An experimental investigation of partial control

Aaron Steven WHITE — *University of Maryland*
Thomas GRANO — *University of Maryland*

Abstract. In Partial Control, the understood subject of an obligatorily controlled complement is construed as properly containing the controlling argument (e.g., *John wanted [PRO to gather at noon]*, where PRO = John and contextually salient others). Since this phenomenon was first systematically described by Landau (2000), many accounts of it have been offered. But some of the diversity in these accounts reflects disagreement over what the facts are. To address this, we use experimental syntax to collect and analyze sentence acceptability judgments that bear on the availability of Partial Control across a wide range of control predicates. Results indicate a substantial amount of gradability in the availability of Partial Control as a function of the choice of control predicate, in a way that depends in part on the predicate's temporal, aspectual, and modal properties. Although these findings are not fully consistent with any existing approach to Partial Control, we suggest that Pearson's (2013) approach lays a promising foundation for further research.

Keywords: control, partial control, propositional attitude predicates, experimental syntax

1. Introduction

Landau (2000) — systematizing observations by Wilkinson (1971); Lawler (1972); Williams (1980); Martin (1996); Petter (1998); Wurmbrand (1998) — identifies a split between E(xhaustive) C(ontrol) predicates and P(artial) C(ontrol) predicates: whereas EC predicates require that the controlling argument exhaustively identify the understood subject of the controlled complement, PC predicates allow the controller to partially identify the subject of the controlled complement. The contrast can be brought out via collective predicates, which crucially are unacceptable with a syntactically singular, non-group-denoting subject, as illustrated in (1). As shown in (2)–(4), a singular, non-group-denoting controller is unacceptable with a collective controlled complement when the control predicate is *try*, but acceptable when the control predicate is *regret* or *want*. Consequently, *try* is classified as an EC predicate whereas *regret* and *want* are classified as PC predicates, the latter group giving rise to control structures in which the controlled position includes the controller but may also include other referents made salient in the context.

(1) a. Kim [solved the problem]. NONCOLLECTIVE PREDICATE
 b. *Kim [solved the problem together]. COLLECTIVE PREDICATE

We are grateful to reviewers and audience members at Sinn und Bedeutung 18 as well as various audiences at the University of Maryland. We would also like to thank Valentine Hacquard, Norbert Hornstein, Jeff Lidz, Alexander Williams, and Hazel Pearson for useful comments as we were developing and interpreting the experiments presented here. This work was supported by an NSF BCS grant (1124338). Aaron Steven White was further supported by an NSF DGE IGERT grant (0801465).

c. *Kim [gathered at noon]. COLLECTIVE PREDICATE

(2) a. Kim **tried** [PRO to solve the problem].
b. *Kim **tried** [PRO to solve the problem together].
c. *Kim **tried** [PRO to gather at noon].

(3) a. Kim **regretted** [PRO solving the problem].
b. Kim **regretted** [PRO solving the problem together].
c. Kim **regretted** [PRO gathering at noon].

(4) a. Kim **wanted** [PRO to solve the problem].
b. Kim **wanted** [PRO to solve the problem together].
c. Kim **wanted** [PRO to gather at noon].

Landau (2000) singles out the control predicate's temporal properties as a crucial factor in determining whether it is EC or PC. In particular, Landau suggests that the EC status of *try* correlates with its inability to support matrix/embedded temporal mismatches and the PC status of *regret* and *want* correlates with their ability to do so, as brought out by the data in (5)–(7). Taking this correlation as central, Landau (2000) argues that the availability of matrix/embedded temporal mismatches are signals of T-to-C movement in the complement clause, which interacts with the determination of PRO's features in such a way that PC may obtain only if T-to-C movement obtains.

(5) **Today*, Kim **tried** [PRO to solve the problem *yesterday/tomorrow*].

(6) *Today*, Kim **regretted** [PRO solving the problem *yesterday*].

(7) *Today*, Kim **wanted** [PRO to solve the problem *tomorrow*].

In the years since Landau's seminal work, a number of researchers have proposed alternative accounts of the EC/PC split. In one vein, Wurmbrand (2002); Barrie (2004); Barrie and Pittman (2004); Cinque (2006); Costantini (2010); Grano (2012) all entertain versions of the view that the split tracks a distinction in whether the control predicate is a restructuring predicate or not. In another vein, some researchers working with the Movement Theory of Control (Hornstein, 1999) have argued that PC arises when a silent associate morpheme is stranded in the controlled complement (Rodrigues, 2008), or when the controlled complement contains a silent commitative expression (Boeckx et al., 2010), or when defective thematic intervention obtains (Sheehan, 2012). (See also Dubinsky and Hamano 2010 for the view that movement is blocked with PC predicates because the complement of a PC predicate has its own event feature.) A number of scholars have also pursued non-syntactic approaches. Jackendoff and Culicover (2003) argue that PC is not syntactically distinct from EC but rather “occurs in contexts where the controller holds a joint intention with respect to the activity described by the complement” (p. 549). Bowers (2008) argues that PC is “not actually a grammatical phenomenon at all” (p. 140) but rather instantiates a kind of metonymy that is found even in raising and monoclausal contexts. More recently, Pearson (2013) argues that the EC/PC split arises as a semantic consequence of the mechanism by which attitude-ascribing control predicates quantify over world-time-individual triplets. On Pearson’s view, in order for a control predicate to be a PC predicate, it must support matrix/embedded temporal mismatches and must also be attitude-ascribing. In a similar vein to Pearson 2013 and also drawing on Bianchi 2003, Landau (2013b) updates his own previous view, suggesting that the fundamental distinction between EC and PC predicates is that only the latter are attitude predicates. For crosslinguistically oriented work on partial control, see also Madigan 2008 (Korean); Witkos and Snarska 2008 (Polish); Modesto 2010 (Brazilian Portuguese).

The motivation for the present study is that one source of the diversity found in the aforementioned approaches is disagreement over what the data and descriptive generalizations are. To take one example, Landau (2000) and Pearson (2013), although they both take a control predicate’s temporal properties as relevant in tracking the EC/PC split, come down differently on whether English *claim* as well as Romance control predicates of belief such as Italian *credere* ‘believe’ is EC or PC. To take another example, Rodrigues (2008) departs from both Landau and Pearson by claiming that modal auxiliaries like *can* are PC. In a more extreme case, Bowers (2008) suggests that (some of) Landau’s EC predicates, as well as (some) raising predicates support PC, and that even direct (monoclausal) predication between a collective predicate and a singular non-group denoting subject is possible in some contexts. (For a response to some of these positions, see Landau 2013a:155–172.)

