# Featural Activity: a New Account of Compound Tensing

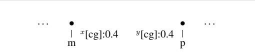
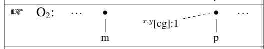
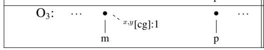
Hyunjung Lee (University of Leipzig)

**Nutshell** Compound Tensing (CT) in Korean exceptionally fails to undergo among 23% of Noun-Noun compounds (Jun 2001; Zuraw 2011; Ito 2014; Kim 2016). A question that arises is whether this exceptionality should be dealt with the grammar or through lexicalization. In this work, I argue for an account in terms of Gradient Symbolic Representations (GSR; 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. **Exceptionality in Compound Tensing** A noun-noun compound ($W_A$+$W_B$) undergoes obstruent tensification (Inkelas & Cho 1994; Kim-Renaud 1974, and among others), when the second conjunct $W_B$ begins with a lenis obstruent (e.g., /pok/ + /kuk/ → [pok.**k’**uk]). However, certain compounds (1a) - (1d) tensify $W_B$ in the predictable way, whereas others in (1e) - (1h) do not.

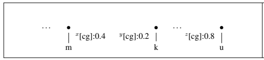
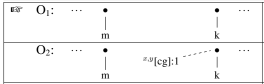
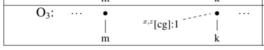
(1) a. /hɛ/ + /pap/ → [hɛ.**p’**ap] ∼ ([hɛt.**p’**ap]) e. /koŋ/ + /pap/ → [koŋ.**p**ap]
b. /hɛ/ + /kuks’u/ → [hɛ.**k’**uk.s’u] f. /koŋ/ + /kuks’u/ → [koŋ.**k**uk.s’u]
c. /pipim/ + /pap/ → [pi.pim.**p’**ap] g. /pipim/ + /kuks’u/ → [pi.pim.**k**uk.s’u]
d. /koŋ/ + /karu/ → [koŋ.**k’**a.ru] h. /hɛ/ + /toci/ → [hɛ.**t**o.ci]

Most previous accounts assume a juncture marker between the two nouns, which triggers tensification, preceded by coda cluster simplification (e.g., /pipim/ + /s/ +/pap/ → /pi.pims.**p’**ap/ → /pi.pi**m**.p’ap/). Although various phonological elements have been proposed as the internal marker (e.g., /s/, /t/, [+cor], [+tense] and others), there is no way to capture the essential phonological distinction between undergoers/triggers and non-undergoers/triggers of CT. For example, *hɛ* triggers CT with both *pap* and *kuks’u*, but not with *toci*, see (1a)-(1b) vs. (1h). The same puzzle can be seen in (1c) vs. (1g) and (1d) vs. (1e) - (1f). This gradient continuum of irregularity cannot be precisely captured by categorical constraints in Optimality Theory frameworks by assuming morpheme-specific phonological constraints (Pater, 2009; Finley, 2009) or different grammars (Inkelas & Zoll, 2005). Listing the exceptional non-undergoers in the lexicon (Bye, 2007) also fails to derive the correct surface forms, unless separately listing whole compounds that block the process of CT. **Theoretical Background** GSR states that phonological elements can have gradient degrees of strength, expressed as numerical activities varying from 0 to 1 in an underlying structure. Output elements have the full activity 1. GSR evaluates grammatical well-formedness through Harmonic Grammar (HG), where constraints are associated with weights, not ranked (Legendre et al., 1990). This grammar allows faithfulness constraints to interact with the partially activated input structures. **Analysis**: The exceptional pattern of CT can be explained if we assume that each conjunct ($W_A$ and $W_B$) may bear a partially activated feature [constricted glottis] ([cg]), whose strength makes each conjunct more likely to be involved in tensification. The gross effects of coalescencing two gradiently activated features ($[cg]_A$ + $[cg]_B$) will determine whether a floating feature [cg] is associated with the edge of $W_B$. The stronger activity $W_{A}/_{B}$ bears, the more likely tensification occurs. I suggest that the feature [cg] with 5 levels of the gradient activity (hypotheticallly given as $0.05 < 0.2 < 0.4 < 0.6 < 0.8$). HG computes the well-formedness of output candidates through the interaction between weighted constraints such as MAX[cg]. The MAX[cg] constraint gives rewards to the candidate in the proportion to the activity of the feature [cg] that makes it to the surface. Other constraints are discrete. IDENT requires that correspondent segments have identical values for the feature [cg]. UNIFORMITY penalizes coalescence. The following tableaux show how CT is determined by the effects of MAX on partially activated [cg] and the counter-acting effects of IDENT and UNIFORMITY. The CODACOND constraint penalizes every segment in coda position specified for laryngeal features such as [cg] and [sg]. Any coalescence that does not preserve the precedence relation in laryngeal tier will violate the LINEARITY (Pater 1999).

