Rethinking the teaching of grammar from the perspective of corpus linguistics*

Chongwon Park · Elizabethada Wright · David Beard · Ron Regal
(University of Minnesota Duluth)

Park, Chongwon, Elizabethada Wright, David Beard, and Ron Regal. 2019. Rethinking the teaching of grammar from the perspective of corpus linguistics in Korean. Linguistic Research 36(1), 35-65. Despite calls from many composition and rhetoric scholars for instructors of writing to stop teaching prescriptive grammar, a vast number of handbooks intended for college writing classes encourage this tradition. For example, Hacker’s Pocket Style Manual has a section on grammar with instructions for students on how to write appropriately. While Hacker may not intend for her instructions to be taken as dictums, they often are, and much time is spent in many classrooms teaching students these rules of grammar. This article uses the Corpus of Contemporary American English (COCA) to support the calls from composition and rhetoric scholars that prescriptive instruction in grammar is more a hindrance to writing instruction than a benefit. Focusing on a few specifics from frequently used grammar handbooks and illustrating how big data shows the “rules” are incorrect at best, this article argues that, just as scholars of English have begun using big data to better understand literary history, scholars of rhetoric and composition might better understand how to help students to write by understanding patterns within big data. Certainly, this argument recognizes that “common usage” may not necessarily be the most eloquent usage. In making this argument, this article focuses on the [neither or either of X] + Verb construction, where the Verb may have either a plural or a singular form. Our findings illustrate that the “real world” writing is different from what textbooks dictate, and we suggest the data-driven observations need to be appropriately incorporated in writing classes.

(University of Minnesota Duluth)

Keywords indefinite pronouns, writing instruction, writing handbook, logistic regression statistic, corpus linguistics

* We extend our thanks to the reviewers and to the staff of the journal, as well as the University of Minnesota Duluth, which supported this project through Dean’s Excellence Funds and Chancellor’s Small Grants.
1. Introduction

As a linguist, one of the authors of this essay has become accustomed to answering his phone warily. Many phone calls, especially from faculty in other departments, are not about collaboration, curriculum, or intellectual work—they circle questions of style, mechanics and grammar. Colleagues in other units want direction about the “correct” way to write a sentence. One example that often arises is about subject-verb agreement in sentences using “neither” as indicated in (1) and (2). Sentences beginning like this have stymied our colleagues:

(1) Neither of the theories adequately explain the phenomena under analysis...
(2) Neither of the theories adequately explains the phenomena under analysis...

Both of these sentences appear to be acceptable in conversational usage as well as in writing. But our colleagues do not trust their intuition about grammar.

Readers of *Linguistic Research* will recognize these behaviors as reflecting “writing anxiety” as articulated by Cheng (2004) and developed by Jee (2018). “Writing Anxiety,” in these second-language contexts, consists of Somatic Anxiety, Cognitive Anxiety, and Avoidance Behavior. Presumably “avoidance behavior” [rewriting the sentence without the “neither” construction] doesn’t seem possible. Even writing in their first language, these doctoral-holding researcher experience “somatic” and “cognitive anxiety” high enough, they feel the need to call our linguist for help.

Our colleagues, our fellow researchers, have become accustomed to imagining that there are “correct” ways to construct a sentence, “correct” ways that they feel they do not know, and fear of making a mistake has paralyzed them. Perhaps this fear arises because they were literally or metaphorically slapped with a ruler for grammatical errors when they were younger. Perhaps this fear arises because they bathed in rivers of red ink on papers in high school, college, and graduate school. Whatever the reason, these kinds of questions make both the linguist and the writing teachers who are collaborating at the heart of this article wince; however, we wince for slightly different reasons. As writing teachers, two of us wince because we are aware of research that shows teaching correctness provides little benefit to students and, in fact, often harms them. As
a linguist, one of us winces because of the conflation of the word “grammar instruction” with the idea of grammatical correctness, while in fact, as linguist Nelson Francis (1954) and writing teacher Patrick Hartwell (1985) have explained, grammatical correctness—or “school grammar”—is only one, impoverished understanding of the complex term.