Consequently, the goal of the work described here is to use experimental techniques to address the following three questions:

(8) a. Do control predicates reliably differ in their tolerance for PC?
b. If so, are Landau (2000) and Pearson (2013) correct that the temporal properties of the predicate reliably predict its tolerance for PC?
c. Do other (nontemporal) properties of the predicate correlate with the acceptability of PC?

In a nutshell, our results indicate affirmative answers to all three questions: control predicates differ in their tolerance for PC, and the relative tolerance of a given predicate correlates not only with its temporal properties but also with aspectual and modal properties to be discussed in more detail below. Our results also indicate a substantial amount of gradability in the tolerance for PC as a function of the choice of embedding predicate, and the statistical analysis reported below suggests that this gradability is not reducible to inter-speaker variation. As far as we know, no existing theoretical account predicts this full range of influencing factors, though we suggest below that Pearson (2013) lays a promising foundation for further research.

The organization of the rest of the paper is as follows. Section 2 describes the methodology for the experiment. Section 3 describes the results and detailed statistical analysis. Section 4 articulates some overall conclusions and speculations toward a theoretical account of the findings.

2. Methodology

2.1. Materials

We constructed items using the the template in (9a), where X and Y are proper names, P is one of the 30 embedding predicates in (10), and Q is one of the five embedded predicates in (11), in either infinitival or gerundive form depending on the selectional properties of P.[1] P mostly consists of uncontroversial control predicates, chosen to represent a wide range of semantic properties, but also includes for comparison one raising predicate (*be likely*$_{to}$) as well as two predicates (*begin*$_{to}$ and *need*$_{to}$) that have raising uses and whose control status is controversial (see, among many others, Perlmutter 1970; Rochette 1999; Fukuda 2012 on aspectual predicates and Bhatt 1998; Wurmbrand 1999; Barbiers 2002 on root modals).

For each of the P/Q pairings, we constructed two variants, one which included *together* in the embedded predicate (the COLLECTIVE embedded predicate types) and one which excluded *together* (the NONCOLLECTIVE types). Because all of the instantiations for Y constitute a syntactically singular non-group-denoting entity, the assumption is that the COLLECTIVE embedded predicate types should be acceptable only in syntactic contexts that support PC. The "preamble" *X said that…* was preprended to all items to facilitate PC by making salient a plurality, i.e., X and Y (Landau, 2000).

(9) a. X said that Y P-ed {to Q / Q-ing} (together).
b. X said that Y Q-ed (together).

(10) P ∈ {be afraid$_{to}$, be eager$_{to}$, be likely$_{to}$, be ready$_{to}$, begin$_{to}$, claim$_{to}$, decide$_{to}$, deserve$_{to}$, expect$_{to}$, hate$_{ing}$, hate$_{to}$, hope$_{to}$, intend$_{to}$, like$_{ing}$, like$_{to}$, love$_{ing}$, love$_{to}$, manage$_{to}$, need$_{to}$, offer$_{to}$, plan$_{to}$, prefer$_{to}$, pretend$_{to}$, promise$_{to}$, refuse$_{to}$, regret$_{ing}$, remember$_{ing}$, remember$_{to}$, try$_{to}$, want$_{to}$}

[1]In (10) and throughout the rest of the paper, we use subscripts *to* and *ing* with every embedding predicate to indicate an infinitival complement frame or a gerundive complement frame respectively. As evident in (10), some predicates (*hate, like, love, remember*) were tested with both infinitival complements and gerundive complements.

(11) Q ∈ {be happy, hang out at the mall, rent an apartment in Baltimore, support the child, take yoga classes}

Each of the 30 embedding predicates was paired with four out of the five embedded predicates. The resulting 240 items (30 embedding predicates X 4 embedded predicates X COLLECTIVE and NONCOLLECTIVE variants) were distributed over 3 experiments, each of which also included 8 'baseline' items using the template in (9b), as well as 80 filler items each.[2] In each experiment, both the embedding predicate (*be afraid*$_{to}$, *be eager*$_{to}$, etc., + BASELINE) and the embedded predicate type (COLLECTIVE vs. NONCOLLECTIVE) were within-subject factors.

2.2. Participants

72 native speakers of American English (24 per experiment) were recruited through Amazon's Mechanical Turk and asked to rate each sentence's acceptability on a 1-to-7 scale. Four of these participants (two from Experiment 1, one from Experiment 2, one from Experiment 3) were excluded from the results for having average log reaction times that were more than two standard deviations below the mean (mean=8.29, sd=0.59).

3. Results and statistical analysis

3.1. Mean ratings by embedding predicate

Figure 1 shows mean raw ratings. Embedding predicates are ordered along the x-axis by difference in mean rating between the COLLECTIVE and NONCOLLECTIVE conditions. We see that the BASELINE as well as every embedding predicate is generally rated near ceiling in the NONCOLLECTIVE condition (albeit with some variability that will be analyzed and controlled for below). As for the COLLECTIVE condition, we see a clear cline in the judgments: predicates like *need*$_{to}$ (mean=2.20), *be likely*$_{to}$ (mean=2.34), *try*$_{to}$ (mean=2.39), *manage*$_{to}$ (mean=2.11), *intend*$_{to}$ (mean=2.45), and *begin*$_{to}$ (mean=2.05) are scored on average between 2 and 2.5. At the other end of the spectrum, *hate*$_{ing}$ (mean=5.47), *regret*$_{ing}$ (mean=5.59), *love*$_{ing}$ (mean=5.76), *like*$_{ing}$ (mean=5.76), and *remember*$_{ing}$ (mean=6.10) fare much better. In some of these cases, the difference in mean rating between the COLLECTIVE and NONCOLLECTIVE conditions is small, reaching a value of 0.62 for *remember*$_{ing}$.

The great difference in acceptability under the COLLECTIVE condition found between predicates like *try*$_{to}$ at the low end vs. *regret*$_{ing}$ at the high end, as well as the placement of the BASELINE

[2]Fillers were constructed by manipulating the test items to contain plural controllers, ECM complements, or *for-to* complements. A quarter of the fillers had *together* in the embedded predicate and a quarter replaced *together* with *alone*.

