

Topic-restricting *as far as* revisited

Robin Melnick
Pomona College

Thomas Wasow
Stanford University

Introduction

Rickford, Wasow, Mendoza-Denton, and Espinoza (1995a, henceforth RWME) investigated a phenomenon that had been condemned by prescriptivists for decades (Fowler, 1926; Bernstein, 1962, 1977; Harper, 1985), but had received only one previous brief discussion in a linguistics journal (Faris, 1962). The phenomenon is the use of topic-restricting *as far as* (AFA) with a following noun phrase (NP) but without a “verbal coda” like *is/are concerned* or *goes/go*. This is illustrated in (1), taken from RWME, with the position of the omitted verbal coda marked by “ \emptyset ”:

- (1) a. As far as ball technique and tactics \emptyset , this area is years behind other areas.
b. As far as filling out the details \emptyset , that isn't a problem.
c. He sounds just like the other kids, as far as general style \emptyset .
d. As far as the temperatures in the Bay Area tonight \emptyset , this is the way I see it.
e. I'll never quit as far as trying to solve the case \emptyset .
f. I need to know about lifestyle, because that's important, as far as where I'm going to be happy \emptyset .
g. As far as how he got shot \emptyset , we don't know yet.

The primary data for the analysis in RWME consisted of about 700 examples serendipitously encountered by John Rickford over a period of eight years, along with about 500 gleaned from online sources, including personal email and some of the few digitized corpora available at the time. RWME also included data from a questionnaire study that elicited acceptability judgments of AFA sentences.

The principal objective of RWME was to investigate what leads speakers and writers to omit the verbal coda. The paper considered both linguistic properties of the AFA phrase and social features of the speaker/writer, as well as whether the example was spoken or written. It also included a section discussing the history of the topic-restrictor use of AFA both with and without a verbal coda.

Collecting, coding, and analyzing the data used in RWME was slow and labor-intensive. The quarter century since that research was conducted has seen remarkable advances in the technologies available for such work. In this paper, we discuss some of the ways in which studies of linguistic usage can now be done faster, on a larger scale, drawing on data from more diverse populations, and with less effort. We illustrate these improvements through a new examination of verbal coda omission in the topic-restricting AFA construction.

By and large, our studies support the conclusions of RWME. Our methods also made possible some new observations about the AFA construction. The lessons of this exercise, both about the phenomenon under consideration and about the value of the new methods, will be summarized in our concluding section. Space constraints limit the degree of detail we provide here of our methods and analyses,

so we invite readers to review the supplemental materials available at <https://tinyurl.com/AFA-revisted-supplemental-pdf>.

A New Corpus Analysis

RWME's corpus was a mixed bag, drawing on examples that Rickford and others had noticed, plus what was readily available through computer searches at the time, and as noted in footnote 7 of RWME, a reviewer had raised a question about the representativeness of that corpus. The reviewer's concern—that the variant without a verbal coda (called "verbless" in RWME and below) might be more noticeable and hence overrepresented in the corpus—was successfully addressed by noting that the percentage of verbless examples in the data extracted from the Switchboard corpus was substantially **higher** than among the examples collected from conversations and reading. But that raises the question of whether the verbless variant was **under**represented among the informally collected examples. It is thus worth considering how representative of contemporary American usage the RWME collection of examples was.

To address this issue, we took advantage of an advance in the technology available for linguistic research, namely, large electronic corpora. In particular, we used the Corpus of Contemporary American English, or COCA (Davies, 2008-), a collection of about half a billion words of writing and speech, divided by genre into five parts of roughly equal size: speech (radio and television transcripts), fiction, magazines, newspapers, and academic writing.

Using this resource and both existing and custom-developed software for extraction and annotation, it was possible to make a substantially larger corpus study with more confidence of its representativeness. And this required far less time and effort than the collection and annotation of RWME's corpus. We do recognize that our software-automated extraction and annotation relies in part on a number of heuristics, and the resulting annotated corpus of AFA examples hence contains some degree of error. Moreover, the corpus lacks some of the demographic information (age, sex, and ethnicity of the speaker/writer) that RWME considered. On the other hand, COCA provides useful information (notably, the five-way genre classification and the year of production) not available in RWME's corpus.

