinations (ϕ = pure activation)

An examination of the data reveals robust stem correspondences between members of a 'group A': the genitive singular and nominative plural and members of a 'group B': the partitive singular, nominative singular and genitive plural, where the partitive plural can pattern with either group. We capture this patterning by strongly weighted group-internal Output-Output correspondence constraints (Benua, 1997) that provide harmonic reward to the system for each segment in an output form of one group that corresponds to a segment in a different output form in the same group.

A reviewer asks how well this kind of constraint would scale to other languages and to what extent it is language-specific. Correspondence between stems across sets of paradigm cells that do not form a natural class occurs not only in Estonian. For example, certain classes of French irregular verbs show stem correspondence among sets of

---

[3] The idea of underlying pure activation in the GSC system was first proposed by Smolensky and Goldrick (2016) for feminine gender in French.

paradigm cells whose feature combinations do not constitute a natural class. In high-frequency verbs like *boire* 'drink', *croire* 'believe', *pouvoir* 'be able to', the 1st and 2nd person plural present indicative forms share a stem that differs from other person-number combinations in the present indicative but mirrors all imperfect indicative forms: e.g., *je bois* 'I drink', *nous buvons* 'we drink', *je buvais* 'I drank'. In such cases, a speaker's grammar would naturally find a way to capture stem correspondence between these groups of forms in the inflectional paradigm. In Estonian, the stem correspondence among certain paradigm cell groups occurs robustly among nouns with few exceptions. Here, this constraint is assumed to be learned outside of the learning algorithm. Estonian comes out better than French in not needing to refer to an inflectional class to determine what kind of stem correspondence occurs. Which O-O constraint is invoked depends on the particular combination of inflectional features but not on the specific noun.

In viewing the paradigm as a total system, we calculate the Harmony across the whole paradigm rather than separately for each cell. A MAX Faithfulness constraint rewards an input segment for surfacing in the paradigm, but only once if it occurs multiple times. To ensure that a stem surfaces throughout the paradigm, pan-paradigmatic correspondence constraints for stems apply across the paradigm, but with a weaker weight than the group-internal ones, since, stems can vary across groups, as shown below in table 8, where groups vary in the occurrence of a Q3 syllable. DEP constraints penalize, with negative Harmony, any deficit between the output activation of a segment in a cell and its input activation, in case the input segment is only partially activated.

Putting these proposed input forms and constraints together, we can calculate the following. The net Harmony for a given blended input segment such as e, t or i to surface once in the paradigm is $Ma_x - D(1-a_x)$, where M and D are the weights of the MAX and DEP constraints and a_x is the input activation of a segment x. Thus, if $a_x > \frac{D}{M+D}$, x will surface somewhere in the paradigm. For it to surface *twice*, would require that $Ma_x - 2D(1 - a_x) > Ma_x - D(1 - a_x)$ or $a_x > 1$, *unless* the segment gets a reward other than MAX for surfacing again: i.e., correspondence to an identical output segment in another cell. For certain constraint weights and input activations, two of the partially activated, blended input segments could each surface twice, iff they occur in pairs of cells that have mutual correspondence rewards. A third such input could surface once somewhere else. Further, if the affix-like segments -d and -e always appear in the nominative plural and genitive plural, we can take them to have full activation in the input, and a strongly weighted ANCHOR constraint will ensure that they occur at the right edge of the word. A phonotactic constraint that prohibits a final [td] sequence will prevent the input /t/ segment from occurring in the nominative plural before the final -d.

5 A learning algorithm for constraint weights and activations

To simulate the learning of constraint weights and activations that derive all the forms of the paradigm, we ran a gradient ascent algorithm on a randomly initialized set of activations and constraint weights, assuming that any segment that occurs in some but not all cells of the paradigm has some partial activation in the lexeme's input. We randomly initialized predicted output forms so that they could range over possible outputs in order to rule out non-occurring ones. A quantization constraint (Tupper et al., 2018) forces outputs to converge to being discrete. The loss function has two components: (a) the negative of the overall Harmony of the system and (b) the sum of squared differences between predicted output and target output activations. The target error is decayed by the square of the iteration number so the model can range over possible outputs. The weight of the quantization constraint increases linearly with each iteration so that it increasingly seeks to make outputs discrete. A sigmoid function is applied (a) to input activations to keep them between 0 and 1, and (b) to the absolute value of the weight of DEP to keep its harmonic penalty negative. Input activations are frozen at a certain point so that the model can explore different possible outputs. Running this simulation on the word *põhjus* resulted in the values in table 5. A proxy $\emptyset$ was used to represent the non-occurrence of an affix in order to facilitate quantization. a_x represents the learned input activation of segment x.

