University of Pennsylvania Working Papers in Linguistics

Volume 25
Issue 1 *Proceedings of the 42nd Annual Penn Linguistics Conference*

Article 11

2-27-2019

Grammars Compete Late: Evidence from Embedded Passives

Daniel Duncan
Newcastle University

Grammars Compete Late: Evidence from Embedded Passives

Abstract
One of the biggest problems for variationist approaches to syntactic variation is the question of where such variation occurs in the grammar, and what type of variation is allowed. Kroch (1994) suggests that syntactic variables are a result of Competing Grammars, in which grammars that derive differing surface outputs are in competition and selected by the speaker. In this paper, I observe an implicit prediction of the Competing Grammars viewpoint as typically described: material above the variable cannot condition variation. I test this prediction in a variationist study of embedded passives (the ëneeds washedí construction) in Pittsburghese, and show that material above the variable does condition variation. This finding suggests that a look-ahead problem arises if a grammar in competition is selected prior to derivation of the variable. To solve this, I propose that both grammars are initially derived, and that the derivation transferred to LF and PF is chosen in Spell-Out from the two possibilities. Grammars still compete; however, the competition selects a variant later than previously thought.

Grammars Compete Late: Evidence from Embedded Passives

Daniel Duncan*

1 Introduction

To study intraspeaker syntactic variation means to ask three research questions: for a given variable, what are the language-internal factors that condition variation? What are the language-external factors that condition variation? Where and how is this variation generated in the grammar? These questions represent a blending of variationist sociolinguistics and generative syntax, two subfields which are often (incorrectly, I submit) thought to be incompatible. We can see this perceived incompatibility in some key assumptions of each approach. Variationist sociolinguistics takes the different variants of a variable to be different ways of saying the same thing (Labov 1972). Community-wide usage rates reflect the individual grammar, which means that the grammar generates each variant. By contrast, generative syntax holds that rules and operations within the grammar apply categorically, and that there is a single output to derivations (Embick 2008). At first glance, the variationist claim that variants are generated in the grammar appears irreconcilable with the generative claim that derivations apply rules/operations categorically to yield a single output. Any account of syntactic variation needs to find a way to reconcile these claims.

One such account of syntactic variation that successfully threads this needle is Kroch's Competing Grammars framework (1994, henceforth CG). CG suggests that each variant has its own grammar; as such, competition between variants is competition between grammars. Variation in usage is thus variation in grammar selection, and language-internal or language-external conditioning on variation is conditioning on grammar selection. This approach to variation is able to successfully account for surface usage rates while holding to principles of generative syntax. For each grammar, rules and operations apply categorically to yield a single output of the derivation.

In this paper, I illustrate a prediction that CG makes about syntactic variation: environments derived subsequent to derivation of the variable cannot be language-internal conditioning factors. I test this prediction on a corpus of embedded passives (the 'needs washed' construction) obtained from fan forums for Pittsburgh sports teams. I show that the kind of language-internal conditioning predicted to be banned by CG in fact occurs. I suggest that by modifying CG such that grammar selection occurs late, we can allow for this language-internal conditioning to be licit. In effect, this means that derivations of both variants are available at Spell-Out, where one is selected to be transferred to LF/PF.

2 Variation in Competing Grammars

CG holds that different grammars generate each variant of a syntactic variable, and that competition between variants is thus a result of competition between grammars. In this section I consider how this operates in practice and identify a key prediction of the framework.

2.1 Competing Grammars in Practice

I make two key assumptions upon which the following discussion of CG in practice and its predictions are quite contingent. First, I adopt the Competence Hypothesis, which states that competence and performance have a maximally transparent relationship (see Oseki 2018 and references therein). I take this to mean that syntactic derivations as theoretically described are a model of production. Second, I adopt the common assumption that derivations are built from the bottom-up.

Recent approaches to CG have observed that in order to select one grammar to derive a variant, there is some decision point prior to which each grammar in competition is available and after which one grammar has been selected (Fruehwald 2012; Wallenberg 2013). Although one may

*Thank you to audiences at PLC 42 and the NYU Syntax Brown Bag, particularly Stephanie Harves, Julie Legate, Laurel MacKenzie, and Mary Robinson, for questions and commentary that helped to clarify the predictions and assumptions underlying this paper.

place this decision point at the beginning of the derivation, it could in practice occur at any point before a variant is derived. In fact, in Fruehwald's (2012) analysis of t/d deletion in English, the competing grammars are found in differing post-syntactic readjustment rules, and the decision point is placed in PF. I follow this approach to adopt the following pipeline for syntactic variation under CG: first, the derivation proceeds through the operation prior to the beginning of the variant. Here the decision point is reached; one grammar or the other is selected to proceed. The derivation then proceeds through the variant and the remainder of the structure before being transferred to PF.

