KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

Coordination of Unlike Categories Creates Grammaticality Illusion*

Nayoun Kim (Sungkyunkwan University) · **Jiayi Lu** (Stanford University)

ABSTRACT

Kim, Nayoun and Jiayi Lu. 2024. Coordination of unlike categories creates grammaticality illusion. *Korean Journal of English Language and Linguistics* 24, 52-61.

Coordinate structures with mismatched DP and CP conjuncts can occupy the complement position of a preposition (e.g., "You can depend on my friend and that she will be on time."). Such examples pose a challenge to previously proposed syntactic generalizations including the Law of Coordination of Likes, and the Wasow's Generalization. Attempting to resolve this problem, recent proposals analyze such DP&CP coordinate structures as underlying DP&DP coordination structures, where the CP conjuncts are contained inside silent nominal shells. Under such proposals, the DP&CP coordination does not violate the Law of Coordination of Likes or the Wasow's Generalization. In this study, we present psycholinguistic evidence against such analyses, and suggest an alternative analysis where the DP&CP coordination is indeed ungrammatical but appears acceptable due to a grammaticality illusion.

KEYWORDS

coordination, syntactic category, DP and CP coordination, grammaticality illusion, experimental syntax

1. Introduction

Law of the coordination of Likes (henceforth LCL; Williams 1981) states that only constituents of the identical syntactic category can be coordinated. Ungrammaticality arises when elements of different categories are conjoined, as shown in (1).

(1)   a. Sue ate [DP sushi] and [DP burger] yesterday.
        b. *Sue ate [DP sushi] and [CP that burger is fresh].

However, exceptions to the LCL have also been noticed in the past literature. For example, in (2), a DP can conjoin with a PP to form a coordinate structure.

(2)   I consider that [DP a rude remark] and [PP in very bad taste].

<div align="right">(Sag et al. 1985, p. 118)</div>

Such exceptions can be captured in a weakened version of the LCL known as the Wasow's Generalization, inspired by Chomsky (1957)'s remark in (3).

(3)   If we have two sentences Z X W and Z Y W, and if X and Y are actually constituents of these sentences, we can generally form a new sentence Z X *and* Y W.   (Chomsky 1957, p. 35)

The Wasow's generalization (formulated in Pullum and Zwicky 1986) and attributed to Thomas Wasow stated below as in (4), preserves the essence of Chomsky's remark in (3).

(4)   *Wasow's Generalization*: If a coordinate structure occurs in some position in a syntactic representation, each of its conjuncts must have syntactic feature values that would allow it individually to occur in that position.   (as formulated in Pullum and Zwicky 1986, p. 752-753)

Note that (4) allows the exceptions to LCL like (2): in the sentence "I consider that [DP a rude remark]," the syntactic position occupied by the bracketed DP constituent can also be occupied by a PP as in "I consider that [PP in very bad taste]." Hence, according to Wasow's generalization, the same position can also be occupied by a coordination of the DP and the PP, forming a coordination of unlike categories as in (2).

In the current study, we examine a particular type of coordination of unlike categories, which, unlike (2), poses a challenge to Wasow's generalization. Consider example (5), which involves a coordinate structure of a DP conjunct and a CP conjunct, occupying the complement position of a preposition.

(5)   You can depend on [DP my assistant] and [CP that he will be on time].

<div align="right">(Sag et al. 1985, p. 165)</div>

Notably, the preposition "on" does not subcategorize for CP complements headed by "that." This is shown in (6). This leads to a violation of the Wasow's generalization in (4): the CP conjunct in (5) cannot individually occur in the identical syntactic position of the full coordinate structure.

(6)   *You can depend on that my assistant will be on time.   (Progovac 1998, p. 4)

Various analyses have been proposed to account for the DP&CP coordination in (5), the most prominent of which was proposed in Bruening and Al Khalaf (2020). In the current study, we present experimental evidence against Bruening and Al Khalaf (2020)'s analysis of the DP&CP coordination and suggest a potential alternative.

2. DP & CP Coordination

In this section, we will review the various properties of the DP&CP coordination, and past attempts at capturing these properties.