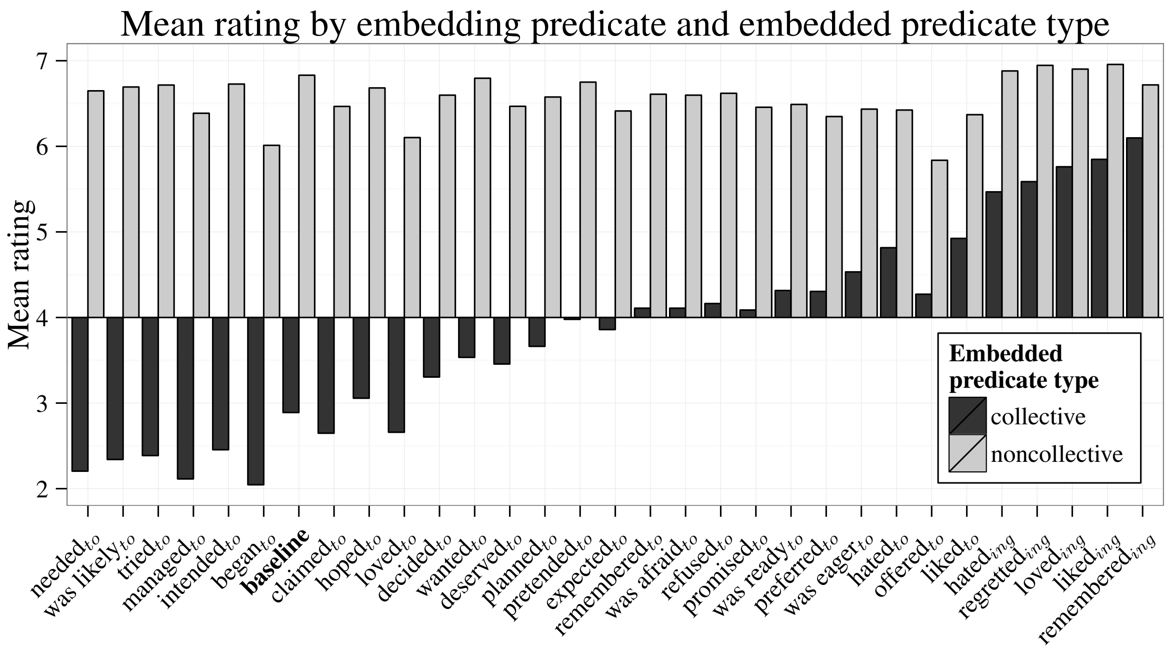


Figure 1: Mean rating of embedding predicates and BASELINE split by embedded predicate type. Embedding predicates are ordered along the x-axis by the difference in the mean rating of the embedded predicate type.

toward the low end of the spectrum, both support the view that PC is a real phenomenon, not reducible to the putative general availability of collective predication with singular non-group-denoting subjects under certain conditions, *contra* Bowers 2008. However, the high degree of gradability found in the COLLECTIVE ratings warrants further investigation, which we turn to presently.

3.2. Understanding gradability in ratings

3.2.1. Participant effects and embedded predicate effects

There are at least four possible sources of the high gradability in the COLLECTIVE condition ratings. First, it could be that the acceptability of PC as a function of the choice of embedding predicate is inherently gradable: certain predicates support PC completely, others fail to support it, and still others fall somewhere in between. Second, it could be that the gradability reflects inter-speaker variation; that is, the predicates that fall in the middle of the spectrum are those that are rated high by some speakers and low by other speakers. Third, because each embedding predicate was tested with only four out of the five embedded predicates, it could be that properties of the

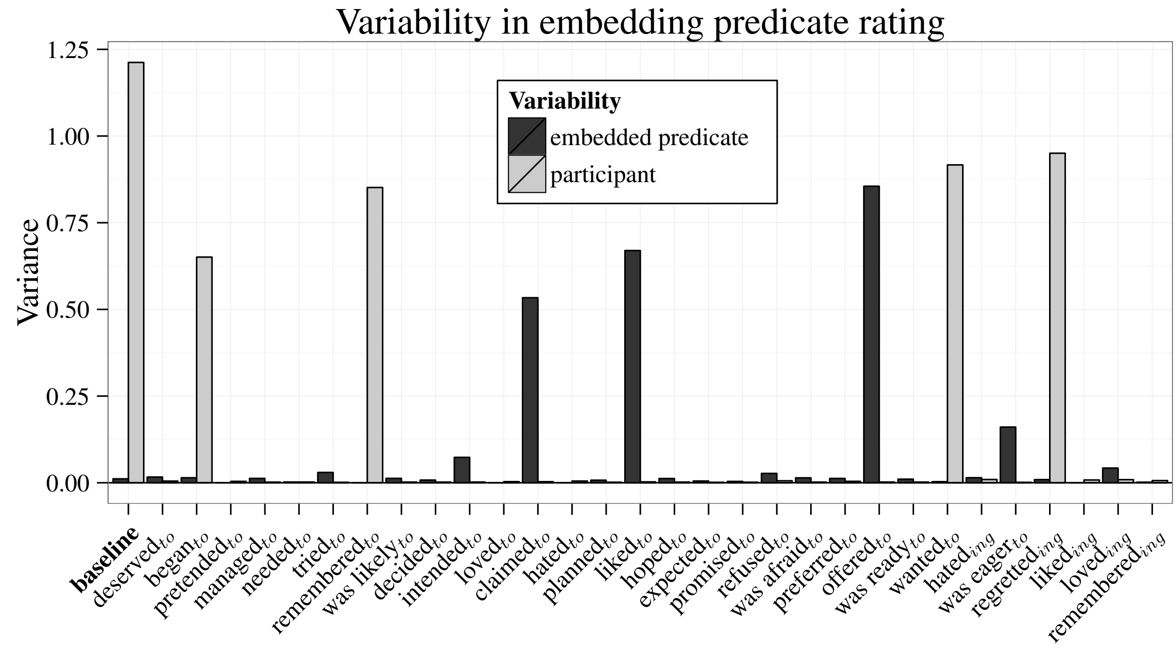


Figure 2: Estimated variability (mode) in embedding predicate rating with a collective embedded predicate due to participants and embedded predicate.

embedded predicates are responsible for the gradability. A final possibility is that there is some combination of the first three.