Extraction

Extracting the examples we wanted from electronic corpora was facilitated by the fact that for the present study, tokens can be identified using a simple search for sentences containing the string *as far as*. In the version of COCA we employed, this resulted in 16,324 hits. However, while many of these are the topic-restricting uses of AFA of interest here, many represent other uses of AFA, including:¹

- Expressions of distance or extent, for example: *The idea goes back at least as far as Socrates.*
- Formulaic collocations, such as: *as far as we know*, *as far as I can tell*, *make it as far as*, and *as far as the eye can see*.

¹ Except where some other source is noted, all examples are from COCA.

- What RWME called “perspective” uses, in which the AFA phrase identifies whose point of view is being expressed, for example: *It was sweet but, as far as I was concerned, useless*. Both the verbless variant and forms of *go* are impossible in these cases.

RWME also note that a verbal coda is obligatory when the NP following AFA is a nominative personal pronoun: *November payments have been paid, as far as I am concerned*/* \emptyset . They go on to suggest (p. 108) that “this restriction appears to be a subset of the perspective restriction”. In other words, all examples with a nominative pronoun immediately following *as far as* are perspective cases. Examination of a few hundred COCA examples convinced us that the converse is also almost true—that is, almost all perspective cases have a pronoun, usually *I*, immediately following AFA. While this is not a categorical restriction, it is a good enough approximation to allow us to employ a simple heuristic for filtering out the perspective cases, namely, excluding all examples in which AFA is immediately followed by a nominative-case personal pronoun.

While this heuristic also eliminates most of the examples with the formulaic collocations, not all of them have pronouns immediately following AFA (e.g., *as far as the eye can see*) so we added several additional filters aimed at these collocations, including: eliminating examples with forms of the verbs *determine*, *know*, *see*, or *tell* within five words after AFA; examples with *is possible* or *was possible* immediately following AFA; examples with *eye* or *eyes* within two words after AFA; and examples with forms of the verb *make* immediately followed by *it as far as*.

Our heuristic for excluding examples in which AFA is used to express distance or extent involved filtering out examples in which a key word close to the occurrence of AFA indicates that it is probably being used in this way. Hence, we filtered out sentences satisfying any of the following conditions:

- AFA immediately follows *almost*, *back*, *from*, *nearly*, or a form of *get*;
- AFA is preceded within two words by a form of one of the following verbs (usually used intransitively): *come*, *flee*, *fly*, *go*, *proceed*, *range*, *reach*, *rise*, *run*, *stretch*, *travel*, *walk*, or *wander*;
- AFA is preceded within four words by a form of one of the following verbs (usually used transitively): *bring*, *drive*, *extend*, *follow*, *move*, *push*, *propel*, *raise*, *ship*, *stretch*, or *take*;
- AFA is followed within four words by *away*.

In order to exclude examples like *as far as my strength will allow me*, we also added a filter ruling out examples in which the first verb in the complement of *as far as* is tensed (other than a form of *go* or *be concerned*). Implementing this involved taking advantage of another technological advance since the period of RWME, the improved accuracy of software-automated syntactic parsing. Here we applied the Stanford NLP Group’s statistical PCFG parser (Klein & Manning, 2003) to our sample of AFA-inclusive sentences. The software was unable to produce any parsed structure for 177 of our sentences due to length or other software errors, so we checked these by hand, excluding those that were not topic-restricting AFA examples.

After filtering, we were left with 6,824 AFA examples from COCA. We hand-checked a random sample of 500 tokens from our pre-filtering corpus to determine how well our heuristics identified the examples we wanted to analyze. We found 38

cases (7.6%) in which our heuristics did the wrong thing—that is, they either filtered out topic-restricting uses of AFA (22 examples) or failed to filter out non-topic-restricting uses (16 examples).²

While our corpus collection of AFA examples is over five times the size of RWME's, it is worth noting that topic-restricting AFA remains nevertheless a relatively rare phenomenon, occurring in COCA less than once in every 65,000 words, thus providing a motivation for employing such large-scale corpora. As a basis of comparison, Melnick's (2017) corpus study of other syntactic alternations found heavy NP shift—an already reputedly rare construction—occurring at over ten times that rate. The occurrence of AFA varies considerably by genre: it is most common in speech, where it is about three times as frequent as in newspapers, about four times as frequent as in fiction and magazines, and about five times as frequent as in academic writing.