As noted in the previous section, the most notable property of the DP&CP coordination is that it does not obey Wasow's generalization. In a sentence like (5), the DP&CP coordinate can appear as the complement of a preposition, while the CP constituent (headed by the complementizer "that") cannot independently occur in that position. Furthermore, there is a restriction on conjunct ordering for the DP&CP coordinate: when the CP conjunct is linearly the first conjunct in the coordinate, the coordinate can no longer appear as the complement of a preposition. This is shown by the contrast between (5) and (7).

(7)   *You can depend on [CP that my assistant will be on time] and [DP his intelligence].
                                    (Bruening and Al Khalaf 2020, p. 2; Sag et al. 1985, p. 165)

Two classes of syntactic accounts have been proposed for the contrast between (5) and (7). Some argue that the first conjunct is more prominent structurally compared to the other conjuncts, and thus its features can percolate up to the coordinate structure and determine its syntactic category for selectional purposes (Johannessen 1998, Munn 1993, Zhang 2009, among others).

Another approach, proposed by Bruening and Al Khalaf (2020), argues that CP conjuncts can be contained inside a null DP shell (or a null NP shell as in Bruening and Al Khalaf (2020)'s original formulation), as shown in (8).

(8)   a. You can depend on [DP my assistant] and [DP ∅ D [CP that he will be on time]].
       b. *You can depend on [DP ∅ D [CP that my assistant will be on time]] and [DP his intelligence].

The null D head bears c-selectional features, but it lacks any S(emantic)-features (e.g., animacy, sentience, etc.) relevant for semantic selection that a regular DP bears (Breuning and Al Khalaf 2020). Breuning and Al Khalaf (2020) further assumes a counter-cyclic derivation of coordinate structures where the first conjunct is merged before the second.

Bruening and Al Khalaf (2020) further assume that unlike the c-selectional features on the selector ("on" in (8)) which needs to be checked against conjuncts, S-features only need to be checked against the first conjunct and then can be deleted immediately. Under the counter-cyclic derivation assumption, while deriving a sentence like (8a), there would be an intermediate step of derivation as shown in (9). At this step, the first conjunct is already merged into the structure. When the first conjunct is a real DP "my assistant" rather than a CP with a DP-shell, it bears S-features and can immediately check and delete the corresponding S-features on the selector "on."

(9)   Counter-cyclic derivation of (8a)
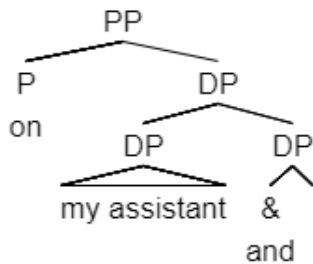
By contrast, when deriving sentence (8b), there would be an intermediate step as shown in (10). The first conjunct is a DP headed by a null D, and therefore does not bear any S-feature. Since the S-feature requirements on the selector P-head need to be checked and deleted immediately, and there is no S-feature-bearing DP in the structure, the derivation crashes at this step. Hence, sentence (8b) is ungrammatical, whereas sentence (8a) is grammatical.

(10)   Counter-cyclic derivation of (8b)
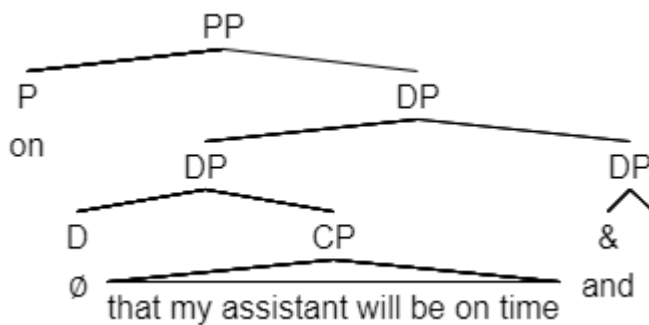
Note that Bruening and Al Khalaf (2020)'s analysis makes the crucial prediction that as long as the first conjunct is a true DP (as opposed to a CP with a silent DP shell), the sentence is grammatical because the featural requirements imposed by the selector P-head is satisfied, and the conjunct ordering beyond the first conjunct does not affect the grammaticality status. In other words, no contrast is predicted between the following examples:

(11)   a. The success of the project depends on [DP a good engineering design], [DP the diligence of the
              workers], and [CP that the contractors will do their part].
        b. The success of the project depends on [DP a good engineering design], [CP that the contractors will
              do their part], and [DP the diligence of the workers].