To test if these values correctly predict the paradigm of *põhjus*, we can observe that most possible output combinations are harmonically

Parameter	value
MAX	0.76
DEP	−0.44
Correspondence	1.37
$*\text{td}]_\sigma$	−0.51
a_e	0.621
a_i	0.562
a_t	0.564
a_v	0.648
$a_\emptyset$	0.526
a_ϕ	0.517

Table 5: Learned parameter values

bounded.[4] Even for a_i, the lowest-activated of the blended input segments, $Ma_i - D(1 - a_i) = 0.76 \cdot 0.562 - 0.44 \cdot (1 - 0.562) = 0.234$, i.e., positive net Harmony. Thus we can exclude any candidate combination in which all three segments do not surface at least once in the paradigm. If the same phoneme π surfaces in a cell-pair that has a pairwise CORRESPONDENCE reward, its net harmonic reward for surfacing a second time is positive: $C - D(1 - a_\pi) = 1.37 - 0.44 \cdot (1 - 0.562) = 1.177$ for the lowest activated segment, so we can rule out any candidate set that lacks two input segments each surfacing twice in corresponding cells. Phonotactic constraint $*\text{td}]_\sigma$ rules out input /t/ in the nominative plural. This narrows down the possible candidate sets to the sets of affixal segments shown in table 6.

	Sg.	Pl.		Sg.	Pl.
Nom.		e/i + d	Nom.	e/i/t	e/i + d
Gen.	e/i	t + e	Gen.	e/i	t + e
Part.	t	e/i/t	Part.	t	

Table 6: Narrowed-down set of possible affixes

The second set is ruled out, since it is more harmonic to have no affix surface in the nominative singular than partitive plural, which can add pure activation to the input segment to reduce the DEP penalty. Because of the correspondence reward to have either e or i in both the genitive singular and nominative plural, the other needs to surface in the partitive plural to reap its MAX reward, ruling out surfacing of t there. This leaves two candidates, of which the first (correct one) is more Harmonic since the DEP penalty for the i to surface occurs only once rather than twice.

Nom. sg.	Gen. sg.	Part. sg.	Nom. pl.	Gen. pl.	Part. pl.	H
	e	t	ed	te	i	7.78
	i	t	id	te	e	7.72

Table 7: Harmonies (details in Appendix A)

6 Flip-flop patterns of gradation

As discussed by Blevins (2008), some Estonian nouns have 'strengthening' gradation patterns, with a Q3 syllable only in the genitive singular and nominative plural, or, a 'weakening' pattern, where the site of Q3 occurrence is reversed. In both cases, the partitive plural usually also has a Q3. A weakening pattern has a vowel-final partitive singular and a strengthening pattern a t-final partitive singular.[5] The following table illustrates the two patterns for *kaev* '(water) well' and *vihje* 'tip'. Q3 syllables are shown bolded. As mentioned in §1 this kind of pattern is not explainable in the account proposed by Rosen (2019).[6]

	Sg.	Pl.	Sg.	Pl.
Nom.	**kaev**	kaevud	vihje	**vih**jed
Gen.	kaevu	**kae**vude	**vih**je	vihjete
Part.	**kae**vu	**kae**vusid	vihjet	**vih**jeid

Table 8: Flip-flop patterns

These patterns suggest that there are highly weighted correspondence constraints for moras on stems among each of the two groups. If a Q3 syllable occurs in *any* of the cells, it will occur in the partitive plural, which suggests a one-way correspondence constraint that rewards the occurrence of a mora in the partitive plural that corresponds to a mora in any of the other cells.

To derive both patterns, we propose to add the following input specification for the partitive singular. Smolensky and Goldrick (2016, 25) propose, following Faust and Smolensky (2017) that

---

[4] Thanks for Jane Lutken (p.c.) for pointing this out.

