February 11, 2019

Born to be gradient: a case study from compound tensing in Korean

Hyunjung Lee
Universität Leipzig
hyunjung.lee@uni-leipzig.de

1 Overview

Morphemes sometimes exhibit an exceptional pattern which cannot be explained by the purely phonological generalizations: one morpheme in a language triggers or undergoes a certain process whereas another morpheme fails to trigger or undergoes the same process, even though the two morphemes cannot be distinguisible from phonological factors. One way for formalizing exceptions is lexically indexed constraints (e.g., Ito and Mester (1990), Pater (2009), Finley (2009)) where every single morpheme in a language can be indexed to different constraints. Lexically-indexed rankings is another method for treating exceptional morphemes differently by allowing those morphemes to select their own sub-rankings (e.g., Orgun (1996), Anttila (2002), Inkelas et al. (2004)). Or the unexpected behavior of certain morphemes is not exceptional anymore if we assume the enriched representations that list this information in the lexicon (e.g., Bye (2007)). In this paper, I will argue for one of the representational approaches, Gradient Symbolic Representations, proposed by Smolensky and Goldrick (2016) with a case study of irregular failure of compound tensing of Korean. The core assumption of this framework is that all phonological elements can have a gradient degree of presence in an underlying structures. The intrinsic property of this system allows us to derive the unexpected compounds where this phonological process fails to apply, not by listing all the combinations of two conjucts of the compounds.

This paper is organized as follows: the section 2 shows the compound tensing data in Korean. I will focus on gradient preference and dispreference pattern of post-obstruent tensification in noun-noun compounds depending on the combination of two conjuncts of compounds. The core point of the data is that this exceptionality cannot be straightforwardly captured in current systems. In section 3, I will breifly introduce the background framework of Gradient Symbolic Representations, and provide an account for the unexpected triggers/undergoers of compound tensding with the partially activated feature in each edge of the nouns. In section 4, I will present an error-driven learning algorithm which also supports that this gradient continum of different degree of activities are learnable. The section 5 concludes the paper.

2 Data

2.1 Gradient inclinations of the process

When a compound consist of two nouns, W_A and W_B, initial plain obstruents of W_B undergo junctural processes including obstruent tensification Inkelas and Cho (1994), Kim

(2001) and among others).[1] As shown in (9), the lenis obstruent of W_B /p/ undergoes tensification into [p']. As can be seen in (9), this applies an initial plain segment of compounds, irrespective of phonological quality of a final segment of the first conjunct.

(1) Compound Tensing

(a)	/hɛ/ + /pic/	→	[hɛ.**p'**it]	post vowel
	'sun' + 'light'		'sunlight'	
(b)	/isɨl/ + /pi/	→	[i.s.ɨl.**p'**i]	post lateral
	'dew' + 'rain'		drizzle	
(c)	/pom/ + /pi/	→	[pom.**p'**i]	post nasal
	'spring' + 'rain'		'spring rain	
(d)	/pok/ + /pi/	→	[pok.**p'**i]	post obstruent
	'luck' + 'cost'		'brokerage'	

Compound Tensing in Korean, however, presents exceptional non-undergoing processes among roughly 23% noun-noun compounds (Ito, 2014, Kim, 2016, Jun, 2018, **?**).

(2) Undergoers and Non-undergoers of Compound Tensing

A.	Regular patterns		
(a)	/pipim/ + /pap/	→	[pi.pim.**p'**ap]
	'mix' + 'rice'		'rice with mixed vegetables'
(b)	/namu/ + /pɛ/	→	[na.mu.**p'**ɛ]
	'wood' + 'boat'		'a wooden boat'
B.	Exceptions		
(a)	/pipim/ + /kuksu/	→	[pi.pim.**k**uk.s'u]
	'mix' + 'noodle'		'noodle with mixed vegetables'
(b)	/komu/ + /pɛ/	→	[ko.mu.**p**ɛ]
	'rubber' + 'boat'		'an inflatable boat'

Not all compounding process triggers tensification, even when they met the environment for applying the rule, and the occurrence or non-occurence of compound tensing is not phonological predictable. Certain compounds, such as [pi.pim.kuk.s'u] and [ko.mu.pɛ] shown in (2)-B, unexpectedly do not undergo post-obsturent tensification, whereas other compounds undergo the same process in the predicted way with the same morphemes /pipim/ and /pɛ/.