We can use the development of English periphrastic *do* as an illustration of this (1). During this period of development, there was variation between sentences in which *do* is inserted into T, and sentences in which the verb in V raised to T (see discussion in Kroch 1989). For speakers during this period, the two grammars generating each variant would have been in competition.

(1) a. I do not like green eggs and ham.
b. I like not green eggs and ham.

Figure 1 illustrates how this variation is derived in our pipeline: the derivation proceeds to NegP before reaching a decision point (A). At this point, it is decided whether to proceed with the *do*-support grammar or the V-to-T grammar (B). After this point, the derivation proceeds in one of two ways, depending on which grammar is selected (C).

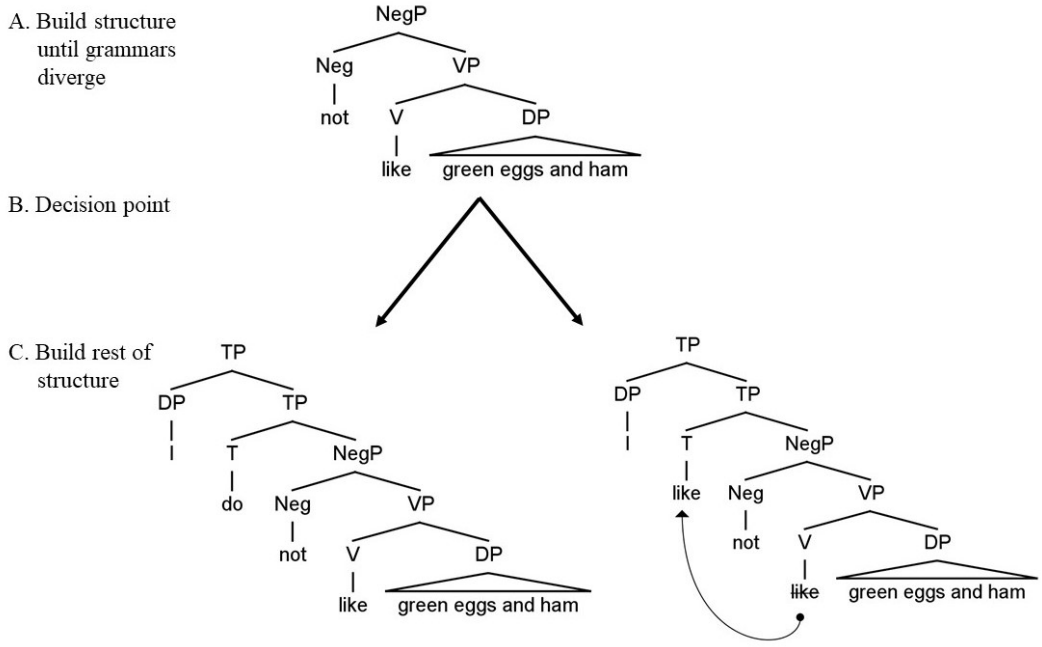

Figure 1. Deriving Variants in CG Framework.

2.2 Prediction

As seen, in our CG pipeline we have committed to a single grammar after the decision point. This means that after the decision point, we have committed to which variant will appear because one grammar yields one variant. The only way the other variant could appear is if rules and operations were not categorical within a grammar. As this is not allowed (Embick 2008), there is no way for both outputs to be derived after the decision point.

This carries implications for the language-internal conditioning of variation, the material or environments that influence the rates at which variants occur. One well-known example of this (ING): the velar [ɪŋ] is more likely to occur in nominal-like categories, while the alveolar [ɪn] is more likely to occur in verbal categories like the present participle (Houston 1985). If syntactic variables are sociolinguistic variables, they must be subject to this type of conditioning. While

often discussed with respect to variants' rates of occurrence, the variationist view of language-internal conditioning as reflecting individuals' grammars makes clear that it is at its core a matter of variant selection. After all, if a linguistic environment influences the rate of occurrence of a variant, the presence/absence of that environment is influencing whether and how often that variant is selected by the grammar. Consider what this means for our CG pipeline. Prior to the decision point, language-internal conditioning is possible. To use the *do*-support example, it is possible that the type of verb influences which syntactic variant occurs.[1] Perhaps, for example, V-to-T movement is more common with unaccusative verbs than unergative verbs. If this were the case, the derivation of one type of verb would then influence which grammar (and therefore which variant) is selected at the decision point.