[5] There is also a small class of weakening-pattern nouns (15 in the dataset) with a coronal-sonorant-final monosyllabic stem that takes final /t/ in the partitive singular, which is phonotactically licit after the sonorant. (e.g., **hiir** 'mouse') This pattern can be explained by the monosyllabicity of the stem, which requires the repair of a Q3 syllable in the affixless nominative singular in order to satisfy a minimal word requirement. (See page 6.)

[6] In addition to partitive singular forms that are vowel-final and t-final, there are also d-final forms for some lexemes; these always have a consistent Q3 syllable throughout the paradigm. For reasons of space, it is not possible to analyse these forms in this paper. (around 1.7% of the data)

in the GSC framework, activation can be shared between alternating segments. In this case, we propose the input form /$\tau \cdot \{t, \mu\}, V$/ for the partitive singular, where t and μ share activation τ and V is a vowel unspecified for Place. All of the input is considered as a single segment with no linear ordering among its elements. A prohibitive DEP violation prevents any of this input from surfacing unless it can coalesce with a matching input in the stem, which we posit does not occur in nouns like *heli* 'sound' which has no visible exponent for the partitive singular, with theme vowel /i/ occurring throughout the paradigm.

If, at the edge of the stem input, there is a partially activated /t$_1$/, it can coalesce with /t$_2$/ in the partitive singular UR with activations $\tau_1 + \tau_2$ and surface, but because it shares activation $\omega = \tau_2 + m_2$ with a mora, the mora cannot contribute to the output.

$$/v\overset{\mu\{m_1 \cdot \mu\}\mu}{i}hj\overset{\mu}{e}\{\tau_1 \cdot t_1\}/ + /\{\tau_2 \cdot t_2, m_2 \cdot \mu\}, V/$$

If a partitive-t-final lexeme with weakening gradation has an extra mora (bracketed above for *vihje*) of partial input activation m_1, a MAX constraint will reward this mora for surfacing once. DEP will penalize each cell in which it surfaces by $D(1 - m_1)$. If the MAX reward for the mora to surface is greater than the sum of DEP penalties for those three cells ($M \cdot m_1 > 3 \cdot D(1 - m_1)$), the Q3 will occur in all three and satisfy group stem correspondence. No stem-mora correspondence constraint will force the extra mora to surface in the other three cells. If the extra mora were instead to surface in the nominative singular, partitive singular, genitive plural and partitive plural, the DEP penalty would be multiplied by *four* instead of three, which would be less harmonic. Notice that the extra mora in the partitive singular UR cannot contribute to the surfacing extra mora here, since shared activation $\omega = \tau_2 + m_2$ has been used up for t$_2$ to coalesce with t$_1$.

But when the partitive singular has a final vowel, it must occur at the right edge of the stem input instead of a /t/. In many such cases, the stem is monosyllabic with no affixal material in the nominative singular, as with *kaev*, shown above. Because of a minimal word requirement, a nominative singular form without a Q3 syllable to form a foot will be illicit (Blevins, 2008), so for it to surface without the extra mora will be less harmonic than for it to surface with the extra mora and incur the DEP penalty. And in the partitive singular, not parsing the /t/ of the $\{t, \mu\}$ with shared activation frees the mora to coalesce with the partially-activated extra mora in the stem and lessen its DEP penalty. Thus, only the genitive plural and partitive plural incur a DEP penalty for the extra mora to surface that is (a) not reduced by coalescence with the mora in the partitive singular and (b) not the least-of-two-evils option for avoiding a minimal word violation.

$$/k\overset{\mu\{m_1 \cdot \mu\}\mu}{a}ev\overset{\mu}{u}/ + /\{\tau_2 \cdot t_2, m_2 \cdot \mu\}, V/$$

To simulate the learning of constraint weights and activations for these patterns, we ran the gradient ascent algorithm for four paradigm types in parallel: lexemes like *kaev* 'water well' that shows weakening gradation, *vihje* 'tip' that shows strengthening gradation, *jõud* 'force' that has a Q3 throughout the paradigm, and *oja* 'stream' with no Q3 in the paradigm.[7] The constraint weights are held constant across the four examples and we assume that there is no extra mora in the input for the never-Q3 type and a fully-activated extra mora in the input of the always-Q3 type. The following table shows the constraint weights and input activations learned by the algorithm. We take the MAX-μ and DEP-μ to be different from the MAX and DEP constraints for segments whose learned values were shown above in table 5.