Most previous accounts assume a single juncture marker which appears between the two nouns in accordance with compounding process. [2] [3] Besides, other various phonological elements have been proposed as the internal markers (e.g., /t/, or floating feature [+cor], [+tense] and others), none of them cannot capture the essential phonological distinction between undergoers/triggers and non-undergoers/triggers of compound tensing. Other previous studies (Im, 1981, Sohn, 1999, Kang, 1989, Ha, 2006) argue that the unexpected unoccurrence of compound tensing is determined by non-phonological factors in terms of their morpho-syntactic, semantic and etymological properties, only to conclude that there is no way to generalize this exceptionality.

Each noun has different inclination or declination of undergoing compound tensing, as

[1] Korean has a three-way distinction in terms of laryngeal contrast in obstruents: voicelss plain, aspirated, and tense obstruents. e.g., [pul] 'fire', [p^hul] 'grass', [p'ul] 'horn' (Cho, 2011)

[2] The most prevalent assumption on the juncture marker is /s/ in an underlying structure, which triggers tensification, and preceded by coda cluster simplification (e.g., /pipim/ + /s/ +/pap/ → /pi.pims.**p'**ap/ → [pi.pi**m**.p'ap]).

[3] In Middle Korean /s/ is assumed to be morphologized as an a genitive marker which shows up with the compound juncture process.

given in (3). For example, the first conjunct /hɛ/ is more likely to trigger compound tensing than other conjuncts /pipim/ and /koŋ/, but the same pattern can be observed from the second conjuncts W_B.

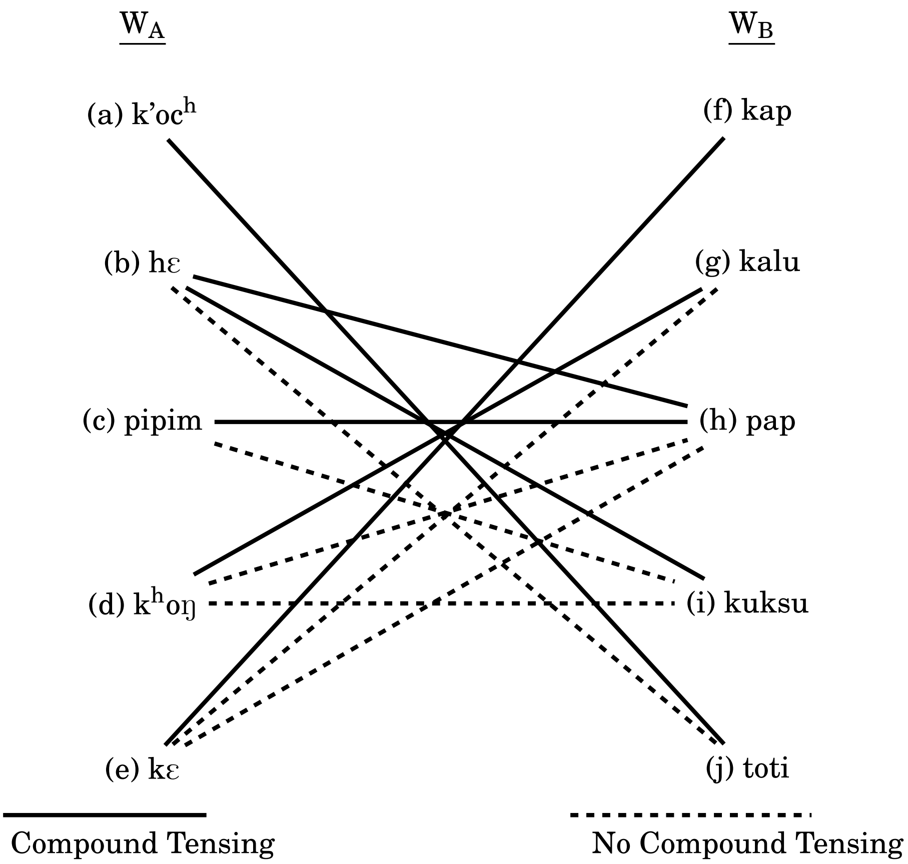
(3)