*$T_1$. /pipim/ + /pap/ → [pi.pim.**p'**ap]*

|  | MAX ([cg]) $w=1$ | IDENT ([cg]) $w=-0.6$ | UNIFORMITY ([cg]) $w=-0.1$ | CODACOND $w=-1$ | H |
|---|---|---|---|---|---|
| $O_1$:  | | | | | 0 |
| ☞ $O_2$:  | (0.4 + 0.4) | 1 | 1 | | 0.1 |
| $O_3$:  | (0.4 + 0.4) | 1 | 1 | 1 | −0.9 |

*$T_2$. /pipim/ + /kuks'u/ → [pi.pim.**k**uk.s'u]*

|  | MAX ([cg]) $w=1$ | IDENT ([cg]) $w=-0.6$ | UNIFORMITY ([cg]) $w=-0.1$ | LINEARITY $w=-1$ | H |
|---|---|---|---|---|---|
| ☞ $O_1$:  | | | | | 0 |
| $O_2$:  | (0.4 + 0.2) | 1 | 1 | | −0.1 |
| $O_3$:  | (0.4 + 0.8) | 1 | 1 | 1 | −0.5 |

$O_2$-$T_1$ receives the 'gradient' reward (=0.4+0.4) from MAX[cg], as it realizes both additive features. Since the value of [cg] of the (fully activated) segment *p* on *pap* is different from the one in the input, it gets a violation of IDENT. Due to the coalescence it is also penalized by UNIFORMITY. The activation value of the [cg] feature on *pipim* is strong enough to trigger the tensification by interacting with the strength of the [cg] on the conjunct *pap*. However, 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 (=0.4+0.2) for triggering tensification. In contrast to the candidatate $O_2$-$T_1$, IDENT[cg] and UNIFORMITY[cg] 'gang-up', so the optimal candidate will be the one without tensification. In other words, CT occurs only if the sum of the activations of the [cg] features on $W_A$ and $W_B$ exceeds the threshold for tensing (i.e. the sum of penalty when tensing occurs). Other possible outputs of coalescence cannot be the optimal candidate; $O_3$-$T_1$ is ruled out by CODACOND. Although $O_3$-$T_2$ gets a reward (=0.4+0.8) more than $O_2$-$T_2$ by coalescencing with much stronger [cg] , it induces a fatal violation of LINEARITY.

| (2) $[cg]_A$ / $[cg]_B$ | 0.05 | 0.2 | 0.4 | 0.6 | 0.8 |
|---|---|---|---|---|---|
| 0.05 | ✗ | ✗ | ✗ | ✗ | ✓ |
| 0.2 | ✗ | ✗ | ✗ | ✓ | ✓ |
| 0.4 | ✗ | ✗ | ✓ | ✓ | ✓ |
| 0.6 | ✗ | ✓ | ✓ | ✓ | ✓ |
| 0.8 | ✓ | ✓ | ✓ | ✓ | ✓ |

This 5-level system of activation values allows us to straightforwardly capture why a certain morpheme tends to undergo CT with the majority of cases, while others show a gradient dispreference for tensification, as shown in (2).

**Learnability** A learning algorithm also demonstrates that this scalar grammar is learnable. The algorithm is built through Convolutional Neural Network (Krizhevsky et al. 2012) with 2 hidden and 1 softmax layer. Each activation level of the conjuncts was set to zero, and could be randomly filled with the numerical activity from 0 to 1. Given the weights of the constraints, each compound in the database is generated and undergoes tensification only if it exceeds the threshold. The output is evaluated to the training set, and (i) if tensing occurs, then it receives a reward, (ii) otherwise, it is penalized. After 14323 iterations the training converged on 5 levels of activity of the feature.

**Conclusion** This analysis provides a unified account of CT by virtue of gradient strength of the features from both conjuncts (not from a single juncture marker). This is a reminiscence of GSR analyses on Liaison and on Rendaku, which also explain irregular patterns observed at the juncture of two words. The intrinsic property of GSR captures the nature of gradient inclination for CT and reforms the division of labor by not postulating a different layer of grammar and avoiding the necessity of listing exceptional cases in the lexicon.