To determine which possibility is likely, we will build a statistical model that attempts to take five things into account simultaneously:[3] (1) idiosyncrasies in participants' use of the Likert scale; (2) differences in how participants rate a specific embedding predicate or embedded predicate; (3) variability in judgments for embedded predicates across syntactic contexts as a whole and in combination with specific embedding predicates; (4) overall higher ratings in a given experiment due to the choice of predicates in that experiment; and (5) the difference between the BASELINE and each predicate in the COLLECTIVE condition once all the other sources of variability are controlled for.

Figure 2 shows the magnitude of the noise added to each embedding predicate's ratings by participants and by embedded predicates.[4] Embedding predicates are ordered along the x-axis according

[3]We use a *cumulative link mixed model*. This family of models is designed for predicting ordered categorical judgments like those we get from a Likert (1-7) scale. It assumes that sentences fall somewhere on a continuous acceptability scale and attempts to convert the discrete Likert scale judgments into the continuous acceptability scale judgments while at the same time filtering out extraneous effects.

[4]More precisely, Figure 2 shows the variance estimate for random slopes for each embedding predicate by participant and embedded predicate. Random intercepts for both participants (mean=2.05; 95% CI=1.37–2.79) and embedded predicates (mean=0.11; 95% CI=0.00–0.35) were also fit, but they are not shown in Figure 2.

to the (filtered) difference from BASELINE for that predicate in the COLLECTIVE condition (cf. Figure 3).

As for participant effects, what we see is that only four of the 30 embedding predicates (*begin*$_{to}$, *remember*$_{to}$, *want*$_{to}$, and *regret*$_{ing}$), plus the BASELINE, show significant amounts of variability in their ratings due to participants. Furthermore, the distribution of this variability does not show a clear pattern with respect to the magnitude of the difference from BASELINE: predicates both on the low and high ends of the spectrum show variability due to participants. The high variability of the BASELINE suggests that some speakers do accept collective predication of a singular non-group-denoting subject, à la Bowers (2008). However, the model also predicts such speakers to be very rare.

As for embedded predicate effects, because, for each experiment, there are fewer embedded predicates per embedding predicate (4) than participants (21 or 22), this estimate is inherently less certain. However, we do see some interpretable results. First, *claim*$_{to}$ shows somewhat high variability (mode=0.53; 95% CI=0.00–4.57) with respect to the embedded predicate, possibly due to the special aspectual restrictions on infinitival complements to *claim* (cf. Wurmbrand, 2012). The similar status of *like*$_{to}$ (mode=0.67; 95% CI=0.04–8.57) may have an explanation along the same lines. Finally, we see that *offer*$_{to}$ has the highest variability of all (mode=0.86; 95% CI=0.00–4.90). This is somewhat surprising, and we do not have an explanation for it.

3.2.2. Filtered ratings

Figure 3 shows the estimated difference from BASELINE in the COLLECTIVE condition for each embedding predicate after filtering out participant and embedded predicate variability.[5] The cline we initially saw in the raw scores represented in Figure 1 appears to remain here. Further, the order stays roughly the same: *need*$_{to}$ (mode=0.45; 95% CI=-0.47–1.29), *be likely*$_{to}$ (mode=0.70; 95% CI=-0.26–1.43), *try*$_{to}$ (mode=0.62; 95% CI=-0.37–1.48), *manage*$_{to}$ (mode=0.43; 95% CI=-0.63–1.14), *intend*$_{to}$ (mode=0.85; 95% CI=-0.26–1.87), and *begin*$_{to}$ (mode=0.02; 95% CI=-1.03–0.91) are all on the low end of PC compatibility. The fact that these predicates' credible intervals include zero means that there is less than a 95% probability that these predicates' true ratings are different from BASELINE. And thus, under an α-value of .05, we fail to reject the null hypothesis that their ratings are different from BASELINE. Furthermore, there are predicates for which we do reject the null hypothesis: *hate*$_{ing}$ (mode=2.42; 95% CI=1.31–3.21), *regret*$_{ing}$ (mode=2.39; 95% CI=1.22–3.39), *love*$_{ing}$ (mode=2.62; 95% CI=1.57–3.53), *like*$_{ing}$ (mode=2.46; 95% CI=1.78–3.57), and *remember*$_{ing}$ (mode=3.00; 95% CI=2.12–3.90) are all reliably better than BASELINE.

[5] More precisely, these effects were included as fixed effects in the model. A variable representing the experiment each embedding predicate was tested in as well as one representing that variable's interaction with predicate type (COLLECTIVE v. NONCOLLECTIVE) were also included.