	W_A	+ W_B	compounds
(a)	/hɛ/	+ /pap/	→ [hɛ.**p’**ap]
(b)	/hɛ/	+ /kuksu/	→ [hɛ.**k’**uks’u]
(c)	hɛ/	+ /kalu/	→ [hɛ.**k**a.lu]
(d)	/pipim/	+ /pap/	→ [pi.pim.**p’**ap]
(e)	/pipim/	+ /kuksu/	→ [pi.pim.**k**uk.s’u]
(f)	/pipim/	+ /kalu/	→ [pi.pim.**k**a.lu]
(g)	/koŋ/	+ /pap/	→ [koŋ.**p**ap]
(h)	/koŋ/	+ /kuksu/	→ [koŋ.**k**uk.s’u]
(i)	/koŋ/	+ /kalu/	→ [koŋ.**k**a.lu]

	W_A	+ W_B	compounds
(a)	/hɛ/	+ /pap/	→ [hɛ.**p’**ap]
(b)	/pipim/	+ /pap/	→ [pi.pim.**p’**ap]
(c)	/koŋ/	+ /pap/	→ [koŋ.**p**ap]
(d)	/hɛ/	+ /kuksu/	→ [hɛ.**k’**uks’u]
(e)	/pipim/	+ /kuksu/	→ [pi.pim.**k**uk.s’u]
(f)	/koŋ/	+ /kuksu/	→ [koŋ.**k**uk.s’u]
(g)	/hɛ/	+ /kalu/	→ [hɛ.**k**a.lu]
(h)	/pipim/	+ /kalu/	→ [pi.pim.**k**a.lu]
(i)	/koŋ/	+ /kalu/	→ [koŋ.**k**a.lu]

Having a closer look at the pattern, the compound tensing exhibits continuum of gradient preferences depending on both the first conjuncts W_As and the second conjuncts W_Bs in the compound. That is, the one side of the gradient preference that each conjunct exhibits is not enough to predict the exceptional cases, but rather the interaction of two conjuncts. For instance, /pipim/ seems to undergo compound tensing with the second conjunct /pap/, but not with /kuksu/. In contrast, the second conjunct /pap/ fails to tensify its initial plain segment /p/ when it consists of a compound with /koŋ/, but /kuksu/ can show the predicted tensified case with the first conjunct /hɛ/.

(4) The continuum of gradient preferences of compound tensing



The entire pattern of this continuum of gradient preferences of compound tensing is illustrated in (4). A solid line between two nouns shows the combinations of two conjuncts which undergo compound tensing and a dashed line presents the noun-noun compounds which exceptionally does not undergo compound tensing.

This pose a challenge for generative theories with rules or categorical constraints like standard rule-based (Chomsky and Halle, 1968) or Optimality theory frameworks (Prince and Smolensky, 1993) These systems only allow features to be binary or privative, and this gradient continuum of irregularity should be dealt with by separately listing whole compounds that unexpectedly fail to undergo the process of compound tensing unexpected compounds in the lexicons (Bye, 2007), by postulating additional subgrammar to refer to all the possible combinations of two conjuncts that fail to undergo this process (Inkelas and Zoll, 2006) or specifying indexed-constraints which refers to speicifc morphemes (Finley, 2009, Pater, 2009).

A question that arises is whether this scalar pattern of exceptionality should be dealt with the grammar or through lexicalization. In this work, I argue for an account in terms of Gradient Symbolic Representations (Smolensky and Goldrick, 2016, Rosen, 2016), where elements can bear different degrees of activity in the input. This assumption allows us to understand the nature of data and to derive exceptions successfully, which is impossible with other systems.

3 Analysis through Gradient Harmonic Grammar

3.1 Gradient Symbolic Representation

Gradient Symbolic Representations (Smolensky and Goldrick, 2016) assumes two layered underlying structures: a symbolic level where and sub-sumbolic level. The discrete phonological elements can be associated with gradient strengths of presence in an underlying structure. These different activation levels are represented as numerical values from 0 to 1, which is the full degree of presence. In an surface structure all output elements have the full activity of 1. It means that a certain input element associated with weak activity is neutralized to fully activated output and the different activities only affect the evaluation of faithfulness constraints.