Annotation and Results

Our 6,824 examples were annotated for the following features: type of coda (*be concerned*, *go*, or none), genre, year produced (ranging from 1990 to 2012), position of AFA in the sentence (initial or non-initial), and length (in words) of the phrase following AFA (not including the verbal coda, when present). This coding was automated, making use of both custom-developed and existing software tools (see discussion in the Conclusions section below). Recognizing, however, that the precision of automated parsing, in particular, remains imperfect, the annotations, like the extraction, inevitably contain some degree of error. (See again <https://tinyurl.com/AFA-revisited-supplemental-pdf> for details.)

Our goal was the same as RWME's: to find factors that correlate with choice of coda and to consider possible explanations for the correlations we found. To this end, we fit a multivariate logistic regression model to our data using R software (R Core Team, 2016),³ predicting coda absence (none) versus presence (combining *be concerned* and *go* cases), based on the several factors mentioned above for which we annotated the data and interactions among them. We replicate a number of RWME's findings, and found nothing in direct conflict with any. On some points, we can add new observations. In what follows, we cite some highlights.

First, as Figure 1 illustrates, increasing the length of the post-AFA noun phrase reduces the use of both types of verbal coda. Controlling for other factors, regression confirms significance of the greater preference for codaless production as NP length grows (estimated beta coefficient $\beta=0.895$, $p<0.0001$). One natural explanation for this correlation is that it is a memory effect⁴: the greater the distance between AFA and the position where the coda would appear, the more likely it is that the speaker or writer will find it difficult to structurally tie a coda

² In the terminology of computational linguists, our precision is 91.8%, with recall of 89.1%, yielding an F-score of 0.9040.

³ The availability of this software is itself yet another technological advance we were able to utilize in our revisited study. (See discussion in conclusions section below.)

⁴ The idea that the sensitivity of linguistic phenomena to length of expressions can be explained in terms of memory is an old one. See Melnick 2017 for a review.

back to the preceding AFA. That is, they may forget. This account would lead one to expect a higher rate of coda omission in speech than in writing, and our statistical models confirm that this differential effect is significant ($\beta=0.277, p<0.0001$). Somewhat surprisingly, though, the rate of increase of coda omission with increased post-AFA NP length is statistically indistinguishable across three of the four written genres sampled via COCA, with the exception, academic writing, largely accounting for the spoken/written statistical difference among our data, as suggested by Figure 2. In the regression model, this is reported as an interaction between genre and length of phrase following AFA, with this effect rising to significance only for academic writing ($\beta=-0.489, p=0.0001$).

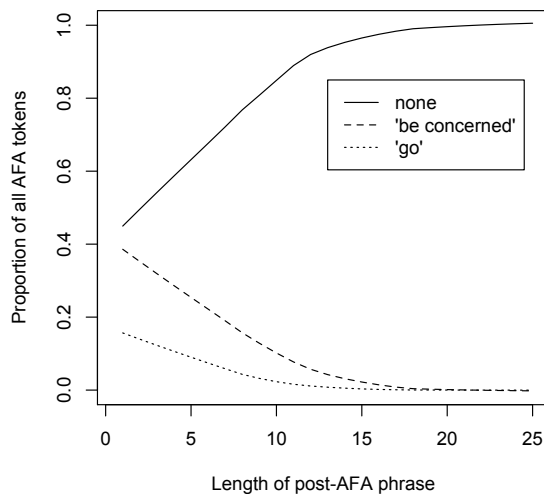


Figure 1: Distribution of coda forms with increasing length of post-AFA phrase.

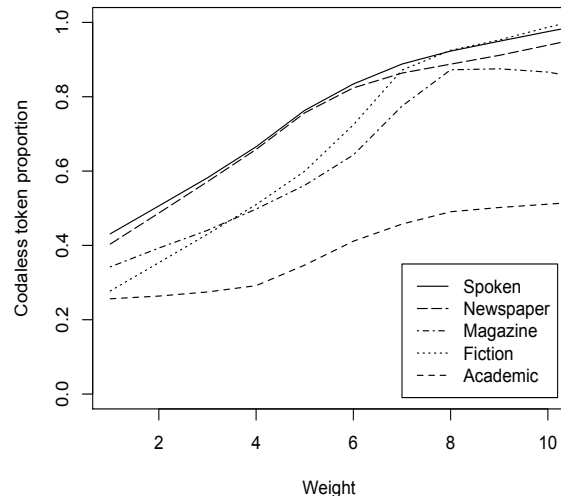


Figure 2: By-genre coda absence with increasing length of post-AFA phrase.