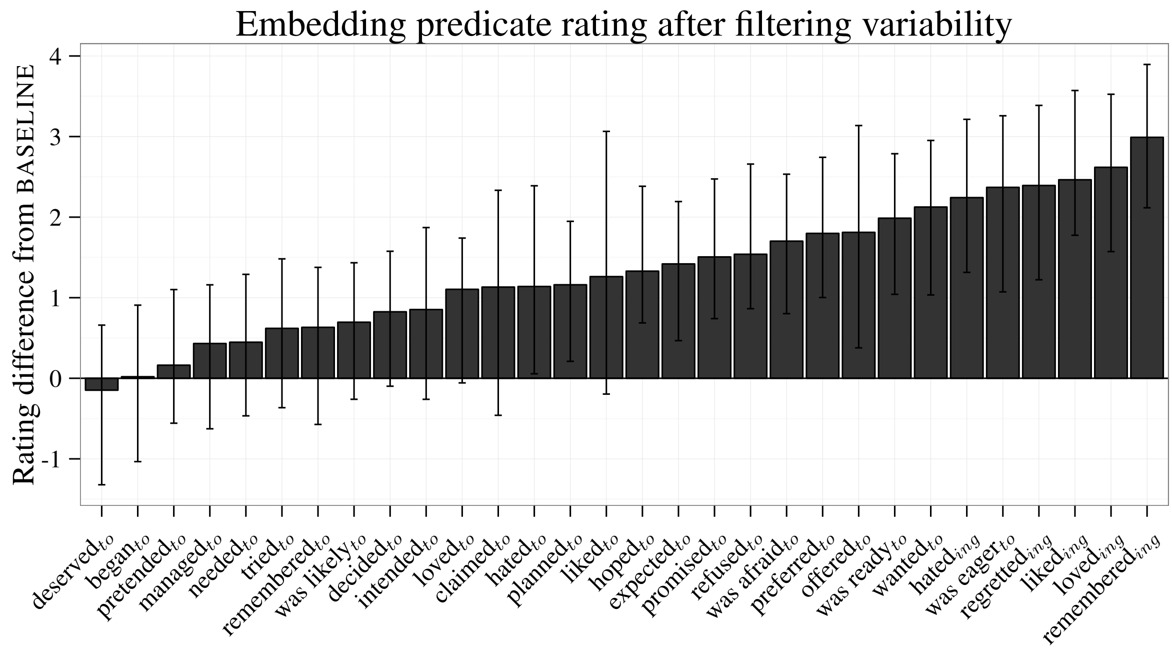


Figure 3: Estimated difference (mode) from BASELINE with a collective embedded predicate. Error bars show 95% credible intervals estimated from 1,200,000 MCMC iterations (burnin=200000; thinning=1000).

However, three qualifications are in order. First, we can make valid inferences about the "PCness" of a predicate but not about its "ECness": failing to reject the null hypothesis is not the same as accepting it. Second, there are differences in the estimates for these predicates that we gloss over by binarizing into EC and PC: a predicate like *remember*$_{ing}$ is still much better than a predicate like *expect*$_{to}$ (mode=1.42; 95% CI=0.47–2.20); in fact, even though *expect*$_{to}$ would be classed as PC under a hypothesis testing EC-PC binarization, it would also be characterized as significantly different from *remember*$_{ing}$. Finally, despite the foregoing, we have not definitively shown that gradability in PC exists: predicates with both low and high ratings have large credible intervals. This is partly due to statistical power, but it could also obscure a true binary split between EC and PC predicates if some predicates' true ratings are at the high end of their credible interval and some are at the low end. To investigate this further, we build a model in Section 3.3 that predicts these filtered ratings from semantic features of the embedding predicates.

3.3. Semantic features of the embedding predicates

In order to investigate the gradability that remains after filtering out participant and embedded predicates effects, we code our embedding predicates for four semantic feature types and build a model that allows us to assess their correlation with the acceptability of PC. These four features were

Feature	Importance	Conditional Importance
attitudinality	35.63	28.70
opacity	36.65	24.88
simultaneity	34.68	25.71
telicity	25.88	19.78

Table 1: Importance of each of the four features. Importance is measured in terms of percent increase in mean squared error due to the feature in a permutation test over the random forest ensemble.

chosen because they play a role in previous accounts of PC: SIMULTANEITY (*non-simultaneous* or *simultaneous*) (Landau, 2000; Pearson, 2013), TELICITY (*telic* or *atelic*) (Rooryck, 2007; Pearson, 2013),[6] OPACITY (*transparent* or *opaque*) (Pearson, 2013; Landau, 2013a), and ATTITUDINALITY (*attitudinal* or *nonattitudinal*) (Pearson, 2013; Landau, 2013a).[7] Predicates exemplifying each opposition, along with the diagnostic used for classification, are illustrated in (12)–(15).

(12) SIMULTANEITY
- a. *simultaneous:* Today, John **managed** to take a yoga class {#yesterday, #tomorrow}.
- b. *non-simultaneous:*
  - (i) Today, John **regretted** taking a yoga class yesterday.
  - (ii) Today, John **planned** to take a yoga class tomorrow.

(13) TELICITY
- a. *telic:* # For five days, John **decided** to rent an apartment in Baltimore.
- b. *atelic:* For five days, John **liked** renting an apartment in Baltimore.

(14) OPACITY[8]
- a. *opaque:* John **wanted** to cut a tomato, but there was no tomato.
- b. *transparent:* John **tried** to cut a tomato, #but there was no tomato.

(15) ATTITUDINALITY

[6] Pearson does not discuss telicity per se, though aspectual properties (and in particular, the presence vs. absence of temporal containment relations) of both the embedding and embedded predicates do play a role in her theory of PC. Rooryck similarly does not discuss telicity per se but identifies an aspectual component to controller selection. Although his focus is on split control rather than PC, many of his examples actually instantiate PC according to Landau's (2000) diagnostics.

[7] Landau (2013a) in fact uses opacity as the diagnostic for attitudinality. In contrast, we follow Pearson (2013) in treating predicates like try_{to} and $manage_{to}$ as attitudinal even though they are transparent.

[8] These examples are due to Sharvit 2003. See also Grano 2011; Giannakidou 2013 for discussion.

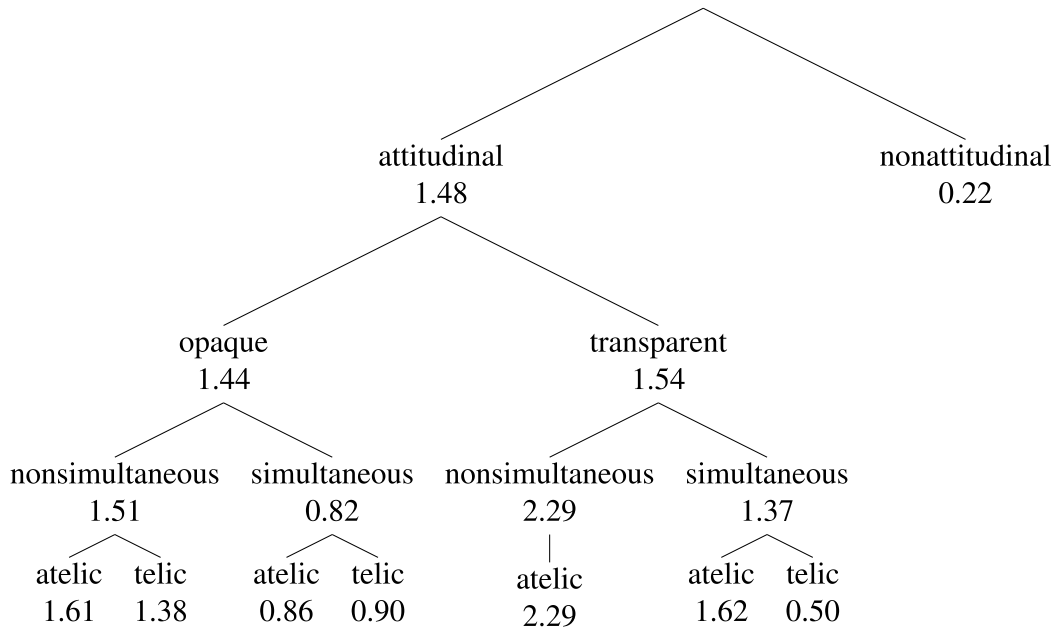


Figure 4: Tree showing predictions for some feature combinations from random forest.