We want students [and our colleagues] to write well and understand that, to do so, they must have facility with the complexities of the language with which they write. We follow the lead of people like Lobeck (2000) who argue that the teaching of grammar in a writing class is not problematic; the teaching of grammatical “correctness” is. Lobeck argues, that a linguistic approach to grammar better teaches how language works. Martha Kolln and Chris Hancock agree, suggesting that the problem stems from writing instructors’ and linguists’ “failure to reach across the divides of our disciplines” (2005: 22). We rectify some of this failure by bringing research in linguistics together with scholarship in rhetoric and composition, a synthesis that occurs in three steps.

First, this article first overviews the writing field’s unfavorable view of grammar pedagogy, then emphasizes the lack of training provided to both high school and university writing instructors in linguistics. With such limited training in what grammar actually is, it is no surprise that writing teachers across the world become dependent on grammar handbooks that dictate what is acceptable in grammatical construction and what is not.

Then, this article illustrates that what is taught in composition classes does not reflect linguistic usage.1 We survey the most popular textbooks in American colleges and universities for the rules addressing the queries discussed in the first paragraph of this article, the use of singular and plural verbs with indefinite

1 Our work echoes the research of Sterling Andrus Leonard (1935) as well as that of his student Robert Pooley (Pooley 1933), who noted that consensus about grammatical correctness among teachers was not grounded in usage. Similarly, 1930 NCTE president Ruth Mary Weeks noted that “Our composition texts are cluttered with requirements no longer observed in the current speech and writing of educated men” (1935) and William Ellery Leonard called for “a realistic presentation of linguistic facts, as distinct from the stupid traditions and pedantic artificialities in so many of the so-called rules of grammar in the textbooks” (1933: 7). In 1988 and 2008 respectively, Robert Connors and Andrea Lunsford (1988) and then Andrea Lunsford and Karen Lunsford (2008) conducted studies on what teachers saw as errors in writing. Both studies found that “teachers vary widely in their thinking about what constitutes a ‘markable’ error” (2008: 784). There is a long history of asking writing teachers to approach grammar more humbly, and we join that history.
pronouns “neither” and “either.” We find fairly widespread banning of construction (1): *Neither of the theories adequately explain the phenomena under analysis...*

We then compare that proscription to contemporary usage through an analysis of the Corpus of Contemporary American English (COCA), an on-line database of 460 million words of text, showing the words’ usages, frequency, collocation with other words, etc. Updating and adding to its total constantly, COCA pulls words from and “is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts” (COCA). Using a well-established statistical method, we demonstrate that both of the examples (1–2) addressed above are acceptable in writing in a statistically significant way. This widespread usage of both occurs despite the banning of the first example by almost all handbooks and textbooks.

We are not arguing a “crowd sourcing” version of grammatical structures, in which widespread usage within COCA determines acceptability of a construction. Instead, the goal of this article is to point out that handbooks discussing “correctness” do not consider what cognitive linguists understand about how language, cognition and context shape meaning. In this particular case of indefinite pronouns, handbooks do not recognize the workings of metonymy. The article concludes by discussing what Cognitive Grammar illustrates about human linguistic choices: textbook notions of correctness reflect neither common usage nor cognitive effectiveness. A writing pedagogy informed by Cognitive Grammar will produce writers who communicate with lower anxiety and greater effectiveness.

2. Overview of the place of grammar in teacher preparation

Many high school and university writing teachers are given insufficient instruction in linguistics, including grammar. For example, Robert Tremmel

---

2 We could have focused on another dictum from handbooks, such as prohibitions of split infinitives, sentences ending with prepositions, the expletive constructions, confusions of “which” and “that,” and so forth.

3 Although COCA contained 460 million words at the time this data was collected, it is significantly larger at the time of publication—and it most likely even larger at the time of the audience’s viewing of this article.
Rethinking the teaching of grammar from the perspective of corpus linguistics

(2001) argues that since the mid-twentieth-century, most writing teachers have been trained with “one methods course in which they [the instructors of the teachers] attempted to cover the whole discipline” (2001: 13). Wright (2017) and Gordon (2005) also argue English education programs marginalize writing instruction. Gordon also observes that many teachers “avoid[ed] the need for grammatical content knowledge by approaching language teaching from a predominantly sociological, rather than linguistic perspective” (2005: 50). Therefore, a large number of writing teachers lack grammatical content knowledge and lack confidence in the teaching of grammar (Blake and Shortis 2010: Myhill, Jones, and Watson 2013). According to Harper and Rennie, pre-service teachers “showed limited understandings in their ability to analyze the parts and structure of sentences, and their knowledge of metalinguistic terms did not seem to extend past the basic concepts of ‘noun,’ ‘verb,’ and ‘adjective’” (2009: 27). As a result, writing teachers end up learning how to teach grammar on the job.