At the symbolic level, the grammatical well-formedness of a structure of the discrete phonological elements is evaluated through Harmonic Grammar (e.g., Legendre et al. (1990), Potts et al. (2010)) where constraints are weighted, not ranked. In Harmonic Grammar, the Harmonic function $H(S)$ calculates thte structure S based on the weighted sum of S's violations of each constraint. The key prediction of this framework is that partially activated input features will interact with a corresponding constraints as much as the amount of the activities they bear. And the interaction of partially-weighted constraints will derive the different outputs, even though the evalutated structures are same in terms of the symbols. This sub-symbolic layer allows us to provide a general account for various patterns of lexical exceptions in the phonology. It can predict exceptional morphemes: exceptional morphemes in the underlying representation of a morpheme can be exceptionally too weak to undergo or trigger a certain phonological process, while the regular morphemes have enough strength to undergo or trigger the same process. There has been previous studies which assume Gradient Symbolic Representations: liaison consonants in French (Smolensky and Goldrick, 2016), semi-irregularity of Japanese Rendaku (Rosen, 2016), exceptional tonal allomorphy in San Miguel el Grande Mixtec (Zimmermann, 2018), phonological length of Raddoppiamento fonosintattico (Amato, 2018), and lexical accent in Lithuanian (Kushnir, 2019).

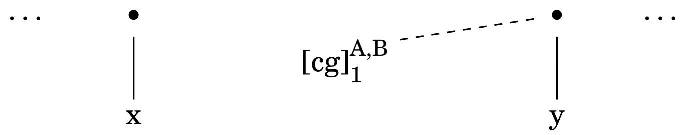
3.2 Proposal

In section 2, it is observed that both first and second conjunct contribute to derive regular and unexpected patterns of compond tensing. I argue for a generalized account for compound tensing in Korean through Gradient Harmonic Representations, where all the phonological elements are allowed to be associated with different activities of presence. I suggest that each edge of nouns in Korean may have a floating feature [constricted glottis] ([cg]) (Zoll, 1996) with different gradient activitities in the underlying structures, as illustrated in (5). Each edge of the W_As and W_Bs contains a floating feature [cg] that is paritally active, expressed by numerical activities τ, where $0 \leq \tau \leq 1$.

(5) An example of an underlying representation

Compound tensing occurs when two stem-specific, partially activated floating [cg] features are coalescenced and this coalescenced feature [cg] docks to the root node of the second conjunct W_B. In a surface structure the feature [cg] which is successfully associated to the root note will be realized with the full acitivity 1, as shown in (6).

(6)



3.3 Analysis

In this analysis the gradient nature of compound tensing, where the preference or dispreference for tensing of a given word is accounted for due to the activation strength of its underlying tensing feature [cg]. Harmonic Grammar computes the well-formedness of output candidates through the interaction between weighted constraints. The core constraints are given as follows:

- MAX[cg]: Input must have output correspondent of the feature [cg].[4]
- IDENT[cg]: The specification for the feature [cg] of an input segment must be preserved in its output correspondent.[5]
- UNIFORMITY[cg]: No feature [cg] in the output has multiple correspondents in the input.

The constraint MAX[cg] will create the positive harmony by interacting with the strengths of [cg] on the left edge of the W_As and the right edge of W_Bs. ; The more acitivity τ the feature [cg] bears in an underlying structure, the more rewards it induces when it realizes in an surface level. The other constraints are discrete. IDENT requires that correspondent root node to have identical values for the feature [cg]. UNIFORMITY penalizes coalescence.

[4]TheMAX constraint here is assumed to be able to refer to the roots nodes.

[5]Featural change is penalized by IDENT constraints, which require that correspondent segments have identical values for a given feature.

The total Harmony score of the structure will be evaluated as follows:

(7) $H(s) = 100 \cdot \mathbb{C}_{\text{Max[cg]}}(s) - 60 \cdot \mathbb{C}_{\text{Ident[cg]}}(s) - 10 \cdot \mathbb{C}_{\text{Uniformity[cg]}}(s)$

Only when the additive combination of the feature [cg] from two conjuncts exceeds a certain threshold which create the best harmony score does tensing occur.[6] The detailed assumption of 5-level of the gradient activities of the feature [cg] is as follows:

(8) A hierarchy of 5-level of activation values of W_A and W_B for compounding tensing

$[cg]_A$ \ $[cg]_B$	0.05	0.2	0.4	0.6	0.8
0.05	✗	✗	✗	✗	✓
0.2	✗	✗	✗	✓	✓
0.4	✗	✗	✓	✓	✓
0.6	✗	✓	✓	✓	✓
0.8	✓	✓	✓	✓	✓

The stronger activity τ W_A or W_B bears, the more likely compound tensing tensification occurs.

