

Evidence for Gradient Input Features from Sino-Japanese Compound Accent

Eric Rosen

Johns Hopkins University

1 Introduction

Predicting the pitch-accent patterns of two-member Sino-Japanese compounds presents an analytic challenge (Kawahara, 2015:460). Although both morphemes (henceforth M_1 and M_2), show general accenting tendencies, I show, using a corpus of compounds from the NHK Accent Dictionary, that deriving their accent patterns without lexically listing each compound accent requires gradient feature values, as provided by the Gradient Symbolic Computation framework. (Smolensky & Goldrick, 2016) (henceforth GSC) Both prosody and morphological identity affect accent in 2-morpheme structures but even with morpheme-specific constraints, both Optimality Theory (Prince & Smolensky, 1993) and Harmonic Grammar (Pater, 2009) fail to provide an explanatory analysis. The problem is that the tendency of a morpheme to trigger accent on its neighbour operates differently left-to-right than right-to-left, as shown by the contrast between accented 社会 *syá-kai* ‘society’ (lit. company-meet) and unaccented 会社 *kai-sya* ‘company’, with morpheme order switched. Prosody alone cannot explain this contrast, given the abundance of pairs like unaccented *ha-tyoo* 波長 ‘wavelength’ and accented *tyóo-ha* 長波 ‘long-wave’ with the opposite accent-prosody correlation.

A GSC analysis succeeds because the formalism naturally affords three locations for accent-affecting propensity: an underlyingly mora-associated position and floating positions on the left and right edges. A simple machine-learning algorithm finds *accent-affecting propensities* = *activations* that collectively work for a set of compounds with frequently-occurring morphemes from the NHK corpus. This analysis provides evidence that gradient input representations are needed to explain these kinds of phenomena.

The paper is organized as follows. §2 discusses how the data that was used was developed and some issues that arose with this set of data. §3 introduces Sino-Japanese compounds and some of their relevant properties. §4 introduces the GSC framework. §5 explains how accent feature coalescence can account for the combined effects of M_1 and M_2 on accent. §6 discusses prosodic constraints that determine Japanese word accent that are proposed by Ito & Mester (2016) and how weighted versions of these constraints are adopted for the present analysis. §7 outlines the kinds of inputs that are proposed for this analysis. §8 discusses how the interaction of Faithfulness constraints with prosodic constraints can account for the accent patterns in the data. §9 gives examples of pairs of compounds with the same two morphemes but switched order and differing accentuation. Here I show why the GSC framework can capture these data but discrete inputs with lexically-indexed constraints cannot, even if constraints are weighted. §10 gives details of the learning algorithm that was tested through computer simulation in order to discover input values and constraint weights that can account for the data. §11 explains how cross-validation was used to test how well a learner might predict unseen forms from exposure to partial data. §12 discusses how some unwanted types of coalescence might be ruled out. §13 concludes.

2 Notes on the data

For data, I started with an .xml file that was taken from the NHK Accent Dictionary (Nippon Hoosoo Kyokai 1998 [Japanese Broadcasting Corporation]) which contains over 116,000 entries, many of which are duplicates. Only a subset of these are Sino-Japanese compounds. Some hand-processing was necessary to filter out both foreign borrowings and Yamato (native) words, which in some cases can be hard to

* Thanks to Paul Smolensky, Matt Goldrick, Michael Becker, Najoung Kim, Matthias Lalisce and Tom McCoy for helpful discussion. Research generously funded by NSF INSPIRE grant BCS-1344269. All errors are my own.

distinguish from Sino-Japanese words without detailed knowledge of the sub-lexica. Automated processing of the original set was done in order to exclude very obscure entries. Only compounds whose constituent morphemes both occurred at least 50 times in the corpus were included, but this number ended up being less than 50 in many cases because of compounds that appeared as duplicates.

Some compounds are listed with multiple possible accent patterns, in order of preference. It is not clear whether this optional patterning means that a given speaker may sometimes choose one pattern and other times another or whether it represents idiolectal differences among a population of speakers of Standard Japanese. To account for the possibility of more than one pattern for a given compound was considered to be beyond the scope of this paper, so in most cases I abstracted away from this issue by either excluding ambiguously accented compounds or else choosing only the most preferred pattern. A goal for further research is to explore optionality of accent in more detail. The set that were tested by the algorithm numbered 1348 compounds. A handful of further compounds were added as a separate set in order to illustrate the effects of morpheme order switches such as those in §9.