RWME found that sentence-initial AFA phrases had a higher rate of coda omission than non-initial ones, speculating (following Faris, 1962) that this could be the influence of *as for*, which also serves as a topic-restrictor and occurs only sentence initially. In our corpus, we replicate this initial-AFA higher rate of coda absence only for the spoken portion of our data (in the regression model, this is the interaction of sentence-initial AFA and spoken genre, $\beta=0.295, p=0.0011$). In fact, for all four written genres, we found the reverse correlation. Since RWME did not investigate the influence of AFA position separately for speech and writing, we do not know whether the same interaction was present in their data, two-thirds of which were spoken. It seems plausible to us that the influence of (or confusion with) *as for* would be greater in speech.

Perhaps the most interesting difference between the results of the two corpus studies concerns change over time. RWME's sample was neither big enough nor collected over a long enough period to test directly for change in AFA usage. Instead, they suggested that coda absence was increasing based on "apparent time", that is, the age of the speaker or writer. Although our corpus covers only 23 years (1990-2012), our analysis showed a clear, significant trend of increasing coda omission over time (main effect of year, $\beta=0.234, p<0.0001$) in all genres except fiction

(interaction of year and fiction genre in the regression model was significant, $\beta=-0.239, p<0.0047$), as illustrated in Figure 3. As Figure 4 shows, the overall increase of coda absence over time has come entirely at the expense of *be concerned*, with the rate of *go* used as the coda in the AFA construction remaining constant.

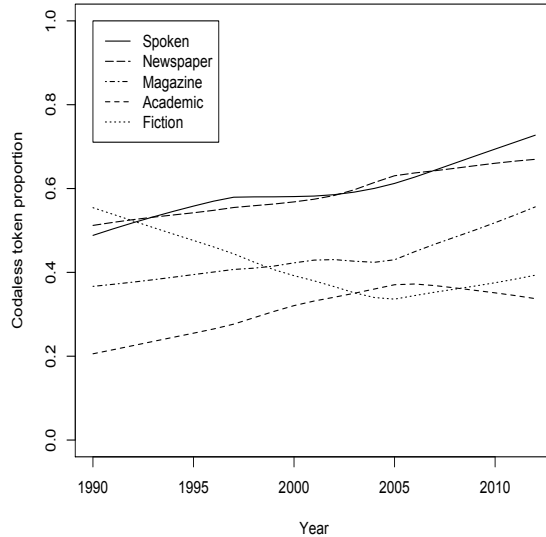


Figure 3: Ø-coda increase over time, by genre.

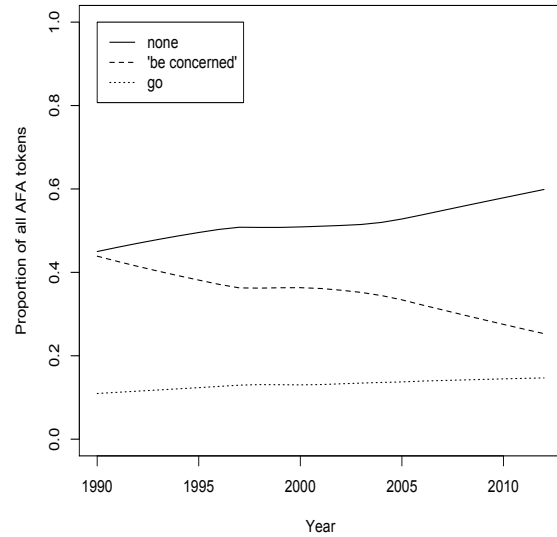


Figure 4: Coda form trend over time.

The Acceptability Judgment Survey

The conclusions of RWME relied primarily on their corpus data. Their survey of acceptability judgments was motivated in large part by the reactions of people the authors talked to, many of whom judged topic-restricting AFA without the verbal coda totally unacceptable. Some (including some linguists) even claimed never to have heard it used. Given the numerous examples without verbal codas in the authors' collection, they included the survey to test just how strong and widespread the negative assessment of such sentences was.