Parameter	value
MAX-μ	2.9
DEP-μ	-0.53
CORRESPONDENCE	0.63
CORRESP-PART.PL.	0.63
*SUBMINWORD	-0.28
a_{μ_stem}	0.39
$a_{\mu\_part.sg.}$	0.39

Table 9: Learned parameter values

The following tableau shows, for a lexeme like *vihje* with a partially-activated extra mora in the input and a t-final partitive singular, how some possible output candidate sets compare for relevant Harmony values with respect to where an extra mora might surface. A μ symbol in the MAX column indicates the extra mora surfacing in that cell. A MAX-μ reward is given as long as at least one cell has the extra mora, in this case, $2.9 \times 0.39 =$

---

[7] Examples of each of the latter two types are found in both the t-final partitive group and the vowel-final partitive group.

1.13. The DEP column tallies DEP-μ penalties: $-0.53 \cdot (1 - 0.39) = -0.32$; the CORR column, Correspondence or Corresp-part.pl. rewards. The latter is a one-way correspondence constraint that rewards the occurrence of a mora in the partitive plural when it occurs in any other cell.

	Q3 everywhere			Weakening			Strengthening		
	MAX	DEP	CORR		DEP	CORR		DEP	CORR
N.sg.	μ	$-.32$	0.63	μ	$-.32$	0.63			0.63
G.sg.	μ	$-.32$	0.63			0.63	μ	$-.32$	0.63
P.sg.	μ	$-.32$	0.63	μ	$-.32$	0.63			0.63
N.pl.	μ	$-.32$	0.63			0.63	μ	$-.32$	0.63
G.pl.	μ	$-.32$	0.63	μ	$-.32$	0.63			0.63
P.pl.	μ	$-.32$	0.63	μ	$-.32$	0.63	μ	$-.32$	0.63
Tot.	1.13	-1.92	3.78	1.13	-1.28	3.78	1.13	$-.96$	3.78
Net H	2.99			3.63			☞3.95		

	Strengthening no Q3 on part.pl.			Weakening no Q3 on part.pl.			No Q3 anywhere		
	MAX	DEP	CORR		DEP	CORR		DEP	CORR
N.sg.			0.63	μ	$-.32$	0.63			0.63
G.sg.	μ	$-.32$	0.63			0.63			0.63
P.sg.			0.63	μ	$-.32$	0.63			0.63
N.pl.	μ	$-.32$	0.63			0.63			0.63
G.pl.			0.63	μ	$-.32$	0.63			0.63
P.pl.									
Tot.	1.13	$-.64$	3.15	1.13	$-.96$	3.15			3.15
Net H	3.64			3.32			3.15		

Table 10: Strengthening pattern wins with final t in partitive singular

The correct strengthening pattern is optimal because it has the lowest DEP penalty while maintaining the full possible Correspondence rewards. Having no Q3 anywhere avoids a DEP penalty but loses mainly from not getting a MAX reward.

In the case of a lexeme with a vowel-final partitive singular and monosyllabic stem, we can assume that a constraint banning noun forms of less than a foot is weighted to rule out such candidates. This means that a DEP violation for an extra mora to surface in the nominative singular is better than violating this constraint. In addition, not parsing the /t/ in the partitive singular cell frees up the activation of the extra mora there to add to the input activation of the extra mora in that cell. This changes the Harmonies in the tableaux as shown in table 11. Instead of the $-0.53 \cdot (1 - 0.39) = -0.32$ DEP penalty in the partitive singular we have $-0.53 \cdot (1 - 0.39 - 0.39) = -0.17$ where 0.39 is the input mora activation in the partitive singular. The contribution of MAX is now $2.9 \cdot (0.39 + 0.39) = 2.26$ when the partitive singular surfaces with the extra mora. And to have a strengthening pattern but with a Q3 in the nominative singular (not shown in tableau) would save the minimal word violation but at the expense of losing the greater Correspondence reward for the nominative singular with the other cells that usually correspond.