There was a small and slightly shifting group of about a dozen compounds that kept showing up as outliers when the algorithm attempted to find activation values and constraint weights that would correctly predict all the compounds in the dataset. These compounds tended to push the learned values in the wrong direction and it was found that it was best to remove these from the set as outliers so that the algorithm would not consider them. Instead, their number was added to that of those that were not successfully learned. Some examples of this set are *nyuu-doo* 入道 ‘entering the priesthood’ (lit. enter-way), *nin-ki* 人氣 ‘popularity’ (lit. person-spirit) and *aku-doo* 悪道 ‘wickedness’ (lit. ‘evil-way’). It is also possible that some of the unexplained irregularity is due to polysemy among a set of what looks like the same morpheme: for example *doo* 道 ‘way’ may have a different semantic sense in *nyuu-doo* 入道 ‘entering the priesthood’ than in *aku-doo* 悪道 ‘wickedness’.

3 Accentuation patterns of two-member Sino-Japanese compounds

Sino-Japanese compounds, which originate from words borrowed from the Chinese, are well recognized as forming a component of the lexicon that exhibits different phonological patterns than other strata. (See, for example, Itô & Mester (1995), who add that “[t]hey are mostly bound forms and occur only compounded with other Sino-Japanese roots.”) Because of their distinct patterns of phonological behaviour, I shall treat them separately in this paper from other Japanese compounds. (See Rosen (2018) for an analysis of gradient patterns in Yamato (native Japanese) compound accent.)

Each single-character morpheme is bimoraic at most, e.g. *kái-gai* 海外 ‘overseas’ (lit. ocean-outside), with each morpheme represented by one Sino-Japanese character. Morphemes show gradient accenting tendencies, as shown below in (1): e.g., *hon* 本 ‘this; main; book’, is accent-friendly, accenting as M₁ in 15/24 compounds in the dataset: e.g. *hón-poo* 本法 ‘this law’, but fails to trigger in *hon-ryuu* 本流 ‘mainstream’. In contrast, *sin* 新 ‘new’, fails to accent in 13/16 compounds: e.g. *sin-poo* 新法 ‘new-law’, but does accent in *sin-pei* 新兵 ‘new-recruit’.

<i>M</i> ₁ accent-friendly	<i>M</i> ₁ accent-resisting
<i>hon</i> 本 ‘this; ‘main’; ‘book’ △△△△△△△△△△△△△△----- 15/24	<i>sin</i> 新 ‘new’ △△△----- 3/16
△ = accented compound	- = unaccented compound
<i>hón-poo</i> 本法 (accented) ‘this law’	<i>sin-poo</i> 新法 (unaccented) ‘new law’
Contrary behaviour	
<i>hon-ryuu</i> 本流 (unaccented) ‘main-stream’	<i>sín-peí</i> 新兵 (accented) ‘new recruit’

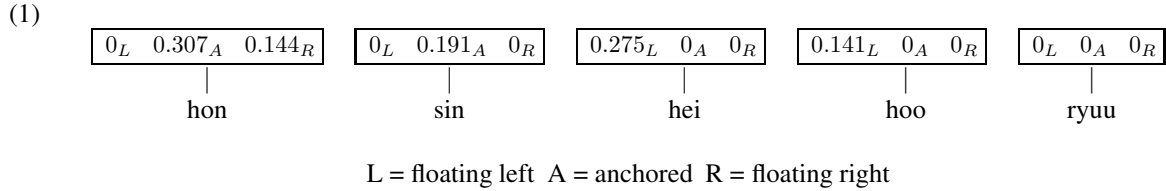
Figure 1: Gradient accenting behaviour of morphemes

I show that accent is determined both by prosody (Ito & Mester, 2016) (henceforth I&M) and combined accenting tendencies of M_1 and M_2 . In the prosodically identical ((H) (H)) and morphologically minimal pairs in (1), with contrasting accentuation shown by shading, accent cannot be determined by M_1 alone or M_2 alone: *hon* 本 ‘main’, *hoo* 法 ‘law’ and *sin* 新 ‘new’ all variably affect accenting.

4 Gradient Symbolic Computation and gradient inputs

In order to capture the observed gradient accenting behaviour of compound elements, I adopt Gradient Symbolic Computation (Smolensky & Goldrick, 2016), a framework that allows *partially-activated* input features and weighted constraints. As shall be shown below, when two accent features coalesce in the output, their effective input activation is the sum of the two activations. This allows *accenting propensities* to be expressed by *input activations*. 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, 2018; Smolensky et al., 2014; Smolensky & Goldrick, 2016; van Hell et al., 2016; Zimmermann, 2017a,b, 2018).

