Modeling super-gang effects in MaxEnt Grammar:

A case study on the nasal super-gang effect in Japanese Rendaku

Seoyoung Kim (Stony Brook University)

Introduction Contrary to strict-ranking OT, weighted constraints in Harmonic Grammar (HG) can easily capture *gang effects*. However, doubly-marked forms are often even worse than the mere addition of each violation (Albright 2012). This *super-gang effect* requires an extra penalty for double-markedness, such as weighted conjoined constraints (Shih 2017).

This study examines how to model super-gang effects in MaxEnt grammar. With a case study on Japanese Rendaku, it argues that nasal's super-gang effect is better captured if the degree of penalty is scaled up by the number of violations, through a power function.

Rendaku and Nasal Paradox In a Japanese noun-noun compound (W_A-W_B), if the initial onset of W_B is voiceless, it often voices (/yama/ + /kata/ → [yama**g**ata] 'mountain area'). Lyman's law allows only one voiced obstruent in a stem. Thus, if W_B already contains one, Rendaku never applies (/yama/ + /ka**j**i/ → [yama**kaj**i] 'mountain fire'), which has been captured as OCP(VOICE) (Itô and Mester 1986).

Japanese has post-nasal (PN) voicing, where voiceless obstruents voice immediately after a nasal (Itô et al. 1995; [tabe-**t**a] 'ate' vs. [ʃin-**d**a] 'died'), which was analyzed as the spreading of the nasal's [voice] to the following obstruent (Itô and Mester 1986). Given OCP(VOICE) and [voice] specification for nasals, a nasal in W_B is also expected to hinder Rendaku but it never does (/ori/ + /ka**m**i/ → [ori**gam**i] 'paper folding'). Japanese nasals behave as if they bear [voice] in PN voicing and as if they lack [voice] in Rendaku, which made paradoxical the specification of [voice] for nasal.

Corpus Study & Results I used *Spontaneous speech Corpus of Japanese* (Maekawa et al. 2000) to investigate phonological restrictions of Rendaku. I first extracted the cases where two Yamato (native Japanese) common nouns occur in a row. From there, I included items *iff* W_B underlyingly begins with a voiceless obstruent and this obstruent is pronounced as itself or as its voiced counterpart (N=2183, %Rendaku=$.29$ ($631/2183$)).

The effect of Lyman's law was confirmed; Rendaku is categorically blocked when W_B contains one or two voiced obstruents ($.001$, $1/598$) whereas 40% of compounds undergo Rendaku if W_B contains none ($630/1585$). There was no significant difference between compounds with zero and one nasal in W_B ($.3$ ($499/1663$) *vs.* $.28$ ($132/465$)), confirming the traditional description that one nasal never blocks Rendaku. However, when W_B contains more than one nasals, Rendaku never applied (0, $0/55$). This result supports Kumagai (2017) who firstly reported the Rendaku-blocking effect of multiple nasals in nonce words. Also, it shows that the multiple nasal effect is robust even in the existing words.

Super-gang effects One nasal does not stop Rendaku but two nasals entirely do. If a nasal is considered a Rendaku blocker, the effect of one is subtle whereas that of two is maximal; nasals not only gang up but super-gang up and completely stop Rendaku.

OCP and CC Jäger and Rosenbach (2006) suggested two kinds of cumulativity. *Ganging-up cumulativity* is when two separate low-ranked constraints jointly beat a high-ranked constraint. *Counting cumulativity* (CC) is when multiple violations of a single weaker constraint overpower a single violation of stronger constraint ($[C_2]^{\geq n}$ » C_1 » C_2).

The failure of one nasal to block Rendaku is not due to nasal's underspecification of [voice] but to the low ranking of OCP(VoiCon), which bans voiced consonants to co-occur. Moreover, the maximal blocking effect of multiple nasals comes from *CC* of OCP(VoiCon); a stem with two nasals and one Rendaku-generated voiced obstruent violates OCP(VoiCon) three times, which overpowers a Rendaku-triggering constraint ($[\text{OCP(VoiCon)}]^3$ » Rendaku » OCP(VoiCon)^1).