	Q3 everywhere			Weakening			Strengthening		
	MAX	DEP	CORR		DEP	CORR		DEP	CORR
N.sg.	μ	$-.32$	0.63	μ	$-.32$	0.63	*		0.63
G.sg.	μ	$-.32$	0.63			0.63	μ	$-.32$	0.63
P.sg.	μ	$-.17$	0.63	μ	$-.17$	0.63			0.63
N.pl.	μ	$-.32$	0.63			0.63	μ	$-.32$	0.63
G.pl.	μ	$-.32$	0.63	μ	$-.32$	0.63			0.63
P.pl.	μ	$-.32$	0.63	μ	$-.32$	0.63	μ	$-.32$	0.63
*SubminWd							-0.28		
Tot.	2.26	-1.77	3.78	2.26	-1.13	3.78	2.26	$-.96$	3.78
Net H	4.27			☞4.91			4.80		

	Strengthening no Q3 on part.pl.			Weakening no Q3 on part.pl.			No Q3 anywhere		
	MAX	DEP	CORR		DEP	CORR		DEP	CORR
N.sg.	*		0.63	μ	$-.32$	0.63		*	0.63
G.sg.	μ	$-.32$	0.63			0.63			0.63
P.sg.		$-.17$	0.63	μ	$-.17$	0.63			0.63
N.pl.	μ	$-.32$	0.63			0.63			0.63
G.pl.			0.63	μ	$-.32$	0.63			0.63
P.pl.									
*SubminWd		-0.28					-0.28		
Tot.	2.26	$-.81$	3.15	2.26	$-.81$	3.15			3.15
Net H	4.32			4.60			2.87		

Table 11: Weakening pattern wins with vowel-final partitive singular (* = *SubminWd violated)

If there were no extra mora at all in the input, the DEP penalties would increase to -0.53 per cell and it would get no MAX rewards. This would optimize the ‘no Q3 anywhere’ candidate, with no DEP penalties, Harmony of 3.15 for the t-final stem and 2.87 for the monosyllabic v-final stem, from Correspondence rewards minus any *SUBMINWD penalties. The only way a different candidate could be more Harmonic would be for it to have Correspondence in 6 cells, which would require the extra mora to surface in at least 3 cells, with a DEP penalty of at least $-0.32 - 0.32 - 0.17 = -0.81$ (when a partitive singular can boost the extra mora input): greater than the extra Correspondence reward of 0.63.

In summary, these flip-flop patterns are derived through a partially-activated extra mora which optimally does not surface in every paradigm cell, but it can surface in one of the two sets of cells whose stem forms are tied together by Correspondence. When no other factors come into play, it is more harmonic for the extra mora to surface in the smaller set, (i.e., genitive singular, nominative plural and partitive plural) which will incur only three DEP violations. But for lexemes with a vowel-final partitive singular, two factors will favour having the extra mora surface in the opposite set of cells: (1) availability of a partially-activated mora in the input of the partitive singular cell to coalesce with the extra mora in the stem UR, (2) the need for the nominative singular to receive an extra mora so that it can be parsed as a foot and satisfy the minimal word requirement.

7 Paradigm uniformity

As mentioned on page 4 this analysis posits pan-paradigmatic Correspondence constraints on stems that ensure that some semblance of the stem input of a lexeme surfaces throughout the paradigm, given that the MAX constraint only rewards one instance of an input surfacing. If a lexeme has a Q3 syllable throughout the paradigm, and there is full input activation on the extra mora (resulting in no DEP-μ penalty), this constraint, weighted weaker than correspondence within each of the two more closely corresponding cell groups, will make it more harmonic for all cells to surface with a Q3 than for just one group of them to do so. Note that without this constraint, there is no harmonic difference between the two possibilities. And as long as this constraint is weighted less than the any DEP penalty in tables 10 or 11, it will not incorrectly cause the Q3 to surface in all cells for those lexemes.

8 Other approaches

In the NLP domain, there have been many recent neural-network-based approaches to predicting forms in inflectional paradigms. Cotterell et al. (2017) create graphical models of implicational relations between paradigm cells inspired by the concept of principal parts of a paradigm that can be used to predict other forms. Annual installments of the Sigmorphon shared task on morphological reinflection (e.g., Cotterell et al. (2016)) contain numerous NLP approaches for morphological prediction. The present work differs from these approaches in that it seeks to explicitly supply linguistic principles and show how they work. Purely neural models have a tendency to be opaque to full linguistic interpretation.