The following example shows gradient accent features for the five example lexemes found in the compounds in table 1 that resulted from a computer-simulated learning algorithm. In the proposed account, input accent features can occur either anchored to a mora or floating at the left or right edge of a morpheme.



A floating accent feature at edge X of one morpheme can coalesce with an anchored feature on a neighbouring morpheme on side X , where $X \in \{\text{left, right}\}$, but no other positions can coalesce. I rule out these other kinds of potential coalescence with the following highly-weighted constraints:

- (2) DEP-PATH-ACCENT: For every accent feature a_i with a path p_j to a mora m_k in the output, there is a corresponding accent feature a'_i with a path p'_j to mora m'_k in the input, such that $a_i \Re a'_i$, $p_j \Re p'_j$ and $m_k \Re m'_k$.

“An accent in the output must have a non-floating correspondent in the input.”

- (3) LINEARITY (strict) If features f_1 and f_2 are elements of morpheme m in the input and $f_1 \prec f_2$ in the input, then $f'_1 \prec f'_2$ in the output, where $f_1 \Re f'_1$ and $f_2 \Re f'_2$.

“A floating feature can coalesce with an anchored feature on an adjacent but not the same morpheme.” (strict LINEARITY in the domain of the morpheme)

- (4) PATHINTEGRITY: If accent feature f' has a path p'_i to mora m'_j in the output, then f has no path to a mora in the input other than path p_i to mora m_j , where $f \Re f'$, $p_i \Re p'_i$ and $m_j \Re m'_j$.

“An anchored feature cannot coalesce with another anchored feature. That is, an accent feature with a path to a mora in the output cannot have more than one path to a mora in the input.”

For simplicity, I only show candidates that respect strongly-weighted constraints requiring the leftmost syllable to be footed (I&M’s INITFT) and disallowing a Foot to span a morpheme boundary except in the case of a minimal word of two light syllables such as 和紙 *wa-si* ‘Japanese paper’.

5 How feature coalescence interacts with constraints in GSC

In GSC, a MAX constraint with weight w_M rewards the occurrence in the output of an input with activation a_i with positive Harmony $w_M \cdot a_i$. A DEP constraint with weight w_D costs negative Harmony equal to the difference between full output activation and input activation, times the weight of the constraint: $w_D \cdot (1 - a_i)$. When two input accent features, a_i and a_j , coalesce in the output and surface with full activation, the Harmonic reward from MAX is the weight of MAX times the sum of the two input activations: $w_M \cdot (a_i + a_j)$. The Harmonic penalty from DEP will be the weight of DEP times the difference between full activation of 1 and the sum of the two input activations: $w_D \cdot (1 - \max(a_i + a_j, 1))$. The winning candidate has the highest Harmony value H.

The combined effect of weighted MAX and DEP with prosodic constraints (see below) determine an epiphenomenal threshold of activation that an (aggregate) input must surpass in order to surface. If a candidate with no accent has greater Harmony than other candidates, an unaccented compound will surface.

The GSC framework is based on a sub-symbolic level (not discussed in detail here, but see Smolensky et al. (2014)), which is a platform not just for optimization but also for **quantization**, which, through a highly-weighted constraint, penalizes non-discrete outputs and favours outputs with full activation. Quantized, discrete output activations percolate to the symbolic level, which contains symbolic descriptions that are familiar from symbol-based linguistic theory. I assume here a highly-weighted quantization constraint that strongly favours discrete outputs. As a result, tableaux presented below will only consider candidates with full or zero activation of features in the output.

6 The effect of prosody on accent

Following their OT analysis, I adopt the following constraints proposed by I&M (but here with weighted versions) to account for prosodic effects, which in the present analysis, combine with lexical effects to determine pitch accent. Constraint (5) is adopted directly from I&M, (6) is adopted by I&M from McCarthy & Prince (1993:81), and constraints (7) and (8) are adopted by I&M from Prince & Smolensky (1993:45).