The first part of the analysis is derive the regular pattern of compound tensing with two conjuncts which contain the floating feature [cg] with paritally-active strengths underlyingly: W_A: /pipim/ - τ: 0.4, W_B: /pap/ - τ: 0.4.

(9) *T_1. /pipim/ + /pap/ → [pi.pim.**p'**ap]*

… m $[\text{cg}]^x_{0.4}$ $[\text{cg}]^y_{0.4}$ p …	MAX ([c.g]) $w = 100$	IDENT ([c.g]) $w = -60$	UNIFORMITY ([c.g]) $w = -10$	H
O_1: … m p …				0
☞ O_2: … m $[\text{cg}]^{x,y}_1$ p …	0.4 + 0.4	1	1	10
O_3: … m $[\text{cg}]^x_1$ $[\text{cg}]^y_1$ p …	0.4 + 0.4	1 + 1		−40
O_4: … m $[\text{cg}]^y_1$ p …	0.4	1		−20

The occurence of compound tensing is the result of two additive features from both two conjucnts as shown in tableaux (9). The faithful candidate O_2-T_1 is excluded because it neither gets the reward from MAX[cg] nor any penalty from IDENT[cg] or UNIFORMITY[cg]. There are other candidates which realize each floating feature [cg] at the surface level. Both candidates O_2-T_1 and O_3-T_1 receive the gradient reward (i.e., (0.4+0.4)·100= 40) from MAX[cg], as it realizes both additive features. Since the candidate O_2-T_1 changes the value of [cg] of the (fully activated) segment [p] on *pap* which is different from the one in the input, it gets a violation of IDENT (i.e., 1·-60= -60), while the candidate O_3-T_1 gets a violation (i.e.,

[6] The threshold here is an epiphenomenon of the gross effects of MAX on partially activated [cg] and the counter-acting effects of IDENT and UNIFORMITY.

(1+1)·-60= -120) twice bigger than O_2-T_1, as it also changes the value of [cg] of the final segment [m] on the first conjunct *pipim*. Due to the coalescence O_2-T_1 is also penalized by UNIFORMITY. The grammar prefers the candidate which realizes the feature [cg] from the both conjunct, otherwise it cannot get bigger reward from MAX. The sum of activation values of the [cg] feature on *pipim* and *pap* is strong enough to trigger the tensification by interacting with the weighted constraints.

The analysis of unexpected compounds is derived in the same way as in (9), but only difference lies on the specific degree of acitivities that two conjuncts bears: W_A: /pipim/ - τ: 0.4, W_B: /kuksu/ - τ: 0.2.

(10) *T_2. /pipim/ + /kuksu/ → [pi.pim.**k**uk.s'u]*

… • (m) $[\mathbf{cg}]^{x}_{0.4}$ $[\mathbf{cg}]^{y}_{0.2}$ • (k) …		MAX ([c.g]) $w = 100$	IDENT ([c.g]) $w = -60$	UNIFORMITY ([c.g]) $w = -10$	H
☞	O_1: … • (m) • (k) …				0
	O_2: … • (m) $[\mathbf{cg}]^{x,y}_{1}$ • (k) …	0.4 + 0.2	1	1	−10
	O_3: … • (m) $[\mathbf{cg}]^{x}_{1}$ $[\mathbf{cg}]^{y}_{1}$ • (p) …	0.4 + 0.2	1 + 1		−60

The total sum of the features [cg] on *pipim* and *kuks'u* in O_2-T_2 is too weak to receive a sufficient reward (i.e., (0.4+0.2)·100= 60) to trigger compound tensing. In contrast to the candidatate O_2-T_1, the grammar prefers to choose the faithful candidate O_1-T_2, since the penalties given from IDENT[cg] and UNIFORMITY[cg] are bigger than the reward from MAX[cg].

The exceptional pattern of compound tensing can be explained if we assume that each edge of the conjuncts (W_As and W_Bs) may bear a partially activated feature [cg], whose strength makes each conjunct more likely to be involved in tensification of the initial segment of the W_Bs. The gross effects of two gradiently activated features with the weighted constraints makes a variation of the junctural morpheme: the compounds with coalescencing [cg] and the others without it.