Because teachers are so marginally prepared, they become dependent on handbooks. As Smagorinsky, Wilson, and Moore (2011) note, “The teaching of writing and grammar . . . typically begins with marginal preparation and is then shaped by school settings in which both the available textbooks and the imperatives of test preparation conspire to mediate teachers’ conceptions of instruction in terms of fragmented parts rather than fluid wholes” (2011: 266). In other words, teachers begin disadvantaged and are constrained by the tools, especially the handbooks, they assign to students (Smagorinsky 2010).

The study we present below supports the idea that this weak training disadvantages teachers—and their students. Despite the lack of training, grammar instruction still occurs in many middle-schools, high-schools, and universities. There is good reason for teachers to maintain this practice in the United States. Smagorinsky, Wilson, and Moore tell us that “writing instruction is heavily geared toward satisfying both internal and external mandates for producing form-centered texts” (2011: 266). Similarly, Kolln and Hancock (2005) observe that new standards resulting from No Child Left Behind demand pre-college students have knowledge of grammatical correctness, and important standardized exams such as the SAT test such knowledge. But that grammar may be taught badly, as teachers are constrained by their tools, most notably
their “grammar” handbooks.

What handbooks and well-meaning grammar guides offer is problematic. To make this claim we will compare a corpus analysis of actual use, in English, of the indefinite pronoun against the handbook rules, guidelines and strictures set out for students in using indefinite pronouns. In so doing we show that what is used in the language of spoken, fiction, popular magazines, newspapers, and academic texts is not what is prescribed in the textbooks used in writing classrooms.

3. Methods, Part I: Constructing a corpus for grammar handbooks

To understand the relationship between [a] the information provided in grammar handbooks commonly used in writing classes and [b] language’s common usage, this project first collected syllabi from writing courses across the United States to see 1) what texts these classes required and 2) what instructions the texts provided concerning indefinite pronoun construction.

Readers of Linguistic Research are familiar with analyses of textbooks; for example, Kim and Paek (2015) conduct a close analysis of five textbooks for teaching English in Korea. Because those five textbooks were published by the Korean Ministry of Education, their methodology was relatively uncomplicated; there is no similar, central corpus of texts for the teaching of writing in the United States, and so we needed to construct a corpus. We needed to ascertain the most frequently used textbooks in writing classes in the United States.

We collected two hundred syllabi. We created a stratified sample of American colleges and universities by referencing the website US News and World Report for the rankings of National Universities, National Liberal Arts Colleges, Regional Universities, and Regional Liberal Arts Colleges. While we initially planned to gather syllabi from the top two hundred institutions (fifty universities in each category), not all institutions make their syllabi available publicly. For this reason, we decided to start at the top of the list and work our way down until we had collected fifty syllabi in each category, totaling two hundred.

We searched for these fifty syllabi using Google via a Python script, written by one of the authors of this essay, with the phrase “site:xyz.edu college writing
syllabus filetype:pdf.” The first part of this search, “site:xyz.edu”, was used to ensure that only pages hosted under the “xyz.edu” web address would be shown. The second component, “college writing syllabus”, was used to limit the results to sites related to a college writing syllabus, a course that would likely assign a handbook or a style guide as a required textbook. The last part, “filetype:pdf”, was to make sure that only pdf files would be shown, excluding web pages for college writing classes, or websites of instructors that teach college writing, among others. The script was designed to pull from a text file with a list of websites, fill the website into the search phrase, and submit the completed search phrase to Google. We manually searched the syllabi to ensure that they were not for creative writing classes; however, the syllabi included courses that veered from those strictly labeled “first-year college writing class” for a number of reasons beyond mere differences in name, especially because many universities have variations to their requirements for college writing. Nonetheless, we ensured that the teaching of writing was one of the main purposes of the courses from which we gathered the syllabi.