Similarity-driven blocking A single nasal endures a Rendaku-generated voiced obstruent whereas a single voiced obstruent or two nasals do not. The tolerance for Rendaku decreases if Rendaku results in a co-occurrence of more similar segments in W_B; the Rendaku application on /pene/ ([**b**e**n**e]) engenders a pair of consonants that agree in two features ([+voi][-cont]), a tolerable degree of similarity; /pege/ ([**b**e**g**e]) engenders three overlapping feature pairs ([+voi][-cont][-nas]) and /peneme/ ([**b**e**n**e**m**e]) even seven (([+voi]$\times 3$)([-cont]$\times 3$)[+nas]), an intolerably high similarity. This result corroborates earlier findings that Rendaku is a way to prevent or resolve a co-occurrence of similar segments (Kawahara and Sano 2014a,b and Sano 2013)

MaxEnt Constraints for Rendaku application are Rendaku ($w = 0$) and Id(Voice) ($w = .8$). Nasal's effect was captured by OCP(VoiCon) ($w = .4$). The weighted grammar could not reproduce the observed pattern: the Rendaku probability of items with one nasal in W_B was too low ($.23$ *vs.* $.28$) and that with two nasals too high ($.13$ *vs.* 0). The fundamental problem lies in the fact that the disparity between one and multiple violations is wide, which cannot be captured if OCP weight is fixed and violations are assessed linearly.

Power Function in MaxEnt A power function is a function of the form $f(x){=}x^c$ where the independent variable x is raised to a constant power c. I substitute the number of violation x with $f(x)$ to scale-up. Then I find a certain c value that best predicts the observed pattern. Japanese corpus data was accurately reproduced with the c value of 4 (one violation:$f(1){=}1^4$; three $f(3){=}3^4$); if one nasal in W_B ($.28$); if two nasals (0).

Conclusion I argued that super-gang effects emerging from multiple violations of a single constraint (CC) are not entirely captured in HG if violations are assessed linearly. I scaled the number of violations through a power function and attempted to capture nasal's super-gang effect in Rendaku. The model with a power function showed a greater explanatory power. I plan to test this constraint evaluation approach on other super-gang phenomena, expecting it to largely help modeling super-gang effects in HG.

References: **[1]** Albright, A. (2012). Additive markedness interactions in phonology. **[2]** Itô, J., Mester, A., and Padgett, J. (1995). Licensing and underspecification in optimality theory. *Linguistic inquiry*, pages 571–613. **[3]** Itô, J. and Mester, R.-A. (1986). The phonology of voicing in japanese: Theoretical consequences for morphological accessibility. *Linguistic inquiry*, pages 49–73. **[4]** Jäger, G. and Rosenbach, A. (2006). The winner takes it all—almost: Cumulativity in grammatical variation. *Linguistics*, 44(5):937–971. **[5]** Kawahara, S. and Sano, S.-i. (2014a). Identity avoidance and lyman's law. *Lingua*, 150:71–77. **[6]** Kawahara, S. and Sano, S.-i. (2014b). Identity avoidance and rendaku. In *Proceedings of the Annual Meetings on Phonology*. **[7]** Kumagai, G. (2017). Super-additivity of ocp-nasal effect on the applicability of rendaku. In *GLOW in Asia XI*. **[8]** Maekawa, K., Koiso, H., Furui, S., and Isahara, H. (2000). Spontaneous speech corpus of japanese. In *LREC*. **[9]** Sano, S.-i. (2013). Violable and inviolable ocp effects on linguistic changes: Evidence from verbal inflections in japanese. *Proceedings of FAJL*, 6:145–156. **[10]** Shih, S. S. (2017). Constraint conjunction in weighted probabilistic grammar. *Phonology*, 34(2):243–268.