a. *attitudinal:* John **wanted** to win. [ascribes an atittude to John]
b. *nonattitudinal:* John **deserved** to win. [does not ascribe an attitude to John]

The model we use is a random regression forest (iter=1000; Breiman, 2001).[9] This type of model allows us to assess the importance of a feature type in predicting compatibility with PC. Table 1 gives the importance of each feature type in terms of percent increase in mean squared error, calculated by a permutation test over each tree. Each feature type shows a reliable increase in the variance explained, suggesting that all of these feature types are correlated with PC acceptability. Strobl et al.'s (2008) conditional importance tests were also run, and the overall pattern of importance shifts slightly, favoring ATTITUDINALITY and SIMULTANEITY. The fact that all of the features remain positive even under the conditional importance test suggests that the importance of each feature type is not due to its correlation with the other feature types, which Strobl et al. (2008) show can result in unimportant features appearing important.

However, the fact that these features are important with respect to PC does not tell us whether the apparent gradability in Figure 3 is due to additivity of these features. It could be that these features interact in such a way that certain combinations are good at about the same level and certain combinations are bad at about the same level. Figure 4 shows the model's prediction for difference from BASELINE for each feature combination within the *attitudinals*. We find that on average, *attitudi-*

[9]Bootstraps for each iteration were stratified by embedding predicate. The number of features tried per split was tuned to avoid overfitting.

nals (mean=1.48; 95% CI=0.49–2.30) are predicted to be better than *nonattitudinals* (mean=0.22; 95% CI=0.03–0.51), suggesting that *attitudinality* does in fact correlate with PC. We now focus in on distinction among the attitudinals.

On average, *opaque* (mean=1.44; 95% CI=0.86–1.62) predicates do not show much distinction from *transparent* predicates (mean=1.54; 95% CI=0.48–2.31), which superficially is somewhat unexpected given recent proposals by Landau (2013b); Pearson (2013).[10] However, what appears to be driving the relatively high prediction for *transparent* predicates is the fact that (departing from Landau 2013b) we classified factive predicates ($regret_{ing}$, $like_{ing/to}$, $love_{ing/to}$, $hate_{ing/to}$, and $remember_{ing}$) as *transparent*; and as can be seen from Figures 1 and 3, these embedding predicates are some of the best with PC, regardless of variability due to participants and embedded predicates.

A likely possibility is that OPACITY interacts with other semantic properties of the control predicate to license PC. Some evidence for this view comes from apparent differences in the strength of our tense and aspect predictors with respect to OPACITY. On the one hand, *transparent* predicates appear to sit on a cline, with *simultaneous telics* worst, *simultaneous atelics* better, and *nonsimultaneous atelics* best, suggesting that both SIMULTANEITY and TELICITY play a role in licensing PC with *transparent* predicates. On the other hand, only SIMULTANEITY appears to matter for *opaque* predicates; *atelics* (mean=1.52; 95% CI=0.84–1.62) are only slightly better than *telics* (mean=1.30; 95% CI=0.89–1.39).[11]

But although *opaque simultaneous attitudinals* like $claim_{to}$ and $pretend_{to}$ are not as acceptable with PC as their *nonsimultaneous* counterparts like $want_{to}$ and $hope_{to}$, they are not predicted to be nearly as bad as *nonattitudinals* or their *transparent* counterparts like $remember_{to}$ and $manage_{to}$. We discuss possible reasons for this in Section 4.2.

4. Conclusions and speculations

4.1. Central conclusions

Our results support four central conclusions:

(16) a. Embedding predicates differ in their tolerance for PC, and these differences form a cline (gradability).
b. Temporal properties of the embedding predicate are a reliable predictor of the availability of PC.

[10]Furthermore, OPACITY alone does not predict PC; *transparent* predicates (mean=1.41; 95% CI=0.03–2.31), regardless of ATTITUDINALITY, are not predicted to be very different from *opaque* predicates (mean=1.24; 95% CI=0.86–1.62) on average.

[11]There may exist some interaction between SIMULTANEITY and TELICITY within the *opaque* predicates, but it is quite small, if it exists.

c. Aspectual properties of the embedding predicate are a reliable predictor of the availability of PC.
d. Modal properties of the embedding predicate are a reliable predictor of the availability of PC.

In what follows, we elaborate on each of these conclusions.

4.1.1. Gradability

Raw mean ratings show a cline in PC acceptability as a function of the choice of embedding predicate, ranging from 2.05 out of 7 for *begin*$_{to}$ to 6.10 out of 7 for *remember*$_{ing}$. This cline remains even after filtering out variability due to participants and embedded predicates and even after building a model of embedding predicate semantic features. The finding that embedding predicates differ in their tolerance for PC supports most previous theoretical work on the topic (with the possible exception of Bowers 2008), but no existing approaches predict the fine gradation that our experiments show (though see section 4.2 below).

4.1.2. Temporal properties

Embedding predicates that support matrix/embedded temporal mismatches (*nonsimultaneous* predicates) are more acceptable with PC than are predicates that do not support matrix/embedded temporal mismatches (*simultaneous* predicates). This finding is consistent with Landau 2000 and (with some qualifications) Pearson 2013. Given that nonsimultaneous predicates include not only future-oriented predicates like *want* but also past-oriented predicates like *regret*$_{ing}$, this finding also challenges the view in Jackendoff and Culicover 2003 that PC is parasitic on intentionality and consequently found only with future-oriented predicates.

4.1.3. Aspectual properties

Among *transparent attitudinal* predicates, being atelic as opposed to telic correlates with a boost in PC acceptability. Although telicity has not played a central role in any existing theoretical approaches to PC, it is indirectly manifest in Pearson's (2013) proposal that progressive aspect improves the acceptability of PC. Also, Rooryck (2007) has identified telicity as a factor in the acceptability of split control.