In this study, we put this prediction to test. In an acceptability judgment experiment, we show that contrary to what the previous accounts predict, even when the first conjunct is a true DP conjunct, the linear proximity of the CP conjunct to the selector P-head still leads to lower acceptability. We speculate that this result might be better captured by a processing-based account, which we present in the next section in detail.

3. A Processing-Based Account

In this study, we propose an alternative processing-level account to explain the contrast between (5) and (7), inspired by the concept of "structural forgetting" proposed by Gibson and Thomas (1999).

Structural forgetting refers to the phenomenon whereby the representation of certain constituent gradually decays in memory as the sentence unfolds. For illustration, in (12), the first NP (NP1), *the ancient manuscript* prompts the parser to expect for a matrix verb (the structurally highest predicate). The second NP (NP2), *the grad student* signals that the sentence involves a center-embedded relative clause and this prompts the expectation for a verb (that is in a dependency with NP2) and an empty category (which is co-indexed with NP2) as part of a relative clause. Similar to the prediction generated by the second NP, the third NP (NP3), *the new card catalog*, also prompts the expectation for a verb and an empty category that are in dependency relations with NP3. Before the predicted verbs and empty categories are encountered by the parser and the dependencies are fully resolved, the three NPs (NP1-3) need to be held in the working memory of the parser, the process of which incurs storage costs.

These storage costs caused by the predictions are contingent on the distance as well as the number of discourse new referents (Gibson 1998, 2000). Let us examine the relevant costs when the sentence that involves three NPs (NP1, NP2, and NP3) is incrementally processed in (12). NP1's prediction is expected to be costless because the relevant prediction is restricted to the main verb. In contrast, NP2's prediction is known to be more expensive relative to NP3. This is because the prediction for the NP2 as well as the discourse new referent (NP3) should be maintained in memory unlike NP3 where the prediction for the NP3 should be made but does not have a new discourse new referent to be introduced (Gibson 1998, 2000, Gibson and Thomas 1999, Huang and Phillps 2021). Given humans' limited memory capacity, the predictions are expensive to simultaneously retain in memory, engendering additional memory costs. The parser hence forgets the previously parsed elements with the highest memory costs with the aim to mitigate the memory demands arising from the costs associated with the predictions (Gibson 1998, 2000, Gibson and Thomas 1999, Huang and Phillps 2021). Hence, the prediction associated with NP2 is forgotten from memory.

In their experiment, they omitted each of the three VPs from the grammatical sentence in a systematic fashion to generate the missing VP conditions as illustrated in (12a-c). An example (12d) refers to the grammatical sentence without the missing VP condition.

(12)   a. [NP1 The ancient manuscript] that [NP2 the grad student] who [NP3 the new card catalog] *Missing VP1* [VP2 was studying in the library] [[VP3 was missing a page].
       b. [NP1 The ancient manuscript] that [NP2 the grad student] who [NP3 the new card catalog] [VP1 had confused a great deal] *Missing VP2* [VP3 was missing a page].
       c. [NP1 The ancient manuscript] that [NP2 the grad student] who [NP3 the new card catalog] [VP1 had confused a great deal] [VP2 was studying in the library] *Missing VP3*.
       d. [NP1 The ancient manuscript] that [NP2 the grad student] who [NP3 the new card catalog] [VP1 had confused a great deal] [VP2 was studying in the library] [VP3 was missing a page].

                                                                                (Gibson and Thomas 1999, p. 237)

Results revealed that the acceptability of (12b) where VP2 is missing is known to be better compared to other three conditions. When the intermediate VP (VP2) is omitted, the sentence is perceived to be acceptable, because the reader selectively forgets some of the partially processed information under high memory costs. These observations are implemented using the computational models in Futrell et al. (2020).

In this paper, we propose a processing-based account as an alternative to the two aforementioned syntactic accounts. We hypothesize that sentences (5) and (7) are both ungrammatical due to the violation of "Wasow's generalization" (a more relaxed version of the LCL) (Pullum and Zwicky 1986): all conjuncts should individually satisfy the requirements of the full coordinate structure's selector. The CP conjunct cannot be individually selected by the matrix predicate *depend on*, resulting in ungrammaticality in both (5) and (7).