This is not the case for material derived subsequent to the decision point. Because at the decision point such material has not yet been derived, it is not visible to the syntax. After the decision point, we have committed to a single grammar and therefore a single variant. Together, these points suggest that this material is not available to influence grammar selection until after the decision point, at which point it is too late. Given our assumptions and how they translate into CG, we have uncovered a prediction of the framework: material derived subsequent to the decision point (and therefore above the variant in the syntax) cannot condition variation. In fact, were such material to condition variation, it would pose a look-ahead problem for CG as described.

This prediction is empirically testable through a variationist study. Such a study requires a variable in which the two variants are a) subject to intraspeaker variation, and b) clearly syntactically different. This unfortunately rules out many variables often described as syntactic in the variationist literature, as they may be alternatively viewed as instances of lexical variation. For example, complementizer deletion (Tagliamonte and Smith 2005) is often labeled as syntactic variation, but could just as easily be described as variation between two lexical items: *that* and *Ø*. Assuming we identify such a variable, the task is to collect a sample of tokens and test for effects of language-internal conditioning resulting from environments derived subsequent to the decision point.

3 Embedded Passives

I suggest that one such variable is the English embedded passive. In some dialects of English, there is variation between the standard embedded passive (EP, 2a) and what Edelstein (2014) calls the alternative embedded passive (AEP, 2b).

(2) a. The car needs **to be washed**.
b. The car needs **washed**.

So far as I can tell, the two variants mean about the same thing. The AEP is most often found with the volitional verb *need* (Murray et al. 1996), although it has also been documented occurring with other volitional verbs *want* and *like* (Murray and Simon 1999; Murray and Simon 2002).[2] The AEP is geographically limited to the Midland within North American Englishes (Murray et al. 1996; Labov et al. 2006). It is especially associated with Pittsburghese (Tenny 1998) and the enregistered linguistic practice surrounding the dialect (Johnstone 2009).

3.1 Variation in Embedded Passives

The majority of the work examining variation in English embedded passives is from a dialectological perspective and asks little more than whether the AEP is used in a particular region. It is unclear then, whether there are any language-internal conditioning factors on variation. It is similarly undocumented whether speakers use both variants, although as a native user of both I can attest they do. It is possible to find examples online in which both variants appear intra-sententially. Such examples provide clear evidence of intra-speaker variation (3).

[1] This is a purely hypothetical example; I make no claims as to whether this is the case.

[2] The AEP is more productive than the literature describes, it is relatively easy to find examples with other verbs like *love*, *hate*, and *deserve* through Internet searches.

(3) I also think Lambo **needs swapped** with Lombardozzi, who then **needs to be given** spots starts here and there to spell people.[3]

We thus have evidence that the first criterion required of a syntactic variable to successfully test the prediction outlined in Section 2.2 has been met.

3.2 Variants Have Different Structures

Murray et al. (1996) describe the AEP as deletion of *to be*. If this were the case, the AEP and EP could be reduced to lexical variation — exactly the situation we wish to avoid in this study. However, Edelstein (2014) offers convincing evidence that the two variants are structurally different. Several tests are available to show this. For example, while the EP is compatible with eventive or stative passives (see Embick's 2004 discussion of eventive, stative, and resultative participles), the AEP is incompatible with stative passives (4).

(4) a. The door needs *(to be) open.
b. You need *(to be) hammered to enjoy this movie.

The EP can occur with negation between the matrix verb and embedded passive; this is not possible in the AEP (5).

(5) a. That car needs to not be washed.
b. *That car needs not washed.

Based on these tests and others, Edelstein (2014) suggests that unlike the biclausal EP, the AEP is a Restructuring phenomenon that has monoclausal behavior (Figure 2). Moving forward, I will adopt this analysis. However, for our purposes, the details of the analysis matter less than the observation that the two variants are clearly structurally different. Variation between the EP and AEP thus fulfills the second of our two criteria required of a variable to test the prediction outlined in Section 2.2.

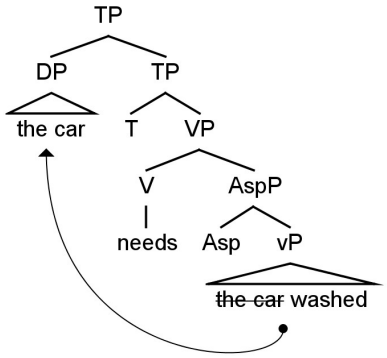

Figure 2. Structure of AEP (Adopted from Edelstein 2014:265).