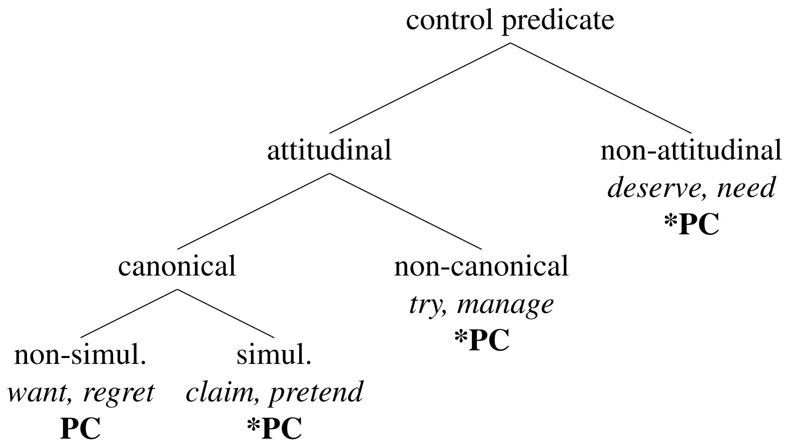
4.1.4. Modal properties

Opacity appears to modulate the effects of temporal and aspectual properties. Within *transparent attitudinal* predicates, being atelic as opposed to telic and being nonsimultaneous as opposed to simultaneous both appear to improve PC compatibility, as evidenced by the cline we see with those predicates. Within *opaque attitudinal* predicates, being nonsimultaneous appears to improve PC compatibility, but being atelic does not.

4.2. Toward a theoretical account: A variant of Pearson's theory

In this final section, we conclude with some speculative remarks on how Pearson's (2013) theory of PC might be adapted to account for some of the gradability that we see in our findings. For Pearson (2013), the semantic dimensions of control predicates that matter for determining the availability of PC are as schematized in (17).

(17)



As seen in (17), the highest level distinction is between control predicates that ascribe attitudes and those that do not. For example, *deserve*, although it passes standard tests for control, does not ascribe an attitude to its subject, and so it is classified as non-attitudinal. Among attitudinal predicates, Pearson proposes a split between what she calls 'canonical' and 'non-canonical' attitude predicates. Drawing on earlier work by Sharvit (2003); Grano (2011), Pearson suggests that non-canonical attitude predicates such as *try* are those that ascribe an attitude but behave unlike canonical attitude predicates such as *want* with respect to opacity, as evidenced by contrasts in existential entailments (see (14) above, and see also Giannakidou 2013 for an approach based on veridicality). Finally, among canonical attitude predicates, Pearson identifies a split similar to Landau's (2000) between those whose tense must be construed as simultaneous with the understood tense of the complement clause and those whose tense can be construed as non-simultaneous with the understood tense of the complement clause. Of all these categories, Pearson argues that only non-simultaneous canonical attitude predicates support PC.

To explain this generalization, Pearson proposes that canonical attitude predicates (to the exclusion of the other kinds of control predicates) all have a semantics of the kind schematized in (18), where the first argument is a proposition with open slots for an individual argument and a time interval argument. Consequently, as illustrated in (19), a complement to such a predicate undergoes lambda-abstraction with respect to both its tense and its PRO subject.

(18) $[\![$attitude predicate$]\!]$ = $\lambda P_{\langle e \langle i, \langle s,t \rangle\rangle\rangle} \lambda x \lambda t \lambda w.[\ldots]$

(19) Jay claimed/wanted [$\lambda x_1 \lambda t_2 \lambda w_3$ [PRO_1 to be $happy_{t2,w3}$]].

Simplifying somewhat, Pearson argues that PC arises as a consequence of the fact that when a canonical attitude predicate imposes a relation of non-simultaneity for its complement's temporal argument, it correspondingly imposes a subset relation for its complement's individual argument, as in (20) and (21). When an attitude predicate imposes a simultaneity relation on its complement's temporal argument, on the other hand, it correspondingly imposes an equality relation on its complement's individual argument, as in (22).

(20) Jay **wanted** [$\lambda x_1 \lambda t_2 \lambda w_3$ [PRO_1 to be $happy_{t2,w3}$]].
a. attitude time $<$ embedded clause time
b. attitude holder $\subseteq$ embedded subject

(21) Jay **regretted** [$\lambda x_1 \lambda t_2 \lambda w_3$ [PRO_1 being $happy_{t2,w3}$]].
a. attitude time $>$ embedded clause time
b. attitude holder $\subseteq$ embedded subject

(22) Jay **claimed** [$\lambda x_1 \lambda t_2 \lambda w_3$ [PRO_1 to be $happy_{t2,w3}$]].
a. attitude time = embedded clause time
b. attitude holder = embedded subject

Among our experimental findings are that simultaneous canonical attitude predicates (in particular, *claim*) are more acceptable than non-canonical attitude predicates like *try* and *manage* or some non-attitudinal predicates like *need*, though all of these types are less acceptable with PC than are most nonsimultaneous canonical attitude predicates like *promise* and *be eager*. To explain this pattern, it is tempting to adopt a framework like Pearson's, but with the modification that a simultaneous canonical attitude predicate can be 'coerced' into accepting PC by tolerating a subset relation for its complement's individual argument despite imposing an equality relation for its complement's temporal argument, albeit with a moderate degradation in acceptability, as schematized in (23). By contrast, non-canonical attitude predicates and non-attitude predicates are not eligible for this coercion, because, by hypothesis, they do not quantify over time-individual pairs.

(23) "Coerced" partial control:
Jay **claimed** [$\lambda x_1 \lambda t_2 \lambda w_3$ [PRO_1 to be $happy_{t2,w3}$]].
a. attitude time = embedded clause time
b. attitude holder $\subseteq$ embedded subject

This modified approach gives us a three-way split between predicates that are inherently PC-compatible (*want, regret*), those that can be coerced at a cost into PC-compatibility (*claim*), and those that are not coercible (*try, manage, deserve, need*). Of course, this three-way split does not account for the full range of variation that we found in the experiment. Furthermore, not all predicates behave as expected on this approach; for example, according to the filtered ratings in Figure 3, *pretend* appears at the low end of the scale, even though, as a simultaneous canonical attitude predicate, it should pattern with *claim* in having intermediate acceptability. Despite these shortcomings, though, this approach strikes us as a promising avenue for further exploration.