However, sentence (5) is deceptively acceptable compared to (7) because the CP conjunct (*that he will be on time*) is distanced from its selector *depend on*, with another grammatical selectee *my assistant* intervening, and thus the selectional failure is less detectable. This effect is akin to the "structural forgetting" effect (Futrell et al. 2020, Gibson and Thomas 1999, Hahn et al. 2022) whereby the comprehender's memory of previously parsed elements fades as processing proceeds. In this case, the comprehenders are gradually "forgetting" the selectional requirements of the matrix predicate as they proceed through the various conjuncts, and the selection failure between the matrix predicate and the CP conjunct becomes less and less likely to be identified.

Our "selectional feature forgetting" proposal predicts that when the CP conjunct is further away from the selector as in (13a) (e.g., when the CP conjunct is the third conjunct), the resulting sentence should be even more acceptable compared to (10b) and (10c) where the CP conjunct is the second and the first conjunct. For memory load reasons, the parser forgets that the CP (*that he will be on time*) is selected by the matrix predicate, *depend on* if the distance between the selector and the CP increases.

(13)   a. The success of the project depends on [DP a good engineering design], [DP the diligence of the workers], and [CP that the contractors will do their part].
         b. The success of the project depends on [DP a good engineering design], [CP that the contractors will do their part], and [DP the diligence of the workers].
         c. The success of the project depends on [CP that the contractors will do their part], [DP a good engineering design], and [DP the diligence of the workers].

Crucially, neither the structural prominence account nor Bruening and Al Khalaf (2020)'s counter-cyclic derivation account predicts any distinction between the various positions of the CP conjunct as long as it is not the initial conjunct.

4. Experiment

We conducted an acceptability judgment experiment to test whether we observe the grammatical illusion for DP-CP coordination sentences by manipulating the distance between the position of the CP and the selector. Specifically, we investigate whether the acceptability improves when the CP conjunct is distanced from the selector that prohibits a CP complement.

4.1 Participants

Thirty-two native speakers of English (Age: 20-50) from the US and Canada were recruited through Prolific (https://www.prolific.com/). All participants self-reported that they are native speakers of English who dominantly used English since the age of 3. Participants provided informed consents before taking part in the experiment and were compensated at a rate approximately equivalent to 10/hr.

4.2 Materials

There are three critical conditions: CP First, CP Middle, and CP Last, where the CP conjunct is the first, second and third conjunct respectively. The first condition was in the order of CP-DP-DP, the second condition was in the order of DP-CP-DP and the last condition was in the order of DP-DP-CP. The conditions varied with the position of the CP with respect to the selector which precludes a CP complement. We included grammatical fillers and word salads. Grammatical fillers consisted of grammatical sentences where the selectors allowed the DP complements (DP-DP-DP) or the CP complements (CP-CP-CP). Words salads sentences consisted of ungrammatical sentences which cannot readily be parsed.

Table 1. Example Stimuli

Condition	Scheme	Example Stimuli
CP-First	CP & DP_1 & DP_2	The success of the project depends on that the contractors will do their part, a good design by the engineers, and the diligence of the workers.
CP-Mid	DP_1 & CP & DP_2	The success of the project depends on a good design by the engineers, that the contractors will do their part, and the diligence of the workers.
CP-Last	DP_1 & DP_2 & CP	The success of the project depends on a good design by the engineers, the diligence of the workers, and that the contractors will do their part.
Grammatical Filler	DP & DP & DP CP & CP & CP	The student is looking for her proof of enrollment, her updated class schedule, and the list of textbooks for her classes.
Ungrammatical Filler	Word Salad	The detective interviewed and the who was who responded first to the victim, the witnesses, investigating the robbery police officer the scene.

The matrix predicates (e.g., *depend on, talk about, account for*, etc.) were chosen such that they only allow DP complements (**depend on that the contractors will do their part*; *depend on a good design by the engineers*). All conjuncts were of comparable length. There is a total of 16 critical items, and three lists were created in a Latin Square fashion. Each list contains 16 critical sentences, 16 grammatical fillers, and 16 word salads.