In brief, the RWME survey asked respondents to rate each of 13 *as far as* sentences plus seven fillers for “acceptability” (as opposed to “grammaticality”), using a four-point scale (from 1 for totally unacceptable to 4 for fully acceptable). The stimuli were all attested examples from their corpus, slightly edited in some cases. Participants came via two different channels: 101 individuals, mostly Stanford undergraduates, took the survey on paper in person, and another 79 responses came as replies to an email solicitation (Rickford et al., 1995a:105).

RWME did not do any statistical modeling of their questionnaire data, but they made some general observations about the responses. In particular, they noted that verbless sentences received lower scores than those with a verbal coda, and that younger participants (those aged 30 or below) tended to give higher scores to verbless examples than older participants.

As with our revisiting of the earlier corpus study, here again we sought to conduct a companion judgment study employing technological advances unavailable at the time of RWME. Most prominently for experimental work, this means

crowdsourcing, though we also took advantage of software-based survey apparatus to improve on a number of structural aspects of the experiment.

First, while we reused RWME’s original 13 test items, this time each sentence was presented not just as it had been found originally attested, but rather in each of three conditions (dynamically selected at random by the survey software), varying the accompanying verbal coda seen by the participant (*go*, *be concerned*, or *none*). Hence, each participant saw each test sentence, but saw each particular item in just one of the three possible conditions. The overall order of items presented was similarly randomized separately for each individual. Combined, these changes sought to minimize the effects of extraneous factors that might affect acceptability judgments (e.g., plausibility, lexical frequency, order of experimental presentation, and so on). The number of fillers was also increased from 7 to 27 to better disguise the target of our inquiry. Each participant thus evaluated 40 sentences, a task that took respondents an average of about 6 minutes. Participants were paid \$1.50 via the Amazon Mechanical Turk crowdsourced-task platform (<https://www.mturk.com/mturk/welcome>).

The crowdsourced approach made available to us a virtually unlimited number of participants. We ultimately chose to take on 400 “workers”, more than double the size of the original RWME survey, with the nature of the crowdsourcing platform being that respondents self-select participation. The 400 participants included 234 men and 166 women, ranging in age from 18 to 70 (mean=34.9). We also collected data on education levels and ethnicity, with Table 1 summarizing these responses.

1	Some high school, no diploma	296	White
48	High school graduate (diploma or equivalent)	27	Hispanic or Latino/Latina/Latinx
107	Some college, no degree	31	Black or African American
64	Associate or technical degree	2	Native American or American Indian
148	Bachelor's degree	36	Asian
32	Graduate or professional degree	8	Other

Table 1: Demographic survey responses—education and ethnicity.

We excluded data from four participants whose first language was not English and from 23 participants who finished the survey in under three minutes (including demographic questions), which we deemed to be too fast to have honestly engaged with the judgment task. This leaves 373 participants, producing 4,849 acceptability scores for the AFA sentences.

We fit a mixed-effects regression model to these data, predicting score (1–4) from the type of coda seen, together with linguistic properties (position of the AFA phrase and length of NP following AFA) of the stimuli and demographic properties (age, sex, ethnicity, and education level) of the participants. We also looked at interactions among these factors (fixed effects), and further employed random effects for both item and participant, allowing the model to take into account any variation in baseline rates of acceptability among individuals or from one stimulus to another (beyond the effects of the fixed-effect factors noted above). Stepwise optimization of model parsimony—that is, removing factors one at a time until all those remaining significantly improve the fit of the model to the data—eliminated

each of our demographic factors, indicating that variation in these elements play no significant effect on judgments of naturalness for this construction.

Examining the factors retained in the model, we found, as RWME had, that stimuli without a coda received significantly lower scores than those with a coda ($t=-6.7$, $p<0.0001$). The model also favors coda absence with AFA in sentence-initial position over non-initial position (in the regression model, this being the interaction of factors for AFA position and stimulus coda form, $t=6.2$, $p<0.0001$); this fits with the higher rates of coda absence in RWME's corpus study and in the spoken portion of our corpus study. Finally, our model also shows that participant ratings increasingly favor the verbless variant when the NP following AFA was longer (i.e., the interaction of NP length and stimulus coda form, $t=6.5$, $p<0.0001$); this converges nicely with both RWME's and our corpus studies, in which the rate of coda absence correlated with NP length. Figure 5 illustrates this last interaction.

