

A gradient blend analysis of English PP verbal dependents

Najoung Kim[†], Kyle Rawlins[†], Paul Smolensky^{†,‡}

[†]Department of Cognitive Science, Johns Hopkins University [‡]Microsoft Research AI
{n.kim, kgr, smolensky}@jhu.edu

The idea that the argument-modifier (or complement-adjunct) dichotomy is problematic is by no means novel in the field of linguistics. In Schütze’s words [12], “argumenthood is not an all-or-nothing phenomenon, [...] it comes in degrees”. Although the distinction between complements and adjuncts seems to have psychological reality [15] and the two concepts are assigned distinct status in many linguistic formalisms [3, 4, 16], researchers have struggled to pinpoint a set of syntactic or semantic criteria for a clear-cut categorization. Many diagnostic tests have been proposed, but none provide necessary or sufficient conditions for complement- or adjuncthood; thus, they are taken only as tendencies rather than deterministic evidence.

Frequently, different diagnostic tests produce conflicting results. To illustrate, let us take two common diagnostic tests OMISSIBILITY (O) and PSEUDO-CLEFT (P). For example, the PP in the following sentence is not omissible (*O), suggesting that it is a complement:

- (1) *Steve pelted Anna [with acorns].* (2) **Steve pelted Anna.*

However, pseudo-clefting the PP is perfectly acceptable (P), implying it is an adjunct:

- (3) *Steve pelted Anna [with acorns].* (4) *What Steve did [with acorns] was pelt Anna.*

Such counterexamples go in both directions; both [*O,P] and [O,*P] are attested. The previous case is [*O,P], whereas *The man strangled the victims into a coma* is [O,*P]:

- (5) *The man strangled the victims.* [O]
(6) **What the man did [into a coma] was strangle the victims.* [*P]

This conflict pointing to the potentially gradient (or at least non-trivial) nature of complement and adjunct status has long been acknowledged [8, 9], but has not as often been formally characterized.

We argue that a **gradient blend** analysis [2, 6, 13] gives a principled explanation to such conflicts, and emphasize that gradience alone is insufficient to capture the observed pattern of conflicts. Formalizing Dowty [5]’s *dual* analysis approach to PP complements and adjuncts, we hypothesize that every English PP is a weighted mixture of *proto-complement* and *proto-adjunct* structures (Fig. 1). These proto-structures bear properties of canonical complements or canonical adjuncts. According to Dowty, the adjunct structure denotes the inherent locative meaning of *to* whereas the complement structure denotes the possibly idiosyncratic, predicate-dependent interpretation of *to*. Under our analysis, the activation strength of each structure determines how strongly complement-like and adjunct-like behaviors manifest. This is distinct from a probabilistic view where one structure is chosen over the other probabilistically [10]; our approach maintains the two structures (i.e., both components of the blend) simultaneously, with different activation values. Having such simultaneously active gradient *blends* enables us to capture all possible patterns of diagnostic test results, when a gradience-only analysis would fail to cover every possible case (compare Figs. 2 and 3).

We claim that different diagnostic tests yield conflicting results because they (1) target different thresholds of activation (*gradience*), and (2) probe different parts of the blend (*blendedness*). For instance, not being omissible is a test for activation of the complement

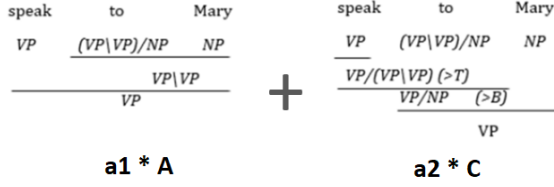


Figure 1: A gradient blend analysis of *speak to Mary* as two simultaneously active proto-structures with different activation values (a_1 and a_2 are the respective degrees of activation; note that it is not necessarily true that $a_1 + a_2 = 1$). The representations are adapted from Dowty’s CCG analysis of adjuncts (left) and complements (right).

structure, whereas being pseudo-cleftable is for activation of the adjunct structure; a non-gradient analysis would predict that only $[*O, *P]$ and $[O, P]$ are possible.

To support our analysis, we collect judgments about PPs across various thematic roles and apparently varying degrees of complement-adjuncthood via crowdsourcing. Participants are asked to give a score on a 7-point Likert scale for the *centrality* (with respect to the main verb) of a given NP under PP, inspired by a previously published protocol [11] that has been shown to effectively replicate linguists’ complement-adjunct judgments. Although the expression *centrality* was used to describe the task in the instructions, the participants had to go through a training procedure prior to the actual study where they were exposed to clear-cut argument- or adjunct-like examples. In order to justify our methodology even further, We also conducted a separate validation study with the same wordings but using relatively uncontroversial argument/adjunct stimuli only. The result of this validation study showed high agreement (88%) with linguist judgments even with minimal training and no explicit definitions of complements/adjuncts given. Moreover, substantial overlap with the thematic hierarchy proposed by [1] is found in the ordering of thematic roles by centrality. For the data collection, 305 stimulus sentences are generated from VerbNet [7] examples, and for each sentence, two diagnostic tests O and P are performed by a linguist. The main findings are (1) the complement- and adjunct-likeness does form a continuum rather than a dichotomy according to the distribution of scores, and (2) the result strongly suggests that the judgment patterns are only adequately analyzable by the gradient blend model and not by either a dichotomous or a gradience-only model (the argument against the latter is also an argument against a probabilistic model in which the probability of two structures sum to one; a one-dimensional scale does not suffice). From the two diagnostic test results, all four possible patterns were observed (supporting Fig. 3 over 2). We hypothesized that (1) centrality is proportional to the ratio of C and A activations (a_C/a_A), and thus higher centrality is entailed by both higher C activation and lower A activation, and (2) $*O$ implies high C and P implies high A, with different thresholds of “high”. The predictions from these hypotheses are borne out by the mean centrality scores of sentences in each judgment pattern group: $[*O, *P] = 0.249$, $[O, *P] = 0.164$, $[*O, P] = 0.099$, $[O, P] = -0.122$. The mean of $[O, P]$ is significantly different from all groups ($p < 0.05$; Kolmogorov-Smirnoff test), except from $[*O, P]$ where $p = 0.059$. A percentile analysis (Fig. 4) shows that each group has a significantly different distribution compared to the null hypothesis (uniformly distributed across quar-