4.3 Procedures

Participants were asked to rate the acceptability of the experimental sentences based on a 7-point Likert rating scale in which 1-point indicated *totally unacceptable* (*very difficult to understand*) and 7-point indicated *totally acceptable* (*very easy to understand*). Participants were instructed to read each sentence at the regular pace and rate the sentence based on their first impression. Each item appears precisely once for each participant. Participants were also presented with one of the three conditions that was randomly instantiated. Prior to the experiment, participants were given 3 practice sentences. It took around 15 minutes for them to complete the experiment.

4.4 Results

We analyzed the results using a linear mixed effect regression model predicting acceptability ratings with the fixed effect of *Condition* (reference level: CP Middle). The model also involves random by-item as well as by-participant intercepts and slopes for the *Condition* fixed effect (Baayen et al. 2008; Bates et al. 2015).

We predicted 3-way distinction in acceptability ratings where the CP-First condition should be the lowest in ratings, followed by the CP-Middle condition, followed by the CP-Last condition. According to the syntactic accounts (Bruening and Al Khalaf 2020), there should be no distinctions between CP-Middle condition and the CP-Last condition as they will both be unacceptable. Thus, under Bruening and Al Khalaf (2020)'s analysis, DP&CP coordination is ungrammatical only when CP conjunct is initial (specifically when CP is adjacent to the selector) with no observable distinctions between CP-Middle and CP-Last conditions. The mean acceptability rating of each condition is shown in Figure 1.
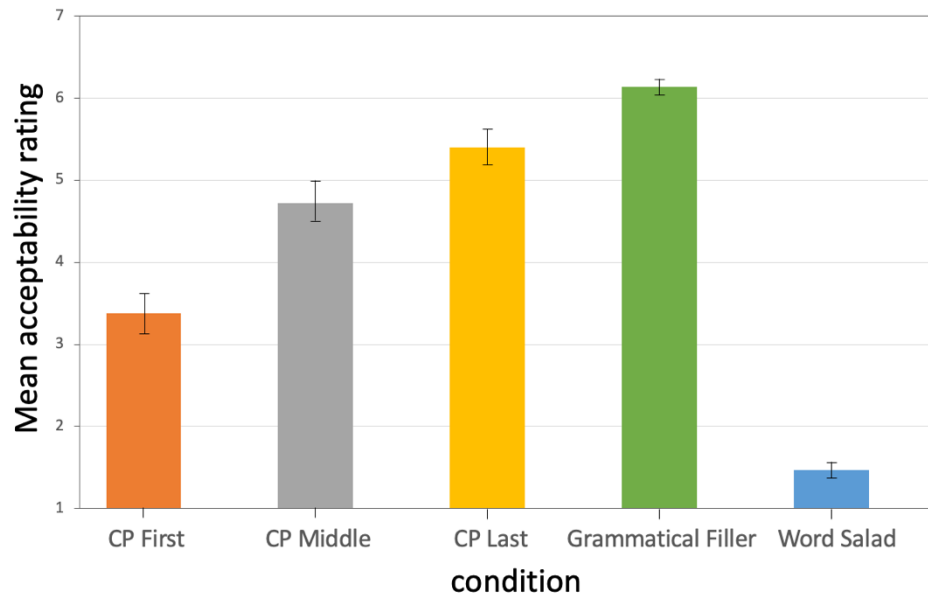

Figure 1. Mean Acceptability for Each Condition, Error Bars Represent 95% Bootstrap Confidence Intervals.

Comparing to the CP Middle condition, the CP Last condition is significantly higher in acceptability (β=0.62, SE=0.16, t=3.81, p<0.001), and the CP First condition is significantly lower in acceptability (β=−1.42, SE=0.25, t=−5.71, p<0.001).

Furthermore, we conducted follow up analyses with the same regression model but with the reference level of *Condition* set to either of the filler conditions. The analyses showed that the CP Last condition, which is rated the highest among the three critical conditions, is still significantly less acceptable than the grammatical filler condition where all three conjuncts match in syntactic category (the three conjuncts are either all true DPs, or all CPs) (β=−0.76, SE=0.19, t=−3.93, p<0.001). The CP First condition, which is rated the lowest among the three critical conditions, is still significantly more acceptable than the word salad condition (β=1.88, SE=0.30, t=6.18, p<0.001).