3.3 Deriving Variants in Competing Grammars Framework

Considering the derivation of the two variants will enable us to more formally state the prediction for language-internal conditioning of embedded passive variation. Recall our pipeline: first the derivation is built until the decision point, then a grammar is selected, and then the remainder of the derivation is built. For the variable in question, this means building the derivation to no higher than AspP (see Figure 2). After this point, the EP or AEP grammar needs to be selected in order to build the rest of the derivation by Merging TP/CP material (EP) or directly Merging V (AEP). According to our prediction, this means that nothing derived subsequent to AspP, and certainly nothing derived after *need*, should condition variation.

[3]This and other online examples come from the sample described in Section 4.

4 Methods

The goal, then, is to test this prediction through a variationist analysis of a corpus of tokens. This is somewhat difficult, as syntactic variables like this are relatively rare. As such, sociolinguistic interviews or similar methods for collecting naturalistic data are not viable for collecting tokens. At the same time, the variable is conditioned by geography. This means that although large scale corpora such as COCA (Davies 2008-) would otherwise be ideal for investigating rare variables, they are likewise not viable as their sampling methodology would include both users of the AEP and speakers for whom the construction is ungrammatical.

To solve this, I offer a methodological innovation for sampling regional variation online. I turn to forums for fans of sports teams as a source. Fandom is highly correlated with location. For example, Figure 3 shows a map of the United States, color-coded for the most-liked Major League Baseball team on Facebook in each county. Particularly in the Northeast and Midwest, teams are liked in relatively small areas; the Pittsburgh Pirates, for example, have a fan base largely consisting of Western Pennsylvania and the Pittsburgh metropolitan area. The same is true for the Houston Astros, who are liked primarily in the Houston metropolitan area. Even among teams with larger geographical footprints, the footprints are largely contiguous with the exception of the New York Yankees and Boston Red Sox.

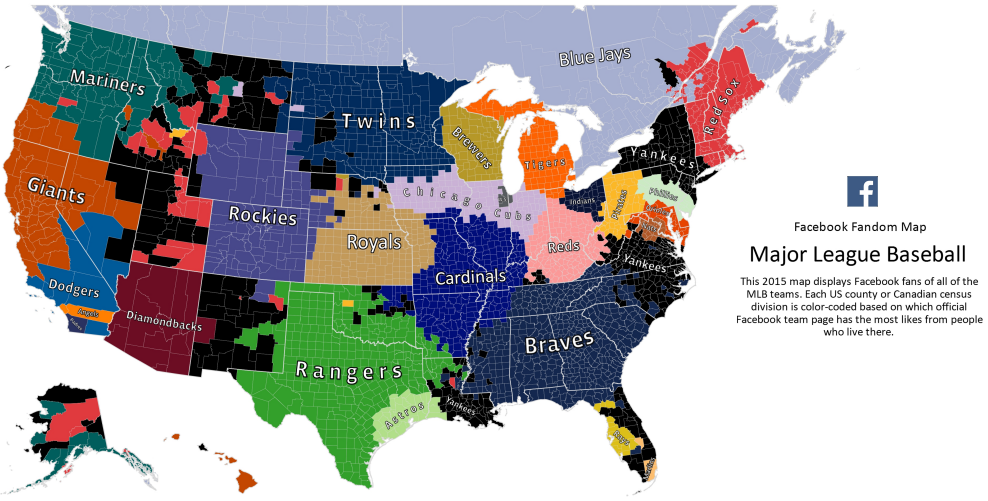

Figure 3. Map of the United States by MLB Fandom (Facebook 2015).

Given the relationship of fandom and geography, we can see that fan forums are an ideal source of regional speech online, using the following syllogistic logic: most posts on fan forums are by fans of the team. Most fans of the team are located in a coherent region surrounding the team's city. Therefore, most posts are from a coherent region surrounding the team's city. This reasoning justifies the use of fan forums as a data source in general, but they are especially useful for the variable of interest. After all, much of fans' discussion of a team's prospects centers on whether players need (to be) benched or whether the manager needs (to be) fired. This data source can thus both give us regional variation and a higher frequency of the variable we are interested in.