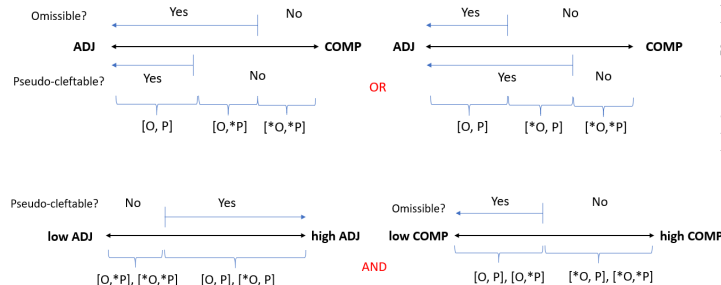


Figure 2: Possible patterns of diagnostic results under gradience-only analysis. Passing a test corresponds to having an activation below a test-particular threshold. (Impossible: $[*O, P]$ (left; Ex. (1–4)), $[O, *P]$ (right; Ex. (5–6))

Figure 3: Possible patterns under gradience *with* blendedness analysis. The (proto-)Comp and Adj structures with respective activation exist simultaneously but do not affect each other.

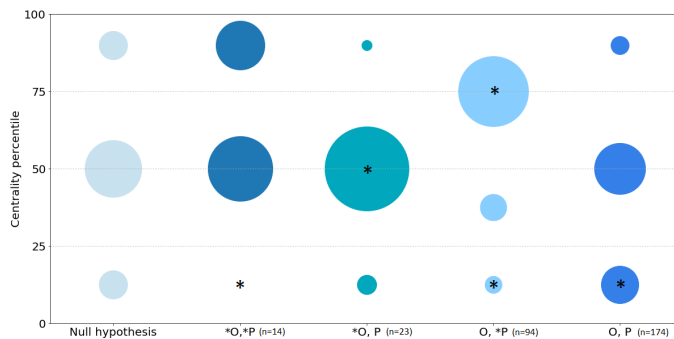


Figure 4: A percentile analysis of the centrality score ranking across the whole dataset. The size of the circles indicate the proportion of the sentences in each group that pertain to the corresponding percentile, and * indicates significant difference ($p < 0.05$) from the null hypothesis (far left).

tiles) in different regions ($p < 0.05$; binomial test with Holm-Sidak correction for multiple comparisons).

There remains a more fundamental question of why speakers need to maintain both structures simultaneously instead of converging to one. An insight into this matter can be drawn from [5] where it is suggested that the inherent meaning of the adjunct form serves as a mnemonic to effectively learning and retrieving the complement meaning with the same form. In other words, it is not coincidental that the ‘directional’ preposition *to* is used in expressions such as speak *to* (‘communicate’) or drink *to* (‘toast’) over other prepositions like *in* or *on*. This division of labor is made more formally explicit by our simultaneously active proto-structures, and how information from both sides are utilized can be expressed via constraint interactions using a gradient variant of Harmonic Grammar [14].

References [1] Baker 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20.4, 513–553. [2] Brehm & Goldrick 2017. Distinguishing discrete and gradient category structure in language: Insights from verb-particle constructions. *J. of Experimental Psychology* 43.10, 1537–1556. [3] Bresnan et al. 2015. Lexical-functional syntax. John Wiley & Sons. [4] Chomsky 1993. Lectures on government and binding: The Pisa lectures. Walter de Gruyter. [5] Dowty 2003. The dual analysis of adjuncts/complements in categorial grammar. In *Modifying adjuncts*. Walter de Gruyter. [6] Goldrick et al. 2016. Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition* 19.5, 857–876. [7] Kipper-Schuler 2005. VerbNet: A broad-coverage, comprehensive verb lexicon. Ph.D. Thesis, UPenn. [8] Koenig et al. 2003. Arguments for adjuncts. *Cognition* 89.2, 67–103. [9] Lang et al. 2003. Modifying adjuncts. Walter de Gruyter. [10] Manning 2003. Probabilistic syntax. In *Probabilistic linguistics*, 289–341. [11] Rissman et al. 2015. Using instruments to understand argument structure: Evidence for gradient representation. *Cognition* 142, 266–290. [12] Schütze 1995. PP attachment and argumenthood. In *MIT working papers in linguistics* 26, 95–151. [13] Smolensky et al. 2014. Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. *Cognitive Science* 38.6, 1102–1138. [14] Smolensky & Legendre 2006. The harmonic mind. MIT Press. [15] Tutunjian & Boland 2008. Do we need a distinction between arguments and adjuncts? Evidence from psycholinguistic studies of comprehension. *Language and Linguistics Compass* 2.4, 631–646. [16] Venneman & Harlow 1977. Categorical grammar and consistent basic VX serialization. *Theoretical linguistics* 4.1–3, 227–254.