Once we collected two hundred syllabi, we checked each manually to determine which style guide or handbook was assigned as required or recommended. The frequency list of handbooks and/or style guides used by the institutions accessed in this essay is provided in Appendix A. After this, we looked at the most frequently used handbooks.

4. Results, Part I: Lack of consensus in the grammar handbooks on indefinite pronouns

For the two hundred syllabi we collected, we found 96 discrete textbooks required; however, the top 13 choices were used in 99 of the course syllabi. For this analysis, we did not look at all 96 handbooks, instead focusing on these top 13. In each instance, we looked at the most recent edition available to us, assuming material in various editions would be consistent and the changes would be to make examples more current, better appealing to the handbook’s contemporary students.

This task was more difficult than it initially appeared to be—not only
because students purchase numerous editions when a handbook is listed on a syllabus. Quickly we found that various handbook editions change content, text and page numbers, even within the same numbered edition. For example, Joseph Williams’ Style has several publishers, co-authors, and subtitles. Similarly, Diana Hacker’s works have “contributing authors” and “specialists” who are not noted in the main bibliographic information. For some handbooks, pagination differs because one text might be with exercises or readings, another without. An even more confusing element of these handbooks is the fact that sometimes authors repeat text from one handbook to the next, yet with slight variations (e.g., Diane Hacker’s Pocket Manual of Style and Rules for Writers). Overall, we found that handbooks, textbooks, and authors disagreed with each other, and sometimes with themselves (across different editions), offering a range from authoritative dictates that indefinite pronoun constructions must agree with a singular verb to rules of some flexibility.

First of all, some popular textbooks offer no directives. A handbook tied for fourth in the stratified sample of syllabi (Graff and Birkenstein’s 2010 They Say, I Say) contains almost no materials relevant to the handbook tradition as we discuss it here. Neither the APA Manual nor the MLA Handbook offer instruction on these pronouns. Some of the most popular manuals and textbooks say nothing about the proper use of indefinite pronouns.

Other handbooks present a “hard and fast rule” for indefinite pronouns, with black and white standards of correctness. For example, Hacker and Sommers’s Sixth Edition of the Pocket Manual of Style (the most frequently used handbook in our stratified sample of handbooks) makes the following blanket statement:

Even though some of the following indefinite pronouns may seem to have plural meanings, treat them as singular in formal English: anybody, anyone, anything, each, either, everybody, everyone, everything, neither, nobody, no one, nothing, somebody, something (2012: 31).

The seventh edition of Hacker and Sommers’s Pocket Style Manual repeats the rule for “neither” and “either,” though it has less to say and allows for some

---

4 Certainly, the deaths of these authors explain some of the reasons for these bibliographic irregularities.
indefinite pronouns (all, any, none, some) to be singular or plural (2016a: 158). Repeating the rule without the exceptions, Strunk and White’s *Elements of Style* states “Use a singular verb form after each, either, everyone, everybody, neither, nobody, someone” (2007: 18). The rule is hard and fast.

While some handbooks offer this rule, others accept that oral speech allows for variations in what the handbook authors still describe as a hard and fast rule for written language. For these handbook authors, written communication remains a realm of rules for correctness. For example, the instructor’s manual for the thirteenth edition of the *Little, Brown Handbook* informs:

> Recent changes in the language have given rise to some confusion over pronoun-antecedent agreement with indefinite pronouns. Moreover, in English, subjects that are singular in form, like audience or the faculty, may be plural in meaning. In speech, even educated speakers sometimes treat indefinite pronouns like anybody and everybody as plural: ‘Everybody ought to pay attention to their own business’ (Fowler and Aaron 2016: 263).

Similarly, both the seventh edition of Hacker’s *A Writer’s Reference* and the eighth edition of Hacker and Sommers’s *Rules for Writers* differentiate oral speech from written standards of discourse “Indefinite pronouns are pronouns that do not refer to specific persons or things… Many of these words appear to have plural meanings, and they are often treated as plural in casual speech. In formal written English, however, they are nearly always treated as singular” (2011: 179, 2016b: 206–207). Interestingly, this passage is rephrased in Hacker and Sommers’s 2016 *Pocket Manual of Style* (see above)—though the 2016 edition does not qualify for spoken English.