I focus on forums for teams based in Pittsburgh. There are three reasons for this: most importantly, the AEP is found in Pittsburgh. As mentioned above, fandom for the Pittsburgh Pirates is highly constrained geographically. Finally, Pittsburghese is highly enregistered and used to perform a Pittsburgh identity (Johnstone 2009). As such, use of the AEP should be both reasonably common and less stigmatized than it may be on forums for other teams. By sampling from Pittsburgh forums we should thus maximize our chances of finding AEP tokens. I collected tokens from two forums: the Unofficial Pittsburgh Pirates Message Board (Pittsburgh Pirates, MLB, http://www.pittsburghsports.net/viewforum.php?f=1) and 5th Avenue Forum (Pittsburgh Penguins, NHL, http://www.fifthavenueforum.com/forum/index.php). I manually searched these for tokens of embedded passives on February 2-4, 2017. Due to how the search function worked on each site,

the two searches were slightly different. The Pirates site did not allow searches for *need*, and instead *needs*, *needed*, and *needing* were searched. The Penguins site, by contrast, was searched for the lemma *need*. Combined, there were 17,504 hits with these searches. Of these, there were 525 tokens of either the EP or AEP.

Tokens were coded for forum and material that would be derived subsequently to the decision point: NEGATION (is negation present), AUXILIARY (does an auxiliary precede *need*), SENTENCE TYPE (is the sentence an interrogative), and CLAUSE TYPE (is the variant in a matrix, embedded, adjunct, or relative clause).[4] Online examples show the AEP is attestable in these environments (4), justifying their inclusion as possible conditioning factors.

(4) a. Our downstairs as I said earlier just doesn't **need cooled**, its cool even on super hot days with a fan going since its under ground for the most part. (Negation)
b. Bonino, Dumo, and Daley may **need picked** up next year. (Auxiliary)
c. Did he **need extracted** from something? (Interrogative)
d. I have been saying for a while that I think Correia **needs removed** from the rotation and replaced by either Lincoln or DCutch if they don't want to make any corresponding moves. (Subordinate Clause)
e. There are issues there that **need addressed**. (Relative Clause)

My intuition as a speaker who varies between the two forms is that the AEP is more likely to appear in relative clauses than in other environments. This result would stand in contrast to our hypothesis that no language-internal factors derived subsequent to the decision point will condition variation. While a significant effect of any language-internal factor would therefore be surprising, an effect of relative clauses may be the least surprising of possible effects.

5 Results

Given that the data sources were selected with the intention of maximizing occurrences of the AEP, the construction appears quite infrequently. Only 97 of 525 tokens (18.47%) were of the AEP. Before testing for effects of language-internal factors, I first consider whether the AEP occurs at different rates depending on forum. This is an especially important question because the search terms used when sampling each forum were slightly different. As seen in Table 1, the AEP appears at the same rate in both forums. This suggests that the different search terms did not affect the overall data. Furthermore, we may be tentatively confident that the two forums represent the same register of speech, and perhaps the same speech community.

	AEP		EP		Total
	n	Rate	n	Rate	n
Pirates	48	18.53%	211	81.47%	259
Penguins	49	18.42%	217	81.58%	266
Total	97		428		525

Table 1. Rate of AEP Use in Pirates and Penguins Forums.

Recall that we expect the AEP to occur at the same rate of roughly 18.5% in all environments. If this were not the case, we would have evidence that at least one language-internal factor conditions variation in embedded passives. This prediction is borne out for the AUXILIARY and SENTENCE TYPE factors (Table 2): for both factors, the rates of AEP use are quite similar across environments and chi-square tests show no significant effects.

[4] I assume the raising analysis of relative clauses (Kayne 1994, Law 2016). Under this analysis, the complementizer would Merge above *need*. Relative clauses therefore are environments derived subsequent to the decision point. Other analyses of relative clauses rely on a null operator that is quite low in the structure (see Bhatt 2002). In these accounts, the null operator would be low enough to condition variation.

	AEP		EP	
	n	Rate	n	Rate
Auxiliary	12	21.43%	44	78.57%
None	85	18.12%	384	81.88%

	AEP		EP	
	n	Rate	n	Rate
Question	4	16.67%	20	83.33%
Declarative	93	18.56%	408	81.44%

Table 2. Rate of AEP Use by Auxiliary Presence and Sentence Type.

At first glance, it appears that there may be an effect of NEGATION. The AEP occurs in 11.11% of negated tokens, but 23.71% of non-negated tokens (Figure 4). The difference does not reach statistical significance in a chi-square test (p=0.2583), most likely because of the small number of negated tokens (n=45). We cannot treat this as an instance of language-internal conditioning, although were this difference to be replicated in a larger sample, it would qualify.