4 A Learning Algorithm

4.1 The Architecture

In this section, I propose a learning model based on Convolutional Neural Network (CNN) (Krizhevsky et al., 2012). The CNN with multiple output nodes produces scores for each label, and a threshold predictor generates a reference point using the scores of the labels. The threshold is then used to for the system to decide whether each label is as relevant or irrelevant. The CNN classifier consists of a two convolutional layer and a single softmax layer. Formally, let be the k-dimensional embedding vector corresponding to the activities of the right edge i of given W_As and the left edge j of given W_Bs. A compound x of the length n can be represented as a matrix $n \times k$:

(11) $x = w_i \oplus w_j$

where $\oplus$ is a concatenation operator. A convolutional operation involves a filter, which is applied to a window of the words. A feature c is generated by

(12) $c_i = f(m \cdot w_{i:i+k-1} + b)$

where is a hyperbolic tangent function f. A max-over-time pooling is operated to take the maximum value $\hat{c} = max\{c\}$ as a representative feature for the filter map. These features are concatenated into 'top-level' feature vector, by passing through to a fully connected softmax layer whose output is the probability distribution over the labels. Given the weights of the constraints, each compound in the trainset is generated and undergoes tensification only if it exceeds the threshold. An output is evaluated to the training set, and if the tensing occurs, then it receives a reward, otherwise, it is penalized.

4.2 Steps for the Learning Algorithm

Each steps of how the learning algorithm is trained as follows:

Step 1. Initialization

- Activation levels for [cg] of the W^A s and W^Bs were initialized at 0.5
- Constraints MAX and IDENT were initialized with unit values
- UNIFORMITY and LINEARITY have fixed values
- The threshold levels for the sum values of [cg] for compounds were set at 0.7

Step 2: Iteration

1. The compounds [W^A+ W^B] are evaluated on each iteration to check whether each gross effect of compound tensing is correctly derived;
   - will get a reward +10 if the correct pattern is derived,
   - will get a penalty -5 if the wrong pattern is derived
2. When two coalescing activations [cg] require adjusting,
   - It randomly refills the both values of [cg] by either decrementing or incrementing them (a stepsize of 0.05)
   - MAX and IDENT adjust their weights slightly adjusted through a simulated-annealing process (de Vicente et al., 2003) [7]

Step 3: Convergence

- After 16533 iterations (i.e., when the algorithm can predict all the training set data of compound tensing corectly) the training of this learning was converged.

[7] The process is proceeded with a decaying temperature T and random Gaussian noise N with $m = 0$ and $s.d. = 0.05$

4.3 Results

Results of running the simulation of the learning algorithm are illustrated in (13).

(13)

Results	
Average of iterations	32
Final Value of MAX	1.121
Final Value of IDENT	0.69
The number of activation levels for [cg] of W^A	5
The number of activation levels for [cg] of W^B	5

Each activity for underlying presence of the floating feature [cg] of both W^A and W^B are converged into 5-level, which is same as in the theoretical modeling.

5 Conclusion

In this paper, I have proposed the unified analysis of both regular and irregular pattern of post-obstruent tensficaiton of the noun-noun compounds in Korean. The main claim in this paper is that exceptionality can be dealt within the simple grammar with the enriched underlying representation, by avoiding the necessity of listing exceptional cases in the lexicon. The intrinsic property of the Gradient Symbolic Representations allows us to capture the scalar irregular pattern which is impossible with other systems. The continuum of presence and absence of compound tensing can be precisely accounted for by virtue of gradient strength of the features from both conjuncts (not from a single juncture marker) without any redundancy rules. This is a reminiscent of other gradient harmonic analyses on French liaison (Smolensky and Goldrick, 2016) and on Rendaku (Rosen, 2016), which also explain irregular patterns observed at the juncture of two words.

References

References

Amato, I. (2018). Length as strength: a new account of Raddoppiamento Fonosintattico. Poster, presented at NELS 2018, at Cornell University, October 05, 2018.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory*, 20(1):1–42.

Bye, P. (2007). Allomorphy selection, not optimization. In Blaho, S., Bye, P., and Krämer, M., editors, *Freedom of Analysis*, pages 63–91. Mouton.