Chuang et al. (2019) present a model specifically for Estonian that uses algebraic matrices to predict word forms from meanings and vice versa. A word form is represented by a binary vector that indicates which triphones it contains among all the triphones of the paradigm. Its semantic form is the sum of a vector representing the lexeme and a vector for each occurring morphosyntactic feature. Form and semantic vectors are mapped to each other through linear transformations. Because a non-ordered set of triphones must be algorithmically reconstituted to a linear string, it is not clear just how those transformations encode grammatical knowledge.

Malouf (2016) provides a recurrent neural network model that predicts, with good accuracy, inflectional forms in seven languages with inputs consisting of vectors representing lexemes and morphological features. He does demonstrate some interpretability with the model in the form of meaningful clustering of representations of phonemes and morphosyntactic feature combinations and correlation of principal components of wordform vectors with syllable and morphological structure. This study differs from his in that it gives lexemes and morphosyntactic combinations phonologically interpretable URs rather than abstract vectors and it has constraints that are interpretable in the wider scope of linguistic theory.

As an alternative to this present model, adapting code from Schlag (2019), we experimented with a neural transformer model that binds fillers to roles with tensor products (Schlag et al., 2019), and which is trained to predict one form from another: e.g., nominative singular forms from partitive singular forms with no other information being given to the model. On the same set of data, this model achieved test accuracies exceeding 96% for three separate predictions among the six case-number combinations. In spite of efforts to examine graphical maps that show attention weights between various vector encodings of segments, we found the model to be too black-box-like to be interpretable to the same extent as the present GSC model.

9 Predicting from partial data

Suppose that a speaker had seen two forms of a paradigm but none of the others. To what extent could they deduce a UR for the lexeme and other surface forms, if they knew URs for certain morphosyntactic combinations shown in table 4 above and repeated below? Consider lexeme *leek* 'flame', which exhibits what Blevins (2008) refers to as 'qualitative gradation' with alternations of consonants k/g in the stem:[8]

	Sg.	Pl.		Sg.	Pl.
Nom.	**leek**	leegid	Nom.		/d/
Gen.	leegi	**lee**kide	Gen.	/V/	/{t,d}e/
Part.	**lee**ki	**lee**ke	Part.	/{t, μ} V/	ϕ

Table 12: Paradigm for *leek* ‘flame’ and URs for affixes

If a speaker knew the nominative singular and genitive singular, they might deduce a UR /lee{k,g}{i}, with the /k/ and /g/ sharing partial activation and the /i/ partially activated. Correspondence between the nominative singular and partitive singular would make **leek** the most harmonic stem for the partitive singular. An /-i/ suffix there would also be more harmonic than a /t/, since parsing the /V/ from the part.sg. input frees up the extra mora activation in its UR to make the Q3 surface and correspond to the nominative singular. In the nominative plural, parsing the /i/ to precede the /d/ suffix is most optimal and the stem /leeg/ would correspond correctly with the stem of the genitive singular. In the genitive plural, having the /i/ surface before suffix /-{d,t}e/ is more phonotactically probable than not, given that the only word-final /ktV/ sequences in the data are in 9 foreign borrowings such as *kontakte* or *konflikte*. Whether a /t/ or /d/ surfaces there is not completely predictable. Partitive plurals are the most difficult to predict. In this case the affixal /-e/ is not completely predictable other than that e/i and i/e alternations between partitive singular and such ‘short’ partitive plural forms are common.[9]

10 Further steps

For reasons of space, this analysis does not cover all the patterns and sub-patterns of noun inflection found in the language but is intended to show how a model that seeks to maximize Harmony across a whole paradigm and that accounts for implicative relations through correspondence constraints could be applied to further data.[10] A further step is to test more cases of how predictions can be made from exposure to partial data within this system.

As mentioned above, another factor that can add to prediction is gradient phonotactics (Hayes and Wilson, 2008, inter alia). For example, a common sub-pattern has nominative singulars with /-ine/ and partitive singulars with /-ist/: e.g., *kuulmine, kuulmise, kuulmist* ‘hearing’ (singular only). If the UR is /kuulmi{n,s}, {e,t}, with blended segments, and if the best MAX reward is to have all four segments surface somewhere, a form like **kuulmint* would be discouraged by a gradient phonotactic dispreference for a word-final *-int* sequence, which occurs only once in all the data and only stem-finally, in lexeme *lint* ‘tape, ribbon’.