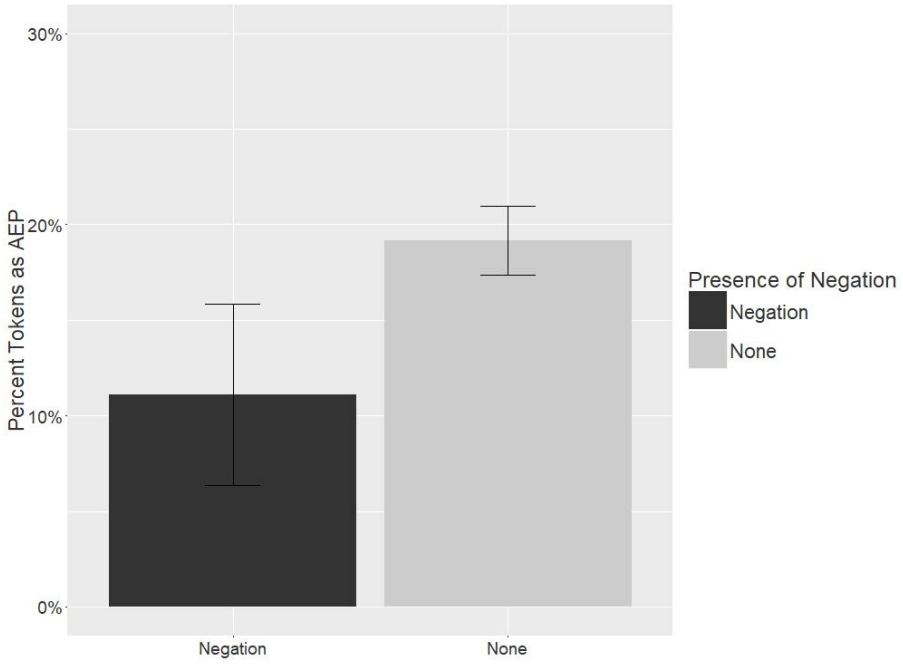

Figure 4. Rate of AEP Use by Negation. Error bars mark standard error.

The initial coding of CLAUSE TYPE considered whether the variant was in a matrix, embedded, adjunct, or relative clause. There were relatively few tokens in adjunct clauses (n=23), and the rate of AEP usage across clause type was approximately similar with the exception of relative clauses. As such, the factor was recoded to indicate whether the variant was in a relative clause or some other clause type. As illustrated in Figure 5, the AEP was more common in relative clauses (26.37%) than in other clause types (16.82%). Unlike NEGATION, there are sufficient tokens that appear in relative clauses (n=91) that the effect of CLAUSE TYPE is statistically significant in a chi-square test (p=0.0470). Such a high p-value is worth some caution, particularly because it would not hold up to corrections for multiple comparisons. Two factors somewhat mitigate this caution. First, this result matches speaker intuition. If we were to find one significant effect, we would expect it to be this one. The second reason is that the effect of clause type still appears when considering the data set as a whole: in a logistic mixed effects regression model following the formula in (5), a token is significantly more likely to appear in a relative clause than any other clause type (β= -0.5476, p=0.0438, Intercept=2.1495). The negative β for relative clauses reflects that the data was coded as a binary in which the AEP was 0 and the EP was 1. Again, there are no significant effects of any other factor considered. That the result still appears within this model is encouraging.

(5) Variant Type ~ Auxiliary + Negation + Sentence Type + Clause Type + (1|Forum)