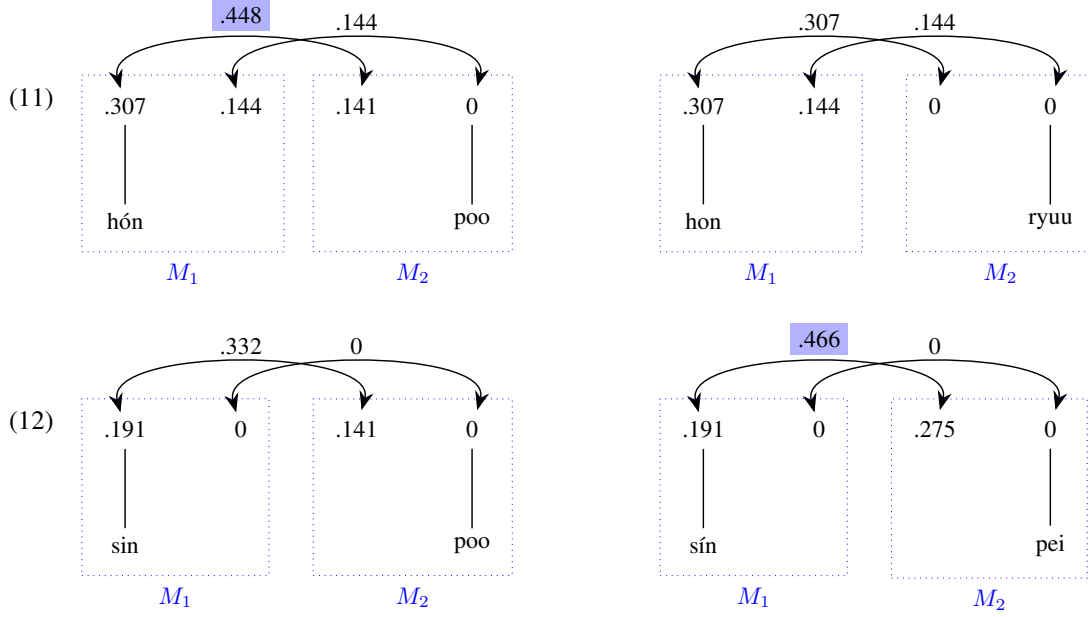
- (5) RIGHTMOST: Violated by any foot following the head foot within the prosodic word.
- (6) INITIALFOOT: A prosodic word begins with a foot (Itô & Mester, 1992:31), Violated by any prosodic word whose left edge is aligned not with the left edge of a foot, but of an unfooted syllable.
- (7) NONFINALITYFOOT: Violated by any head foot that is final in its PrWd ‘– final’ in the sense that the right edge of FT’ coincides with the right edge of PrWd.
- (8) PARSE- σ : All syllables are parsed into feet.

I also propose the following two novel constraints that award positive Harmony to candidates that respect them.

- (9) WORDACCENT: A prosodic word contains a prominence peak.
- (10) PREJUNCTURAL: A compound juncture is immediately preceded by a pitch accent. (See Kawahara (2015)).

7 Proposed inputs

I posit underlying accent features with gradient activation that are anchored to moras or float at the left and/or right morpheme edge. The following diagrams show how learned activations coalesce to result in aggregate inputs that correctly derive the accentuation patterns of the examples that were shown above in (1). Arcs indicate coalescence of a floating feature with an anchored feature and the number at the top of an arc indicates the aggregate input activation from the two sources. Aggregate activations that result in accentuation are shaded in blue.



8 Interactions of Faithfulness with prosodic constraints

The following tableaux show derivations of the four compounds in (1) with feature activations and constraint weights that were learned by the algorithm.

Apart from the effects of input activations, the net effects of the last 4 prosodic constraints give accent on M_1 a Harmony advantage of 0.027 over accent on M_2 .

($w_{\text{Prejunc}} - w_{\text{Parse}} = -0.063$ for M_1 .)

($w_{\text{NonFin}} = -0.090$ for M_2 .)

(13)

hon+hoo	MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
ㄣ (hón)-poo	0.492	-0.498		-0.181	0.118	0.098		0.029
$0.307_A + 0.141_L$								
(hón)-(poo)	0.492	-0.498	-0.183		0.118	0.098		0.027
$0.307_A + 0.141_L$								
(hon)-(póo)	0.158	-0.772				0.098	-0.090	-0.606
$0.144_R + 0_A$								
(hon)-(poo)								0

In contrast to accented *hón-poo* ‘this-law’ in (13) above, compound *sin-poo* ‘new-law’ in (14) below, does not accent, because of lower anchored input activation on *sin* ‘new’ (0.191 vs. 0.307).