5. Discussion & Conclusion

In this study, we tested the acceptability of sentences with a P-head taking as its complement a DP&CP coordinate structure with three conjuncts (two DP conjuncts and a CP conjunct). We found that the acceptability of the sentences correlates with the distance of the CP conjunct to the selector P-head: there is a three-way distinction among the critical conditions. When the complement position of a predicate that only selects DP is a coordinate structure with one CP and two DP conjuncts, the acceptability is the highest when the CP conjunct is the third conjunct (and thus furthest from the selector), intermediate when CP conjunct is the second conjunct, and the lowest when the CP conjunct is the first conjunct (and thus immediately adjacent to the selector). Furthermore, we found that grammatical filler sentences without category mismatch between conjuncts are rated more acceptable than the DP&CP coordination even when the CP conjunct comes last.

These results pose a challenge for the previous accounts for the DP&CP coordination. Under either the structural prominence account (Johannessen 1998, Munn 1993, Zhang 2009) or the counter-cyclic derivation analysis (Bruening and Al Khalaf 2020), there is no distinction predicted between the various positions of the CP conjunct as long as it is not the initial conjunct. Both CP Middle and CP Last sentences are predicted to be perfectly grammatical, while the CP First condition is ungrammatical. Furthermore, neither accounts predict any contrast between a grammatical all-DP coordination and the DP&CP coordination when the CP conjunct does not come first. The three-way acceptability distinction among the three critical conditions, as well as the contrast between the grammatical fillers and the CP Last condition directly challenge the predictions of the two previous accounts.

By contrast, our processing-level account is supported by the experimental results. The CP Middle and CP Last condition, despite having relatively higher ratings than the CP First condition, are in fact both significantly less acceptable than the grammatical filler where all three conjuncts match in category. Such contrast can be easily explained as a violation of the LCL (or Wasow's generalization): since a CP cannot independently occupy the complement position of a P-head, any DP&CP coordination in the complement position of a P-head would be degraded.

Furthermore, under the processing-based account, the acceptability degradation due to the failed selection should fade as the selector-to-selectee distance increases (Futrell et al. 2020, Gibson and Thomas 1999, Hahn et al. 2022). The three critical conditions, despite all being ungrammatical, show a three-way acceptability distinction: the penalty of violating the Wasow's generalization gradually decreases as the selector-to-selectee distance increases because the mental representation of the selectional features that the selector P-head bears are less clearly represented in the comprehender's memory.

References

Baayen, R. H., D. J. Davidson and D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390-412.

Bates, D., R. Kliegl, S. Vasishth and H. Baayen. 2015. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bruening, B. and E. Al Khalaf. 2020. Category mismatches in coordination revisited. *Linguistic Inquiry* 51, 1-36.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Futrell, R., E. Gibson and R. P. Levy. 2020. Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science* 44, e12814.

Gibson, E. 1998. Syntactic complexity: Locality of syntactic dependencies. *Cognition* 68, 1-76.

Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita and W. O'Neil, eds., *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*, 94-126. Cambridge, MA: MIT Press.

Gibson, E. and J. Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14, 225-248.

Hahn, M., R. Futrell, R. Levy and E. Gibson. 2022. A resource-rational model of human processing of recursive linguistic structure. *Proceedings of the National Academy of Sciences* 119, p.e2122602119.

Huang, N. and Phillips, C. (2021). When missing NPs make double center-embedding sentences acceptable. *Glossa: A Journal of General Linguistics* 6(1).

Johannessen, J. B. 1998. *Coordination*. Oxford: Oxford University Press.

Munn, A. B. 1993. *Topics in the Syntax and Semantics of Coordinate Structures*. Doctoral dissertation, University of Maryland, College Park.

Progovac, L. 1998. Structure for Coordination. *Glot International* 3, 3-6.

Pullum, G. K. and A. M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language*, 751-773.

Sag, I. A., G. Gazdar, T. Wasow and S. Weisler. 1985. Coordination and how to distinguish categories. *Natural Language & Linguistic Theory* 3, 117-171.

Williams, E. S. 1981. Transformationless grammar. *Linguistic Inquiry*, 645-653.

Zhang, N. N. 2009. *Coordination in Syntax* (Vol. 123). New York: Cambridge University Press.