While “either” and “neither” are not included in the exceptions, several handbooks recognize the diversity of written language in dealing with indefinite pronouns, noting that some indefinite pronouns can take either singular or plural verbs. Two such handbooks are Andrea Lunsford’s *St. Martin’s Handbook* (used in five of the syllabi and the ninth most frequently used handbook) and her *Everyday Writer* (2016). While in both Lunsford states that “most” indefinite pronouns take the singular verb, including “neither” and “either” among those that take the singular (2008: 617; 2016: 377), she also observes that “[a]ll, any,
enough more, most, none, and some can be singular or plural, depending on the noun they refer to” (2008: 617, 1995: 123, 2016: 378). Hacker’s 2007 A Writer’s Reference, also allows for some indefinite pronouns (“all,” “any,” “none,” “some”) to take the plural (1995: 124, 2007: 168). However, regarding the indefinite pronouns that are the focus of this paper, “either” and “neither,” Hacker states that they should be treated as singular (1995: 123, 2007: 167).

Various editions of handbook number three, Joseph Williams’s Style: Lessons in Clarity and Grace, offer several interpretations of how to deal with the indefinite pronouns. The 1990 edition, with two chapters co-authored with Gregory Columb, though not discussing “either” or “neither”, specifically states “As flexible as English is, it does have a problem with indefinite subjects. Unlike writers of French, who have available an impersonal pronoun that does not seem excessively formal, English has no convenient indefinite pronoun” (1990: 29). In his revision to Joseph Williams Lessons in Clarity and Grace (2017), Joseph Bizup (like Lunsford and Hacker) notes the exception of irregular pronouns “none” and “any,” which he states that, though “originally singular,” are “today” mostly used as plural (2017: 18). Bizup, however, makes no mention of “either” or “neither.”

In his revision of the same book, Columb notes the problems with these unusual and sometimes conflicting rules (Williams 1990: 21). For example, Columb states that rules should not be considered “draconian” (xiv) and goes into a lengthy discussion of relationships between correctness and power. Attempting to help readers understand what rules to follow, the text goes on to describe several types of rules, with the types overlapping. Joseph Bizup’s 2016 revision of Williams’s Style similarly observes that there are different kinds of rules, labeling them “folklore” and “elegant options.” The folklore options, Bizup writes, are themselves regularly violated by good writers, and “it is not the writers who should change their usage, but grammarians who should change their rules” (Williams 2016: 16). In his description of “elegant options,” the handbook introduces yet another rule, the “real” rule, stating that elegant options “complement Real Rules.”

5 The citation information for the books to which Bizup and Columb contributed does not mention their names, but the forewords and introductions make clear their contributions.
Most readers do not notice when you observe these Real Rules, but do when you violate them (like that). On the other hand, few readers notice when you violate these elegant options, but some do when you observe them, because doing so makes your writing seem just a bit more self-consciously formal. (Williams 2016: 16)

Bizup proceeds to state that a person “can’t predict good grammar or correct usage by logic or general rule. You have to learn the rules one-by-one and accept the fact that many of them are arbitrary and idiosyncratic” (Williams 2016: 21).

In summary, then, while the handbooks prescribe a variety of rules for the use of indefinite pronouns, they also sometimes comment on the problems with these many rules. An analysis of COCA, too, demonstrates writers have a range of choices regarding the rules, and similarly suggests that the rules in the most prescriptive handbooks are problematic.

5. Methods, Part II: An analysis of COCA for indefinite pronoun use

Our work in this study in corpus linguistics participates in what Park and Nam (2017: 428) call “a rigorous methodology used to describe structural, lexical, and variational linguistic phenomena.” Corpus analysis of texts within the COCA has been well-established; for example, in Linguistic Research, an essay by Kim and Moon (2014), use analysis of texts within the COCA to ground their discussion of the SKT construction in English. Like Kim and Moon, we seek to ground our linguistic analysis of the “neither” constructions in a statistically significant sample of English-language usage.

We extracted eight types of indefinite pronoun constructions from COCA.6 Listed below in (3), accompanied with their corresponding examples, are the schematic structures of the eight types of constructions we examined.