(14)

sin+hoo	MAX	DEP	WDACC	RMOST	PARSE	PRJNC	NONFIN	H
(sín)-poo	0.365	-0.603		-0.181	0.118	0.098		-0.203
$0.191_A + 0.141_L$								
(sín)-(poo)	0.365	-0.603	-0.183		0.118	0.098		-0.205
$0.191_A + 0.141_L$								
(sin)-(póo)	0	-0.902				0.098	-0.090	-0.894
$0_R + 0_A$								
ㄣ (sin)-(poo)								0

Compound *hon-ryuu* ‘main-stream’ in (15) does not accent, in contrast to *hón-poo* ‘this-law’ in (13) above, because of lack of input activation on *ryuu* ‘flow; et al’.

(15)		MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
	hon+ryuu	+1.098	-0.902	-0.183	-0.181	+0.118	+0.098	-0.026	
	(hón)-ryuu $0.307_A + 0_L$	0.337	-0.625		-0.181	0.118	0.098		-0.253
	(hón)-(ryuu) $0.307_A + 0_L$	0.337	-0.625	-0.183		0.118	0.098		-0.255
	(hon)-(ryúu) $0.144_R 0_A$	0.158	-0.772				0.098	-0.026	-0.606
	𠬞 (hon)-(ryuu)								0

Compound *sín-pei* ‘new-recruit’ in (16) does accent, in contrast to unaccented *sin-poo* ‘new-law’ in (14) above, because of higher left floating input activation on *hei* ‘soldier’ (0.275 vs. 0.141).

(16)		MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
	sin+hei	+1.098	-0.902	-0.183	-0.181	+0.118	+0.098	-0.090	
	𠬞 (sín)-pei $0.191_A + 0.275_L$	0.512	-0.482		-0.181	0.118	0.098		0.065
	(sín)-(pei) $0.191_A + 0.275_L$	0.512	-0.482	-0.183		0.118	0.098		0.063
	(sin)-(péi) $0_R + 0_A$	0	-0.902				0.098	-0.090	-0.894
	(sin)-(pei)								0

9 Change in accentuation when morpheme order switches: not explainable in OT or HG, with lexically-indexed constraints

Recall again the examples given in §1, in which pairs of compounds that have reversed morpheme order exhibit contrasting accentuation.

- | | | |
|------|--|---|
| (17) | (a) 字数 <i>zi-súu</i> ‘# of written characters’
ACCENTED
(LH) | (b) 数字 <i>suu-zi</i> ‘numeral’
UNACCENTED
(HL) |
| | (c) 波长 <i>ha-tyoo</i> ‘wavelength’
UNACCENTED
(LH) | (d) 长波 <i>tyóo-ha</i> ‘long-wave’
ACCENTED
(HL) |

Prosody cannot explain this contrast, given the opposite correlation between prosody in accent across the two pairs.)

HG (e.g. Pater (2009)) or OT (Prince & Smolensky, 1993) with lexically-indexed constraints¹ (Pater, 2000) is insensitive to switching the morpheme order, *unless* edge-aligned floating features and coalescence occur, but this then leads to the following problems.

In OT or HG, for accented *zi-súu* 字数 ‘number of written characters’, there must be some underlying accent feature somewhere since no accent surfaces in prosodically identical (light-heavy) *ha-tyoo* 波长 ‘wavelength’. Yet accent doesn’t surface on *suu-zi* 数字 ‘numeral’, in which the morphemes are reversed from in *zi-súu*. And if HG were to try to explain the latter contrast through a weighting of constraints that favours accent in a final heavy syllable, then the contrast between *ha-tyoo* and *tyóo-ha* would be unexplained.

Any ranking or weighting of MAX-ACC_{zi}, MAX-ACC_{suu}, *ACC_{zi} and *ACC_{suu} that derives accented *zi-súu* should predict *suu-zi* to also be accented. The only possible resort would be coalescence of anchored and floating features, as in the present GSC account, to distinguish the left from the right edge of a morpheme. Consider how we might try to make this work.

Suppose there is a floating accent feature on the right edge of *zi* ‘character’ as follows:

- (18) ACC ACC
 |
 zi

¹ See Round (2017) for an argument against indexing constraints to morphs rather than phonological elements.