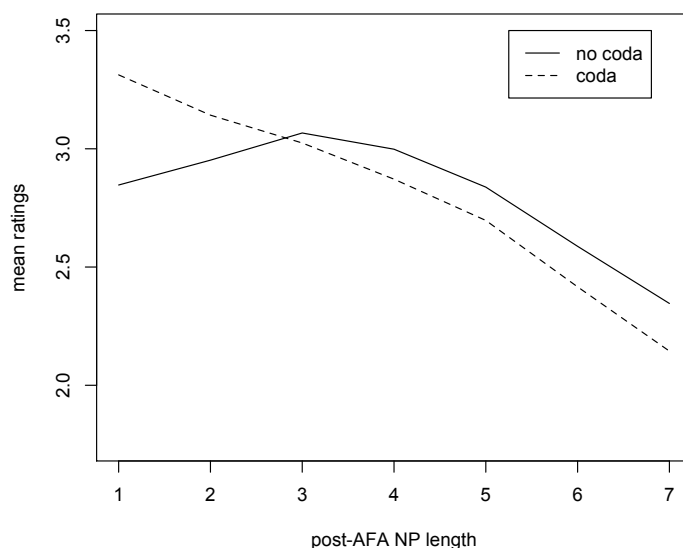


Figure 5: Mean scores with and verbal coda with increasing NP length.

The one place where the results of our questionnaire study clearly diverge from RWME's findings concerns age. Our model displays no significant interaction between participant age and ratings given to verbless examples, whereas the age data from RWME, both in the corpus and the questionnaire, had suggested that younger people are more tolerant of coda omission.

Conclusions

We had several ends in mind in revisiting the AFA construction. First and foremost, we wanted to illustrate how the computational tools developed in the past quarter century can facilitate the study of language usage. Second, we wanted to determine whether RWME's conclusions would be supported by a similar pair of studies using the new tools. And third, we hoped to learn something new about the construction.

Taking these in reverse order, the following novel observations emerged from our studies:

- Topic-restricting AFA is a rare construction. Hence, it is not surprising that it took Rickford eight years to collect about 700 examples.
- The verbal coda is being omitted at an increasing rate, in both speech and writing. RWME conjectured that this was the case on the basis of apparent time, and our corpus study found direct evidence of the change in real time.
- The rate of usage of the coda *go* has remained constant, revealing that the increase in coda omission has come entirely at the expense of *be concerned*.
- Judgments of the naturalness of coda omission do not seem to be affected by the sex, ethnicity, or level of education of the speaker. The absence of an effect of education surprised us, particularly given the condemnation of the verbless variety by prescriptivists. RWME found no effect of sex or ethnicity on their corpus data, however, which fits well with our survey results.
- Judgments are affected by length of the NP following AFA and by the position of the AFA phrase. RWME noted correlations of length and position in their usage data, but did not test for them in their questionnaire study.

Turning to a comparison of our findings with RWME's, our work generally supported their conclusions. In particular, like them, we found that:

- Longer NPs following AFA correlate with higher rates of coda omission.
- In speech, initial AFA phrases have higher rates of coda omission than non-initial ones. RWME made the more general claim that initial position of the AFA phrase favored coda omission, but they did not check separately for speech and writing.
- Sentences containing verbless AFA phrases are generally rated as less natural than ones containing AFA phrases with a verbal coda, despite the fact that the verbless variant is the most common one in the corpus data. Divergences between judgment data and usage have been noted elsewhere (e.g., Asudeh, 2011), raising questions about what theories based entirely on judgments are theories of; we will not venture into this minefield here.

The most prominent point of divergence between our findings and those of RWME is with regard to the effect of age on judgments: they found a negative correlation, where we failed to surface any significant effect. It is worth noting here that our result was based on a multivariate statistical model in which age was a continuous variable, whereas theirs was based on the simple observation that the average ratings of survey respondents 30 and under tended to be higher than those of respondents over 30. It is possible that the noted difference in results is a consequence of the difference in methods.