---

6 The number of words in each genre is very close (Spoken 95,565,075; Academic 91,066,191; Fiction 90,429,400; News 91,717,452; Magazine 95,558,725) and our interest is the distribution between the singular and plural forms, so there is no reason for additional normalization.
Indefinite pronoun constructions examined

a. [Neither of NounPL + VerbPL] (ex: Neither of the theories are supported.)
b. [Neither of NounPL + VerbSG] (ex: Neither of the theories is supported.)
c. [Either of NounPL + VerbPL] (ex: Either of them are fine.)
d. [Either of NounPL + VerbSG] (ex: Either of them is fine.)
e. [Neither + VerbPL] (ex: Neither are fully accurate.)
f. [Neither + VerbSG] (ex: Neither is fully accurate.)
g. [Either + VerbPL] (ex: Either are acceptable.)
h. [Either + VerbSG] (ex: Either is acceptable.)

Once we extracted these eight types of constructions, we calculated the frequency of each, focusing on the constructions followed by a “to be” verb (the copula), excluding verbs with a neutral number value such as modals (Neither of them will be sufficient.) and non-copula verbs (Neither of them studied rhetoric solidly.) because the past tenses of many non-copula verbs do not differentiate between singular and plural constructions; if we were to include regular verbs, we would only be able to assess the differentiation in the present tense, confusing our results.

For this calculation, logistic regression statistic appeared to be our best choice since the outcome of the dependent variable of our analysis is dichotomous (with all the options being either plural or singular) while there were several factors potentially affecting the choice of the plural verb form. We classified instances in the corpus of any of these four indefinite pronoun constructions as plural or singular—the dependent variable—, thereby yielding a total of eight different types of constructions. We then assessed the linearity of the continuous variable year with Hosmer, Lemeshow, and Sturdivant’s lack of fit test (2013).

That said, we found a simple logistic regression was not sufficient for our analysis because to account for chances of plural or singular forms potentially varying from source to source within a particular genre, a random source effect needed to be included in the models. The reason for the need to incorporate random effects in our models is that simple logistic regression analysis requires that instances of plural or singular, the dependent variable, be independent for all instances under the same conditions of independent variables. Because of the

---

7 The independent variables were 1) Year (1990-2012), 2) indefinite pronoun construction ([Neither of NounPL + Verb], [Either of NounPL + Verb], [Neither + Verb], [Either + Verb]), and genre.
variability of chances of the plural form across different sources, the multiple instances of plural/singular within a given source are not independent. Consider, for example, in search for instances of [Neither of NounPL + Verb] in 2012 for the magazine genre, we found an extreme case where the plural or singular form is always the same within a source. This might be due to the editorial guideline; the periodical might choose an editorial style regarding indefinite pronouns for authors—or a single author might have become accustomed to a particular usage. In this case, the chance that the second instance is plural is entirely dependent on whether the first instance was plural. If the first instance is plural, the second instance will also be plural, whereas if the first instance is singular, the second instance is also singular; therefore, instances of plural/singular are clearly not independent. In less extreme cases, instances from some sources have greater chances of following the plural/singular format than instances from other sources. In these less extreme cases, we observe that the chance of one instance being plural is not independent of whether another instance is plural within the same source. The generalized linear mixed models were fit with SAS (version 9.4) procedure GLIMMIX, with maximum likelihood (rather than quasi-likelihood) fitted via the Laplace method. For this standard model, source to source variability is modeled as normally distributed in the log odds scale. Convergence was confirmed with small gradients at the final maximum likelihood estimates.

A simple logistic regression analysis requires that for instances of plural or singular, the dependent variable be independent for all instances under the same conditions of independent variables. Consider instances of [Neither of NounPL + Verb] in 2012 for the spoken genre. Suppose E = “First instance of [Neither of NounPL Verb] is plural”, and F = “Second instance of [Neither of NounPL Verb] is plural” within a given source. Consider an extreme case where the plural or singular form is always the same within a source; some sources always use plural while other sources always use singular. Under this extreme case, P(second instance is plural | the first instance is plural) = 1 while P(second instance is plural | the first instance is not plural) = 0. Therefore, instances within a source are clearly not independent; within a particular genre for the same indefinite pronoun construction and the same year, the chance of the next instance being plural depends heavily on whether the first instance is plural. In
less extreme cases where some sources have greater chances of being plural than others, \( P(\text{the second instance is plural} \mid \text{the first instance is plural}) > P(\text{the second instance is plural} \mid \text{the first instance is not plural}) \). In this less extreme case as well, we observe that instances within a source are dependent. That is to say, the chance of one instance being plural is not independent of whether another instance is plural within the same source.