If there is the same kind of Path Integrity constraint that was proposed above that prevents an anchored feature from migrating to coalesce with another anchored feature, that right floating feature on *zi* can coalesce onto *suu* to cause accenting on *suu* with the ranking below, but not vice versa in *suu-zi*, where no coalescence can occur. (Here, we interpret a constraint like MAX-ACC_{zi} to mean that an accent on morpheme *zi* in the input must surface *somewhere*, but not necessarily on *zi* itself.)

$$(19) \quad *ACC_{zi} \gg \text{MAX-ACC}_{zi} \gg *ACC_{suu} \gg \text{MAX-ACC}_{suu}$$

$$(20) \quad \begin{array}{ccc} \text{ACC} & \text{ACC} & \longrightarrow \text{ACC} \\ | & & | \\ zi & & suu \end{array} \quad \begin{array}{ccc} \text{ACC} & \cancel{\longrightarrow} & \text{ACC} \quad \text{ACC} \\ | & & | \\ suu & & zi \end{array}$$

If constraints are weighted rather than categorically ranked:

$*ACC_{zi} > \text{MAX-ACC}_{zi}$ is needed to prevent accent on *zi*.

$*ACC_{suu} > \text{MAX-ACC}_{suu}$ is needed to prevent *suu* from accenting in *suu-zi*.

But notice that there cannot be any leftward floating feature on *zi*. If there were, the ranking above would incorrectly derive accent on $*s\acute{u}u-zi$, since it would be able to coalesce with the anchored accent feature on *suu* and the ranking $\text{MAX-ACC}_{zi} > *ACC_{suu}$ would force accentuation on the morpheme *suu*. A lexically indexed MAX-ACC_X on morpheme *X* means that any accent in the input on *X* needs to have a correspondent in the output, but it does not necessarily need to surface with a path to a mora that is on morpheme *X*.

This causes a problem because now, having no leftward floating accent feature on *zi* prevents it from triggering accent on an M_1 when *zi* is an M_2 :

$zy\acute{u}u-zi$ 十字 (‘ten’ + ‘character’) ‘cross’, (accented), requires the ranking $\text{MAX-ACC}_{zyuu} \gg *ACC_{zyuu}$ with an underlying accent on *zyuu* since *zi*, with no left floating feature, is unable to help accent surface on it. If constraints are weighted, it must be that $\text{MAX-ACC}_{zyuu} > *ACC_{zyuu}$. $*ACC_{zi} > \text{MAX-ACC}_{zi}$, will, as before, prevent accenting on *zi*. The constraints on *zyuu* and *zi* cannot interact if there is no possible coalescence.

It then should become impossible for a compound with *zyuu* ‘ten’ as M_1 to be unaccented, since whatever the input form or indexed constraint ranking of M_2 , the compounds should surface with accent on *zyuu*. This is contradicted by examples like unaccented *zyuu-moku* 十目 ‘all eyes’ (unaccented) (lit. ten-eye).

GSC, which allows gradient activations, is able to derive all four compounds, as shown in the tableaux below, since different activations can occur on different accent features: floating at L and/or R edges (shown by *L* and *R* subscripts) and μ -anchored (*A* subscript).

(21) Learned input accent activations for *suu*, *zi*, *zyuu* and *moku*

0.092 0.156	0.317	0.359 0.261	0
<i>zi</i>	<i>suu</i>	<i>zyuu</i>	<i>moku</i>

Unlike the case of OT or HG in which input representations must have discrete values of 0 or 1, gradient values in GSC allow a weak floating accent activation at the left edge of *zi*, which allows it to accent variably. OT/HG needs to prevent any floating feature on the left of *zi* so that *suu-zi* doesn’t accent.

(22)

	MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
suu+zi	+1.098	−0.902	−0.183	−0.181	+0.118	+0.098	−0.090	
(súu)-zi $0.317_A + 0.092_L$	0.449	−0.533		−0.181	0.118	0.098		−0.049
(súu)-(zi) $0.397_A + 0.092_L$	0.449	−0.533	−0.183		0.118	0.098		−0.051
(suu)-(zǐ) $0_R + 0_A$	0	−0.902				0.098	−0.090	−0.894
ㄌㄞ (suu)-(zǐ)								0

(23)

	MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
zyuu+zi	+1.098	-0.902	-0.183	-0.181	+0.118	+0.098	-0.090	
ㄹㄹ (zyúu)-zi $0.359_A + 0.092_L$	0.495	-0.495		-0.181	0.118	0.098		0.035
(zyúu)-(zi) $0.359_A + 0.092_L$	0.495	-0.495	-0.183		0.118	0.098		0.033
(zyuu)-(zí) $0.261_R + 0_A$	0.287	-0.667				0.098	-0.090	-0.372
(zyuu)-(zi)								0

As shown in (22) and (23) above, *zyuu* accents before *zi* but *suu* doesn't, because anchored activation on *zyuu* of 0.359 exceeds that of 0.317 on *suu*.