Finally, we return to our point about the new tools for usage research. The availability of large electronic corpora such as COCA—among many others in the increasingly active discipline of corpus linguistics—makes it possible to conduct in minutes searches of usage patterns that once would have required years. In the early 1990s, when RWME's research was underway, there were some electronic corpora available, but they were hundreds of times smaller than those now available. Further, many contemporary corpora have been annotated in ways that can be very helpful in searching. COCA, for example, comes pre-tagged for part-of-

speech, which permits far more general searches than would be possible using only word strings.

Moreover, the ready availability of increasingly accurate software for automated part-of-speech tagging, syntactic structure and dependency parsing, sentiment analysis, and other features makes it possible to rapidly apply such annotations to virtually any electronically available collection of data. In the present work, for example, we applied syntactic parsing software to our initial extraction of AFA sentences from COCA, a corpus that comes POS-tagged but without syntactic structure annotation. For investigations into grammatical patterns of usage, such parsing information is especially useful. Searching such parsed corpora in turn requires search tools that recognize the notation used in the parses. In the present studies, we employed Tgrep2 (Rohde, 2005), designed to search parse trees formulated according to the Penn Treebank format (Marcus et al. 1993). After parsing our initial COCA extraction, we were able to use the parse information both in the filtering and annotation stages of our corpus study. Extracting and annotating the quantity of data in our corpus study would have taken hundreds of person-hours without these tools.

With regard to our questionnaire study, the principal enabling tool was crowdsourcing. We were able to attract a far larger and more diverse set of participants than would have been possible using leaflets or personal contacts. It was also extraordinarily fast: We had over 200 responses in under an hour, with the remainder of our our target number of 400 responses in less than two hours.

Data from both of our studies were also analyzed using statistical software that did not exist when RWME's research was conducted. While VARBRUL, which RWME used for their corpus study, was a sophisticated multivariate tool at the time it was created (1974), the R programming language (see Baayen, 2008, for a review of its application to linguistic data analysis) made possible some analytical choices that would have been difficult or impossible with VARBRUL. In particular, we were able to treat year of production in the corpus study and age of participant in the judgment study as continuous variables, rather than requiring binning; we could also easily check for interactions among variables, which proved important for some of our findings; and lastly, it supports "random effects", allowing correction for baseline biases by participant, stimulus, or other elements not easily accommodated using a small number of fixed categorical choices.

Summing up, thanks to modern technology, we were able to conduct updated versions of RWME's investigation of the AFA construction more quickly and at a larger scale. Our studies largely supported RWME's conclusions, and we added some new observations. If the price of the improved efficiency of studies like ours is that computerizing data extraction and annotation often requires tolerating a degree of error in the data set, the general convergence of our results with RWME's indicates that the advantages are worth this price.

References

Asudeh, Ash. 2011. "Local Grammaticality in Syntactic Production". In E. Bender & J. Arnold (eds) *Language from a Cognitive Perspective*. Stanford, CA: CSLI Publications.

Baayen, Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge, UK: Cambridge University Press.

Bernstein, Theodore. 1962. *More language that needs watching*. Manhasset, NY: Channel.

Bernstein, Theodore. 1977. *Dos, don'ts and maybes of English usage*. New York: Times Books.

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

Faris, Paul. 1962. "'As Far as Halfbacks, We're All Right'." *American Speech* 37.3, 236-238.

Fowler, Henry Watson. 1926. *A dictionary of modern English usage*. Oxford: Press.

Dan Klein and Christopher D. Manning. 2003. "Fast exact inference with a factored model for natural language parsing." In *Advances in Neural Information Processing Systems 15 (NIPS 2002)*, Cambridge, MA: MIT Press, pp. 3-10.

Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. 1993. "Building a large annotated corpus of English: The Penn Treebank." *Computational Linguistics*, 19(2), 313-330.

Melnick, Robin. 2017. *Consistency in Variation: On the Provenance of End-Weight*. Unpublished Stanford University dissertation.

Morris, William and Mary Morris (eds). 1985. *Harper dictionary of contemporary usage*. New York: Harper and Row.

R Core Team. 2016. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Rohde, Douglas. 2005. "Tgrep2 User Manual". URL <https://tedlab.mit.edu/~dr/Tgrep2/tgrep2.pdf>