In summary, this study shows how inflectional paradigms, including those with complex flip-flop patterns can be generated through (a) input forms of lexemes with partially-activated affix-like segments and (b) correspondence constraints that capture patterns of cell interrelatedness (e.g., genitive singular with nominative plural), where Harmony is maximized across a whole paradigm.

Acknowledgements

Thanks to Paul Smolensky for advice and suggestions at all stages of this research. Also to Nebojsa Jojic, Roland Fernandez and Jake Russin for discussion of an earlier version of this work. Three anonymous reviewers, as well as audience members at a JHU Cog. Sci. brown-bag lunch talk provided valuable suggestions. Thanks also to Caitlin Smith, Matthias Lalisse, Tom McCoy, Najoung Kim, Paul Soulos and Coleman Haley of the JHU Neurosymbolic Computation group for helpful feedback. Research was generously sponsored by Microsoft Research and NSF INSPIRE grant BCS-1344269. All errors are my own.

---

[8] This k/g alternation is not simply final obstruent devoicing, since *k* occurs as an onset in some cells. There are cases of qualitative gradation where stem allomorphs diverge even more: e.g., *pidu* (nom.sg.), ***peo*** (gen.sg.) ‘party’. These facts make it difficult to maintain the hypothesis of Albright (2002) that inflectional paradigms have single bases and that their UR’s must be based on a single base form.

[9] A reviewer asks “whether and how optimal weights can be learned from partial exposure to inflected forms”. As well as creating input-output and output-output derivations in a harmonic grammar, the model is also a proxy for morphological prediction as uncertainty minimization through implicative relations as described by Ackerman et al. (2009b), which “explains phenomena in terms of the dynamics of interdependencies within complex adaptive systems.” An input representation of a lexeme learned from partial data will make predictions of unseen forms based on constraint weights that were learned from data encountered so far. Those learned weights and representation will produce comparative harmonies for possible candidates for unseen forms, which can be translated into probabilities. There is not necessarily an optimal set of weights that is necessary for making such predictions. Ackerman et al. (2009b) see the task as reducing the conditional entropy of a paradigm cell based on knowledge of other cells.

[10] The idea that morphological forms are derived through gradiently-weighted Output-Output Correspondence constraints was also proposed by Burzio (1999) for derivational morphology.

References

Farrell Ackerman, James Blevins, and Robert Malouf. 2009a. *Analogy in Grammar: Form and Acquisition*, chapter Implicative Relations in Word-Based Morphological Systems. Oxford University Press.

Farrell Ackerman, James Blevins, and Robert Malouf. 2009b. *The Oxford Handbook of Morphological Theory*, chapter Word and Paradigm Morphology. Oxford University Press.

Farrell Ackerman and Robert Malouf. 2013. Morphological Organization: the Low Conditional Entropy Conjecture. *Language*, 89(3):429–464.

Farrell Ackerman and Robert Malouf. 2016. *Cambridge Handbook of Morphology*, chapter Implicative Relations in Word-Based Morphological Systems. Cambridge University Press, Cambridge.

Adam Albright. 2002. *The identification of bases in morphological paradigms*. Doctoral dissertation, University of California Los Angeles.

Laura Benua. 1997. *Transderivational Identity: Phonological Relations between Words*. Doctoral dissertation, University of Massachusetts.

James Blevins. 2006. Word-based Morphology. *Journal of Linguistics*, 42(3):531–573.

James Blevins. 2008. Declension Classes in Estonian. *Linguistica Uralica*, 44(4):241–267.

James Blevins. 2016. *Word and Paradigm Morphology*. Oxford University Press.

Luigi Burzio. 1996. *Current Trends in Phonology: Models and Methods*, chapter Surface Constraints versus Underlying Representation. European Studies Research Institute, University of Salford Publications.

Luigi Burzio. 1999. Surface-to-Surface Morphology: when your Representations turn into Constraints. In *Paper presented at the 1999 Maryland Mayfest, University of Maryland, College Park.*

Yu-Ying Chuang, Kaidi Lõo, James Blevins, and Harald Baayen. 2019. Estonian case inflection made simple. A case study in Word and Paradigm morphology with Linear Discriminative Learning. PsyArXiv, May 6, 1-40.