Failing to include random effects in the models can result in clearly erroneous inferences particularly in the direction of claiming that effects exist where the evidence in the data do not support that conclusion. If we ignore sources in our analysis, the genres are significantly different with \( p = 0.003 \). We could account for variability from source to source by either 1) including sources within genres as fixed effects or 2) including sources within genres as random effects. The fixed effect independent variables, factors, in our model were year, indefinite pronoun construction, and genre. We are making inferences only about the particular levels chosen for each factor.

If we treat sources as fixed effects, then we are investigating only if these two particular, fixed magazines tend to use the plural form more than these two particular, fixed fictional works. With a source within genre fixed effect, the genres are significantly different, \( p = 0.0003 \). However, our real interest is in whether usage of singular versus plural is different for magazines and works of fiction in general. Using sources as random effects treats these two particular magazines as representative of magazines in general and these two particular works of fiction as representative of fiction in general. Ideally, the two magazines would be chosen randomly from all possible magazine, hence the terminology of random effect, but at least we need to believe that the magazines in the corpus are somehow representative of a larger population of possible magazines. Again, treating all instances as independent, not including source random effects, genres would be significantly different, \( p = 0.0003 \) with a source within genre fixed effect or \( p = 0.003 \) ignoring sources. If we include random effects for sources within each genre, then \( p=0.46 \) for a difference between genres. Consistent with little evidence in this pretend example that magazines in general are different from works of fiction in general, the mixed model p-value which includes random source effects is a much more reasonable assessment of the evidence that magazines in general are different from works of fiction in general based on just these four sources. Failing to use
random effects would lead us to claim a significant genre difference when there is little evidence in these data for that inference.

More detailed information concerning mixed model logistic regression is provided for example by Barr et al. (2013) and Jaeger (2008).

6. Results, Part II: The distribution of “Neither +” / “Either +” constructions in COCA

The raw frequencies of the four indefinite pronoun constructions are illustrated in the four tables (Table 1–Table 4) below.

| Table 1. Total instances of [Neither of Noun PL + Verb] per genre (raw frequency) |
|---|---|---|---|
| Genre | VerbPL | VerbSG | Total |
| Spoken | 128 | 135 | 263 |
| Academic | 51 | 165 | 216 |
| Fiction | 81 | 325 | 406 |
| News | 37 | 131 | 168 |
| Magazine | 46 | 198 | 244 |
| Total | 343 | 954 | 1,297 |

| Table 2. Total instances of [Either of Noun PL + Verb] per genre (raw frequency) |
|---|---|---|---|
| Genre | VerbPL | VerbSG | Total |
| Spoken | 48 | 28 | 76 |
| Academic | 11 | 31 | 42 |
| Fiction | 11 | 58 | 69 |
| News | 4 | 13 | 17 |
| Magazine | 3 | 34 | 37 |
| Total | 77 | 164 | 241 |

| Table 3. Total instances of [Neither + Verb] per genre (raw frequency) |
|---|---|---|---|
| Genre | VerbPL | VerbSG | Total |
| Spoken | 11 | 64 | 75 |
| Academic | 15 | 221 | 236 |
| Fiction | 11 | 96 | 107 |
Table 4. Total instances of [Either + Verb]

<table>
<thead>
<tr>
<th>Genre</th>
<th>(\text{Verb}_{\text{SG}})</th>
<th>(\text{Verb}_{\text{PL}})</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>10</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Academic</td>
<td>11</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>Fiction</td>
<td>4</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>News</td>
<td>5</td>
<td>32</td>
<td>37</td>
</tr>
<tr>
<td>Magazine</td>
<td>2</td>
<td>39</td>
<td>41</td>
</tr>
</tbody>
</table>

The singular verb is most frequently used in all genres, and the genre that uses the plural most is the spoken one \((p < 0.0001\), with an overall odds ratio of 5.52 versus written genres). Even in academic texts, the indefinite pronouns take both singular and plural verbs, though with less frequency than they do in spoken texts.