References

Barbiers, S. (2002). Current issues in modality: An introduction to modality and its interaction with the verbal system. In S. Barbiers, F. Beukema, and W. van der Wurff (Eds.), *Modality and its interaction with the verbal system*, pp. 1–17. Amsterdam, The Netherlands: John Benjamins B.V.

Barrie, M. (2004). Moving toward partial control. In K. Moulton and M. Wolf (Eds.), *Proceedings of the 34th meeting of the North East Linguistics Society, vol. I*, pp. 133–146. Amherst, MA: GLSA.

Barrie, M. and C. M. Pittman (2004). Partial control and the movement towards movement. *Toronto Working Papers in Linguistics 22*, 75–92.

Bhatt, R. (1998). Obligation and possession. In H. Harley (Ed.), *UPenn/MIT Roundtable on Argument Structure and Aspect, MIT Working Papers in Linguistics (MITWPL) 32*, pp. 21–40. Cambridge, MA: MITWPL.

Bianchi, V. (2003). On finiteness as logophoric anchoring. In J. Guerón and L. Tasmowski (Eds.), *Temps et Point de Vue / Tense and Point of View*, pp. 213–246. Nanterre: Université Paris X.

Boeckx, C., N. Hornstein, and J. Nunes (2010). *Control as movement*. Cambridge: Cambridge University Press.

Bowers, J. (2008). On reducing control to movement. *Syntax 11*, 125–143.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Cinque, G. (2006). *Restructuring and Functional Heads*. Oxford: Oxford University Press.

Costantini, F. (2010). On infinitives and floating quantification. *Linguistic Inquiry 41*, 487–496.

Dubinsky, S. and S. Hamano (2010). Framing the syntax of control in Japanese (and English). In N. Hornstein and M. Polinsky (Eds.), *Movement Theory of Control*, pp. 183–210. Amsterdam: John Benjamins.

Fukuda, S. (2012). Aspectual verbs as functional heads: Evidence from Japanese aspectual verbs. *Natural Language & Linguistic Theory 30*, 965–1026.

Giannakidou, A. (2013). (Non)veridicality, epistemic weakening, and event actualization: *Evaluative* subjunctive in relative clauses. In M. Taboada and R. Trnavac (Eds.), *Nonveridicality and evaluation: Theoretical, computational, and corpus approaches*, pp. 17–49. Leiden: Brill.

Grano, T. (2011). Mental action and event structure in the semantics of *try*. In N. Ashton, A. Chereches, and D. Lutz (Eds.), *Semantics and Linguistic Theory (SALT) 21*, pp. 426–443. eLanguage.

Grano, T. (2012). Control and restructuring at the syntax-semantics interface. Ph.D. Dissertation, University of Chicago.

Hornstein, N. (1999). Movement and control. *Linguistic Inquiry 30*, 69–96.

Jackendoff, R. and P. W. Culicover (2003). The semantic basis of control in English. *Language 79*, 517–556.

Landau, I. (2000). *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer.

Landau, I. (2013a). *Control in generative grammar: A research companion*. Cambridge: Cambridge University Press.

Landau, I. (2013b). A two-tiered theory of control. Unpublished ms., available at http://ling.auf.net/lingbuzz/001937.

Lawler, J. (1972). *A problem in participatory democracy*. Bloomington, IN: Indiana University Linguistics Club.

Madigan, S. W. (2008). Control constructions in korean. Ph.D. Dissertation, University of Delaware.

Martin, R. (1996). A minimalist theory of PRO and control. Ph.D. Dissertation, University of Connecticut.

Modesto, M. (2010). What Brazilian Portuguese says about control: Remarks on Boeckx & Hornstein. *Syntax 13*, 78–96.

Pearson, H. (2013). The sense of self: Topics in the semantics of de se expressions. PhD dissertation, Harvard University.

Perlmutter, D. M. (1970). The two verbs *begin*. In P. Rosenbaum and R. Jacobs (Eds.), *Readings in English transformational grammar*, pp. 107–119. Waltham, MA: Ginn-Blaisdell.

Petter, M. (1998). *Getting PRO under control*. The Hague: Holland Academic Graphics.

Rochette, A. (1999). The selection properties of aspectual verbs. In K. Johnson and I. Roberts (Eds.), *Beyond Principles and Parameters*, pp. 145–165. Dordrecht: Kluwer.

Rodrigues, C. (2008). Agreement and flotation in partial and inverse partial control configurations. In W. D. Davies and S. Dubinsky (Eds.), *New Horizons in the Analysis of Control and Raising*, pp. 213–229. Dordrecht: Springer.

Rooryck, J. (2007). Control via selection. In *New horizons in the analysis of control and raising*, pp. 281–292. Springer.

Sharvit, Y. (2003). Trying to be progressive: The extensionality of *try*. *Journal of Semantics 20*, 403–445.

Sheehan, M. (2012). A new take on partial control: defective thematic intervention. *Cambridge Occasional Papers in Linguistics 6*, 1–47.

Strobl, C., A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis (2008). Conditional variable importance for random forests. *BMC bioinformatics 9*(1), 307.

Wilkinson, R. (1971). Complement subject deletion and subset relations. *Linguistic Inquiry 2*, 575–584.

Williams, E. (1980). Predication. *Linguistic Inquiry 11*, 203–238.

Witkos, J. and A. Snarska (2008). On partial control and parasitic PC effects. *SKASE Journal of Theoretical Linguistics 5*, 42–75.

Wurmbrand, S. (1998). Infinitives. Ph.D. Dissertation, MIT.

Wurmbrand, S. (1999). Modal verbs must be raising verbs. In S. Bird, A. Carnie, J. Haugen, and P. Norquest (Eds.), *West Coast Conference on Formal Linguistics (WCCFL) 19*, pp. 599–612. Somerville, MA: Cascadilla Press.

Wurmbrand, S. (2002). Syntactic vs. semantic control. In J.-W. Zwart and W. Abraham (Eds.), *Studies in comparative Germanic syntax: Proceedings of the 15th Workshop on Comparative Germanic Syntax*, pp. 93–127. Amsterdam: John Benjamins.

Wurmbrand, S. (2012). Tense and aspect in English infinitives. Ms., University of Connecticut.