As shown in (24) below, *zyuu-moku* doesn't accent because *moku* has no activation to add to the activation on *zyuu* to bring it above the threshold.

(24)

	MAX	DEP	WDACC	PARSE	PRJNC	RMOST	NONFIN	H
zyuu+moku	+1.098	-0.902	-0.183	-0.181	+0.118	+0.098	-0.090	
(zyúu)-moku $0.359_A + 0_L$	0.395	-0.578		-0.181	0.118	0.098		-0.149
(zyúu)-(moku) $0.359_A + 0_L$	0.395	-0.578	-0.183		0.118	0.098		-0.151
(zyuu)-(móku) $0.261_R + 0_A$	0.287	-0.667				0.098	-0.090	-0.372
ㄹㄹ (zyuu)-(moku)								0

zi-súu in (25) below accents unlike *suu-zi* in (22) above because the floating input activation on the right side of *zi* is greater than on the left (0.156 vs. 0.092).

(25)

	MAX	DEP	RMOST	PARSE	PRJNC	WDACC	NONFIN	H
zi+suu	+1.098	-0.902	-0.183	-0.181	+0.118	+0.098	-0.090	
(zí)-suu $0_A + 0_L$	0	-0.902		-0.181	0.118	0.098		-0.867
(zí)-(suu) $0_A + 0_L$	0	-0.902	-0.183		0.118	0.098		-0.869
ㄹㄹ (zi)-(súu) $0.156_R + 0.317_A$	0.519	-0.475				0.098	-0.090	0.052
(zi)-(suu)								0

10 Learning Algorithm for constraint weights and accent activations

The following is pseudocode for a computer-simulated algorithm that learned constraint weights and input activations that correctly predicted accent patterns of a set of compounds.

```

 $\eta = 0.03$  ▷ Stepsize for changing activations and weights
for  $m_i \in \text{stems}$  do
  for  $a_{ij} \in \text{accent activations}$  do
     $a_{ij} \leftarrow 0.5$  ▷ Initialize accent activations on each stem at 0.5.
  end for
end for
 $w_{Max} = w_{Dep} \leftarrow 1$ 
for  $c \in \text{constraints}, c \neq Max, c \neq Dep$  do
   $w_c \leftarrow 0$ 
end for
loop
   $errors \leftarrow 0$ 
  for  $w_k \in \text{compound words}$  do
    for  $w_{kl} \in \text{output candidates}$  do
       $H_{w_{kl}} = \sum_c H_{w_{kl}|c}$  ▷ Harmony for each candidate is sum of Harmonies resulting from each
    end for
    if  $\arg \max H_{w_{kl}}$  has the actual accent pattern then
      continue
    else
       $errors \leftarrow errors + 1$ 
      Increase by stepsize  $\eta$  all activations and constraint weights that favour the desired winner.
      Decrease by stepsize  $\eta$ , if they stay above 0, all activations and constraint weights that favour
      the false winner.
    end if
  end for
  if  $errors == 0$  then
    break ▷ Learning completed.
  end if
   $\eta = \eta \times 0.96$  ▷ Decay the stepsize.
  for  $a_{ij} \in \text{accent activations}$  do
    if  $a_{ij}$  did not need adjustment in this loop iteration then
       $a_{ij} \leftarrow 0$  ▷ Aim to maximize number of zero activations.
    end if
  end for
end loop

```

11 Cross-validation

In order to test how well a learner could predict unseen forms after exposure to partial data, I ran cross-validation on the dataset as follows. Compounds were chosen randomly from the dataset one at a time. If both M_1 , M_2 and the prosodic pattern of this compound occurred elsewhere in the dataset, that compound was added to the test set until 300 compounds had been sampled. The residue was retained as a training set. Over ten randomized runs, the average accuracy of prediction on the test set was 85.8%.

12 Ruling out some types of association of accent features

A question that might come to mind about this analysis is, what prevents a floating accent feature on the left of a morpheme from associating to a morpheme on its right, and vice versa, if the other accent features do not surface?

Given the fact that ordering of the constituents in the input makes a difference, both semantically and with respect to accent, I take constituents of a compound to be ordered in the input. We still do not lose the ability to rule out coalescence of features from the same morpheme through strict linearity if we make the morpheme the domain of a strong LINEARITY constraint and the prosodic word the domain of a weaker LINEARITY constraint.