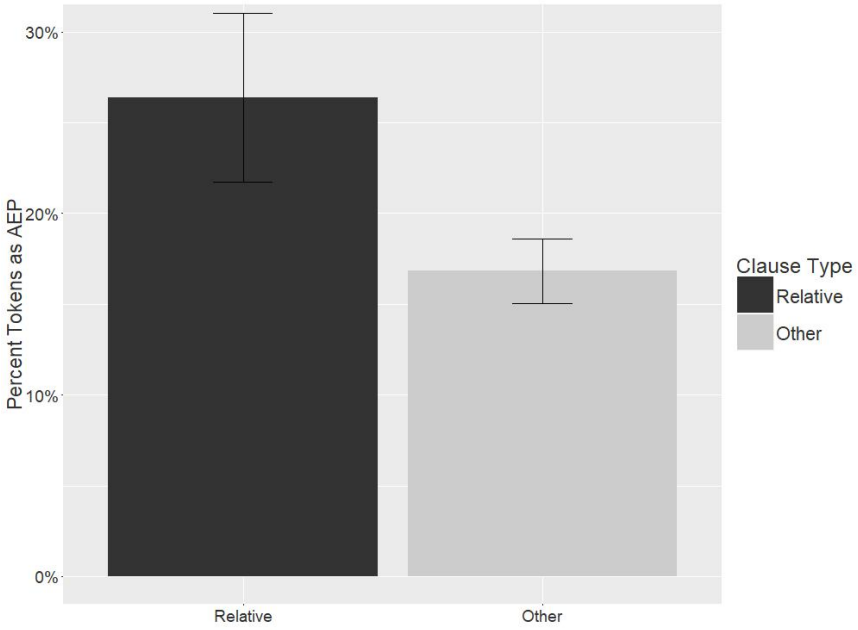

Figure 5. Rate of AEP Use by Clause Type. Error bars mark standard error.

6 Discussion

To summarize, we expected to find no language-internal factors that significantly predict the rate at which the AEP occurs. However, we appear to have found one: the type of clause in which the variant occurs. This is a problematic result for CG as described in Section 2. In this section, I note that relative clause effects have been found across dialects and variables. I then offer a revised CG, in which the decision point is located in Spell-Out, in order to account for the quantitative results.

6.1 Other Phenomena with Relative Clause Effect

Adger (2014) observes that data from Buckie, Scotland, concerning the Northern Subject Rule shows a relative clause effect similar to that found in this study. This is an unrelated morphosyntactic phenomenon in which English verbs with plural subjects variably take an *-s* ending. In the Buckie data the non-standard variant (*-s* ending) is more common in subject relative clauses than non-relatives. That relative clauses have the same effect in the same direction on unrelated phenomena in unrelated samples is noteworthy. While it is beyond the scope of this paper to pursue, studies of morphosyntactic variables in which a relative clause effect could occur should test for it. If it is a generalizable finding beyond these two examples, it is deserving of an explanation.

6.2 Accommodating Results within Competing Grammars Framework

It is certainly possible that the apparent relative clause effect would disappear given additional data. Supposing that it holds, the question at hand is how to account for the effect within CG. The problem is as follows: assuming a raising analysis of relative clauses (Kayne 1994), the relative clause structure is derived subsequent to Merging *need* into V. Because this is after our decision point, the relative clause structure is not visible to the syntax when we commit to one grammar. We therefore have a look-ahead problem: how does the presence of a relative clause condition variation if it is not available to do so when the variant is selected?[5]

[5]Under the null operator analysis of relative clauses, there is no problem; the operator is inserted prior to the decision point and can therefore condition variation. If you subscribe to this account, the following discussion is best read as a hypothetical: supposing that the effect of negation bears out across a much larger sample, how would we account for that?

Our goal is to address this effect without abandoning our assumptions about syntax and variation. Therefore, any proposal must keep rules and operations categorical and maintain that a derivation has a single output (Embick 2008). I likewise seek to maintain the view of syntax as a model of production in which derivations are constructed from the bottom-up. Assuming that CG is the correct approach to syntactic variation, how can we change the model and hold to these principles? There are three possibilities that spring to mind (7):

(7) a. Allow for speakers to recognize a derivation as leading to a stigmatized variant, crash it, and re-derive the sentence.
b. Model the syntax as derived top-down instead of bottom-up.
c. Select the variant later than previously thought.

I suggest that the third option is the most tenable. In order to formalize the first scenario, we would need to first suppose that there is some rate A at which the non-standard variant is initially derived and a rate B at which the speaker notices they are about to use it. The simplest implementation of these rates is to assume that the standard variant is categorically re-derived, in which case the true rate of occurrence would be A*B. Our data set shows that this scenario would require arbitrary rates of the speaker noticing a stigmatized construction per environment; there would apparently be a B_1 when the variant is in a relative clause and a B_2 in other environments. Such arbitrariness would make it difficult to conceive of a way to model syntactic variation moving forward.

The second option is also problematic, despite there being broader arguments for modeling syntax in a top-down manner (see Chesi 2015, for example). The reason is that to do so simply reverses the direction of the prediction we found to be unsupported. Environments derived after the decision point would still be predicted to not condition variation; however, such environments would now be below the decision point in the derivation rather than above. It seems likely that such environments may also condition variation, which would lead to the same problem we are attempting to solve.