Cho, Y.-m. Y. (2011). Laryngeal Contrast in Korean. In van Oostendorp, M., Ewen, C. J., Hume, E. V., and Rice, K., editors, *The Blackwell Companion to Phonology*, volume 5, pages 2662—2684. Wiley Online Library, Malden, MA.

Chomsky, N. and Halle, M. (1968). The sound pattern of English. New York: Harper & Row.

de Vicente, J., Lanchares, J., and Hermida, R. (2003). Placement by thermodynamic simulated annealing. *Physics Letters A*, 317(5-6):415–423.

Finley, S. (2009). Morphemic harmony as featural correspondence. *Lingua*, (119).

Ha, S.-K. (2006). *The Morphology and Phonology of 'Sai-sori' phenomena in Modern Korean*. PhD thesis, Seoul National University.

Im, H.-P. (1981). For the solution of sai-sios problem [Sai-sios mwunceyuy haekyelul wihaye]. *Kukehak*, 10:1–35.

Inkelas, S. and Cho, Y.-M. Y. (1994). Post-obstruent tensification in Korean and geminate inalterability. In Kim-Renaud, Y.-K., editor, *Theoretical issues in Korean linguistics*, volume 45. Center for the Study of Language.

Inkelas, S., Orgun, O., and Zoll, C. (2004). The implications of lexical exceptions for the nature of grammar. In *Optimality theory in phonology: A reader*, pages 542–551. Wiley Online Library.

Inkelas, S. and Zoll, C. (2006). Reduplication: Doubling in Morphology. In *Cambridge Studies in Linguistics*, volume 106. Cambridge University Press.

Ito, C. (2014). Compound tensification and laryngeal co-occurrence restrictions in Yanbian Korean. *Phonology*, 31(3):349–398.

Ito, J. and Mester, A. (1990). he structure of the phonological lexicon. In Tsujimura, N., editor, *The Handbook of Japanese Linguistics*, pages 62–100. Blackwell, Malden.

Jun, J. (2018). Morpho-phonological processes in Korean. In Aronoff, M., editor, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press, New York.

Kang, O.-m. (1989). Conditions on-/t/-Insertion in Korean Noun Compounds. Master's thesis, University of Washington.

Kim, G.-R. (2001). Obstruent alternations in sub-compounds. *Studies in Phonetics, Phonology and Morphology*, 7(2):315–331.

Kim, S. (2016). Phonological trends in Seoul Korean compound tensification. Master's thesis, Seoul National University.

Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In *NIPS*.

Kushnir, Y. (2019). *Prosodic Patterns in Lithuanian Morphology*. PhD thesis, Univeristy of Leipzig.

Legendre, G., Miyata, Y., and Smolensky, P. (1990). Harmonic grammar – a formal multilevel connectionisttheory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the 12th annual conference of thecognitive science society*, pages 388–395.

Orgun, C. O. (1996). *Sign-based morphology and phonology with special attention to Optimality Theory*. PhD thesis, UC Berkeley.

Pater, J. (2009). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Parker, S., editor, *Phonological Argumentation: Essays on Evidence and Motivation*, pages 123–138. Equinox, London.

Potts, C., Pater, J., Jesney, K., Bhatt, R., and Becker, M. (2010). Harmonic grammar with linearprogramming: From linear systems to linguistic typology. *Phonology*, pages 77–117.

Rosen, E. (2016). Predicting the unpredictable: Capturing the apparent semi-regularity of rendaku voicing in Japanese through Harmonic Grammar. In Clem, E., Dawson, V., Shen, A., Skilton, A. H., Bacon, G., Cheng, A., and Maier, E. H., editors, *Proceedings of BLS 42*, pages 235–249. Berkeley Linguistic Society, Berkeley.

Smolensky, P. and Goldrick, M. (2016). Gradient symbolic representations in grammar: The case of French Liaison. ROA 1286.

Sohn, H.-M. (1999). The Korean Language. In *Cambridge Language Surveys*. Cambridge University Press, Cambridge.

Zimmermann, E. (2018). Being Exceptional is Being Weak: Tonal Exceptions in San Miguel el Grande Mixtec. In Gallagher, G., Gouskova, M., and Yin, S. H., editors, *Proceedings of the 2017 Annual Meeting on Phonology*.

Zoll, C. (1996). *Parsing below the segment in a constraint-based framework*. PhD thesis, UC Berkeley.