The following constraint, highly weighted, would prevent A_1 from coalescing on M_2 below, even if A_2 , A_3 and A_4 were not to surface: “If A_i and A'_i stand in correspondence, and A_i is left-aligned with morpheme M_j in the input, then A'_i is left-aligned with morpheme M_j in the output.” The same with right alignment.

If A_1 , which is floating to the left of another accent on M_1 , associates with the second morpheme, it will no longer be left-aligned with M_1 if we consider both the accentual tier and moraic tier at once and measure alignment there. Whether or not A_4 is there in the input does not matter. If A_3 coalesces, it is right-aligned with M_1 in both the input and the output so it satisfies the constraint.

(26)

$$\begin{array}{l} \text{input: } \left[\begin{array}{ccc} A_1 & A_2 & A_3 \\ & | & \\ & \mu_1 & M_1 \end{array} \right] \left[\begin{array}{ccc} A_4 & A_5 & A_6 \\ & | & \\ & \mu_2 & M_2 \end{array} \right] \\ \\ \text{output: } \left[\begin{array}{ccc} & & \\ M_1 & \mu_1 & M_2 \end{array} \right] \left[\begin{array}{c} A_{3,5,*1} \\ | \\ \mu_2 \end{array} \right]_{M_1} \end{array}$$

13 Conclusion

The foregoing analysis shows that the pitch accent patterns of two-member Sino-Japanese words exhibit subtle, gradient tendencies, where the surfacing of accent depends not just on prosodic factors but is lexically influenced by underlying properties of both M_1 and M_2 , where neither of M_1 or M_2 alone can single-handedly determine where and if accent surfaces. The facts of Sino-Japanese pitch accent provide support for models such as GSC with partially-activated input features, given the argument in §9 above that models that only allow discrete inputs and lexically-indexed constraints cannot capture all of the data. A further question to be explored with respect to the issue of gradiently-activated features vs. lexically-indexed constraints is of what predictions each model makes about separate phonological processes that operate in tandem. This question is raised by Zimmermann (2018), who finds that in Moses Columbian Salish, multiple processes (vowel deletion and stress assignment) show parallel exceptional behaviour. In her handout she comments that if one tried to explain these effects through lexically-indexed constraints, “[i]t is a coincidence that at least two different constraints are indexed to the same class of (exceptional) morphemes.”

A further step in comparing GSC to other frameworks will be to find other examples across languages in which multiple gradient effects occur in parallel, and to compare how both GSC and competing models would handle these kinds of phenomena.

References

- 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.
- 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.
- 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.
- Itô, Junko & Armin Mester (1992). Weak layering and word binarity. *A New Century of Phonology and Phonological Theory. A Festschrift for Professor Shosuke Haraguchi on the Occasion of his Sixtieth Birthday*, Kaitakusha, 26–65.

- Itô, Junko & Armin Mester (1995). The core-periphery structure of the lexicon and constraints on reranking. *University of Massachusetts Occasional Papers in Linguistics* 18, 181–209.
- Ito, Junko & Armin Mester (2016). Unaccentedness in Japanese. *Linguistic Inquiry* 47:3, 471–526.
- Kawahara, Shigeto (2015). *Handbook of Japanese phonetics and phonology*, De Gruyter Mouton: Berlin, Boston, chap. The phonology of Japanese accent, 445–492.
- McCarthy, John & Alan Prince (1993). Generalized alignment. *Yearbook of morphology* 79–153.
- Müller, Gereon (2017). Gradient symbolic representations in syntax. Handout, IGRA 02: Syntax II.
- Pater, Joe (2000). Non-uniformity in English secondary stress: the role of ranked and lexically-specific constraints. *Phonology* 17:2, 237–274.
- Pater, Joe (2009). Weighted Constraints in Generative Linguistics. *Cognitive Science* 33, 999–1035.
- Prince, Alan & Paul Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Tech. rep., Rutgers University Center for Cognitive Science and Computer Science Department, University of Colorado at Boulder.
- 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 (2018). Predicting semi-regular patterns in morphologically complex words. *Linguistics Vanguard* 4:1. ROA 1339.
- Round, Erich R. (2017). *On looking into words (and beyond)*, Language Science Press: Berlin, chap. Phonological exceptionality is localized to phonological elements: the argument from learnability and Yidiny word-final deletion, 59–98.
- Smolensky, Paul & Matthew 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.
- 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 (2018). Gradient Symbolic Representations in the output: A case study from Moses Columbian Salishan stress. *Proceedings of NELS*, vol. 48.