Ryan Cotterell, C. Kirov, and John Sylak-Glassman. 2017. Neural graphical models over strings for principal parts morphological paradigm completion. In *EACL*.

Ryan Cotterell, Christo Kirov, John Sylak-Glassman, David Yarowsky, Jason Eisner, and Mans Hulden. 2016. The SIGMORPHON 2016 shared Task — Morphological reinflection. In *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 10–22, Berlin, Germany. Association for Computational Linguistics.

Eesti Keele Institut. 2020. Basic Estonian Dictionary. (Institute of the Estonian Language) www.eki.ee/dict/psv/.

Noam Faust and Paul Smolensky. 2017. Activity as an alternative to autosegmental association. In *Manchester Phonology Meeting*, Manchester, United Kingdom.

Raphael Finkel and Gregory Stump. 2007. Principal parts and morphological typology. *Morphology*, 17(1):39–75.

Morris Halle and Alec Marantz. 1994. Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics*, 21:275–288.

Bruce Hayes and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379 – 440.

Robert Malouf. 2016. Generating morphological paradigms with a recurrent neural network. *San Diego Linguistic Papers*, 6:122–129.

Alan Prince and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical report, Rutgers University Center for Cognitive Science and Computer Science Department, University of Colorado at Boulder.

RGCL. 2018. Morphologically disambiguated corpus. Research Group of Computational Linguistics, University of Tartu, www.cl.ut.ee/korpused/morfkorpus/.

Eric Rosen. 2019. Learning complex inflectional paradigms through blended gradient inputs. In *Proceedings of the Society for Computation in Linguistics*, volume 2.

Imanol Schlag. 2019. TP-Transformer. https://github.com/ischlag/TP-Transformer.

Imanol Schlag, Paul Smolensky, Roland Fernandez, Nebojsa Jojic, Jurgen Schmidhuber, and Jianfeng Gao. 2019. Enhancing the transformer with explicit relational encoding for math problem solving. In *Proceedings of the thirty-third Conference on Neural Information Processing Systems (NeurIPS)*, volume 33.

Andrea D. Sims and Jeff Parker. 2016. How inflection class systems work: On the informativity of implicative structure. *Word Structure*, 9(2):215 – 239.

Paul Smolensky and Matt Goldrick. 2016. Gradient Symbolic Representations in Grammar: The case of French Liaison. Rutgers Optimality Archive 1552, Rutgers University.

Paul Smolensky, Matthew Goldrick, and 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.

Paul F. Tupper, Paul Smolensky, and Pyeong Whan Cho. 2018. Discrete symbolic optimization and Boltzmann sampling by continuous neural dynamics: Gradient Symbolic Computation. *CoRR*.

Appendix A: Detailed calculations of Harmonies in table 7

Parameter	value
Max	0.76
Dep	−0.44
Correspondence	1.37
*td$]_\sigma$	−0.51
a_e	0.621
a_i	0.562
a_t	0.564
a_v	0.648
$a_\emptyset$	0.526
a_ϕ	0.517

Table 13: Learned parameter values (repeated from table 5)

Cand.	Nom. sg.	Gen. sg.	Part. sg.	Nom. pl.	Gen. pl.	Part. pl.	H
(a)		e	t	ed	te	i	7.78
(b)		i	t	id	te	e	7.72

Table 14: Calculated Harmonies of relevant segments (repeated from table 7)

	Max	Dep	Corresp	Harmony
Candidate (a)				
e + V (gen.sg.)	0.76		1.37	2.13
t (part.sg.)	0.43	−.19	1.37	1.61
e (nom.pl.)	0.47	−.17	1.37	1.67
t (gen.pl.)	0.43	−.19	1.37	1.61
i + ϕ (part.pl.)	0.76			0.76
Total				7.78
Candidate (b)				
i + V (gen.sg.)	0.76		1.37	2.13
t (part.sg.)	0.43	−.19	1.37	1.61
i (nom.pl.)	0.43	−.19	1.37	1.61
t (gen.pl.)	0.43	−.19	1.37	1.61
e + ϕ (part.pl.)	0.76			0.76
Total				7.72

Table 15: Calculation of Harmonies for relevant segments