I instead suggest that the decision point is in Spell-Out — later than currently proposed in CG. The effect of this is to suggest that initially both grammars in competition yield derivations for both variants. After this, one of the competing derivations is then selected for transfer to LF/PF in Spell-Out. Because the entirety of both derivations are visible to the syntax when the variant is selected, language-internal constraints derived both prior to and after the variant is derived can condition variation. Such a change, I argue, maintains our assumptions about syntax while accounting for actual variation data. Because one grammar derives one variant and another grammar the other, rules and operations still apply categorically within a grammar. At the same time, derivations still have a single output. While there is variation in which derivation is transferred to LF/PF, that derivation has one output.

Embick's (2008) review of approaches to morphosyntactic variation draws contrasts between CG and other approaches relying on probabilistic grammars (for example, Manning 2003). He notes that although they are quite different with respect to their theory of the grammar, these approaches nonetheless make the same surface predictions. This is because a single grammar with probabilistic variation in outputs is indistinguishable from probabilistic variation in grammars that yield single outputs. The above proposal brings the two closer together from a theoretical perspective with respect to where variation is located in the grammar without sacrificing the syntactic principles maintained in CG.

7 Conclusion

In this paper, I outlined the process by which an utterance is derived in the CG framework in order to identify a key prediction of it: environments derived subsequent to the decision point cannot condition variation. A variationist study of embedded passives did not support this prediction, as the type of clause that a variant occurs in was found to be a significant predictor of which variant occurs. To account for this, I offer a modified version of CG in which the decision point is late. In effect, this means that variant selection occurs in Spell-Out, after both variants have been derived.

I make this proposal acknowledging that the effect it addresses may be a result of a small sample size. Additional data, both on this variable and others, will help to refine the proposal as well as shed light on whether it is in fact necessary.

References

Adger, David. 2014. Variability and grammatical architecture. In *Linguistic Variation in the Minimalist Framework*, ed. M.C. Picallo, 179–196. Oxford: Oxford University Press.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics* 10:43–90.

Chesi, Cristiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44:65–89.

Davies, Mark. 2008-. *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. Available online at https://corpus.byu.edu/coca/.

Edelstein, Elspeth. 2014. This syntax needs studied. In *Micro-Syntactic Variation in North American English*, ed. R. Zanuttini and L. Horn, 242–267. Oxford: Oxford University Press.

Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35:355–392.

Embick, David. 2008. Variation and morphosyntactic theory: Competition fractionated. *Language and Linguistics Compass* 2(1):59–78.

Facebook. 2015. *Facebook fandom map: Major League Baseball*.

Fruehwald, Josef. 2012. Redevelopment of a morphological class. *U. Penn Working Papers in Linguistics 18.1: Proceedings of PLC 35*, ed. J. Fruehwald, 77–86.

Houston, Ann Celeste, 1985. Continuity and Change in English Morphology: The Variable (ing). Doctoral dissertation, University of Pennsylvania.

Johnstone, Barbara. 2009. Pittsburghese shirts: Commodification and the enregisterment of an urban dialect. *American Speech* 84(2):157–175.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K. Beals et al., 1–23.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English*. New York: Mouton de Gruyter.

Law, Paul. 2016. The syntax of Tagalog relative clauses. *Linguistics* 54(4):717–768.

Manning, Christopher D. 2003. Probabilistic syntax. In *Probabilistic Linguistics*, ed. R. Bod, J. Hay, and S. Jannedy, 289–342. Cambridge, MA: MIT Press.

Murray, Thomas E., Timothy C. Frazer, and Beth Lee Simon. 1996. Need + past participle in American English. *American Speech* 71(3):255–271.

Murray, Thomas E., and Beth Lee Simon. 1999. Want + past participle in American English. *American Speech* 74(2):140–164.

Murray, Thomas E., and Beth Lee Simon. 2002. At the intersection of regional and social dialects: The case of *like* + past participle in American English. *American Speech* 77(1):32–69.

Oseki, Yohei. Syntactic structures in morphological processing. Doctoral dissertation, New York University.

Tagliamonte, Sali, and Jennifer Smith. 2005. No momentary fancy! The zero 'complementizer' in English dialects. *English Language and Linguistics* 9:289–309.

Tenny, Carol. 1998. Psych verbs and verbal passives in Pittsburghese. *Linguistics* 36(3):591–597.

Wallenberg, Joel. 2013. A unified theory of stable variation, syntactic optionality, and syntactic change. Paper presented at DiGS 15, University of Ottawa.

School of English Literature, Language and Linguistics
Percy Building
Newcastle University
NE1 7RU
United Kingdom
daniel.duncan@ncl.ac.uk