More precisely, there was significant variation of fractions of plural form across sources (texts) within genres \((p = 0.026\), using a random effects likelihood ratio test based on a mixture of chi-squares). Thus, sources within genres were not independent. However, the variation from source to source was not significantly different across all three indefinite pronoun constructions \((p = 0.712\) or across genres \((p = 0.586\).

Figure 1 shows percentages of plural form for each indefinite pronoun construction within each genre.

---

8 Over all indefinite pronoun constructions, the spoken genre had significantly higher. Individually, for all indefinite pronoun constructions, spoken had greater use of the plural form: \(p < 0.0001\) for both [Either of Noun\(_{\text{SG}}\) + Verb\(_{\text{PL}}\)] and [Neither of Noun\(_{\text{SG}}\) + Verb\(_{\text{PL}}\)], \(p = 0.0027\) for [Neither + Verb\(_{\text{SG}}\)].
The percent axis in Figure 3 is plotted in scale of ln(Odds). In this scale, differences correspond to odds, and parallel lines indicate equal odds ratios. For example, the lines from Academic to Spoken for [Either of NounPL + VerbPL] and [Neither of NounPL + VerbPL] are fairly parallel, indicating similar odds ratios for Spoken versus Academic, 5.14 for [Either of NounPL + VerbPL] and 6.06 for [Neither of NounPL + VerbPL]. The line for [Neither + VerbPL] has a shallower slope indicating a smaller odds ratio for Spoken versus Academic, an odds ratio of 2.50 for [Neither + VerbPL]. Odds ratios compare odds under different conditions. For example, the observed chances or fractions of the plural form for (Academic, Spoken) are respectively fractions of (0.259, 0.643) for [Either of NounPL + VerbPL], (0.197, 0.597) for [Neither of NounPL + VerbPL], and (0.063, 0.145) [Neither + VerbPL]. The corresponding odds are (0.350, 1.801), (0.245, 1.483), and (0.0675, 0.169), which are calculated by the following formula in (4).

(4) Plural odds and odds ratio
The odds ratios for Spoken versus Academic are then $1.801 / 0.350 = 5.14$ for [Either of NounPL + VerbPL], 6.06 for [Neither of NounPL + VerbPL], and 2.50 for [Neither + VerbPL]. For example, with [Either of NounPL + VerbPL], the odds of plural are estimated to be over five times greater for Spoken language than for Academic writing.

The error bars in Figure 3 are standard error bars. 95% confidence intervals would be roughly twice as wide as the standard error bars. Although the estimated odds ratios, slopes, of Spoken versus Academic are larger for [Either of NounPL + VerbPL] or [Neither of NounPL + VerbPL] than for [Neither + VerbPL], given the uncertainty in the estimated odds, they are not significantly different. For example, the ratio of odds ratios for [Either of NounPL + VerbPL] versus [Neither + VerbPL] is $5.14/2.51 = 2.05$ with a 95% confidence interval of (0.62 to 6.81). The possibility that the odds ratios are the same, a ratio of odds ratios of 1.0, is included inside the confidence interval, so we cannot reject the possibility of equal odds ratios. The corresponding p-value is 0.24.

![Figure 2. Source to source distributions of fractions plural in Either of X plus Verb](Fractions plotted in Ln(Odds) scale)
Figure 2 illustrates the overlap from source to source within each genre for “Either of X plus Verb.” Each curve is the estimated normal probability distribution function for Ln (Odds) for the plural form. This figure shows that spoken sources use the plural form more often than do other sources. Academic texts use the plural at the second highest rate, almost overlapping newspapers’ usage. Surprisingly, where the plural is found least frequently is in news and magazine genres, genres that often quote the spoken genre within their texts.

7. Findings from COCA partially disconfirm the handbook tradition

When we conducted our analysis of the use of the indefinite pronoun “neither” with singular and plural verbs in COCA, we found results that contradicted the prescribed information found in the handbooks from first-year composition college classes; however, the verb number with the indefinite pronoun “either” agreed with the handbooks—to a degree. Overall, we call our results a partial disconfirmation of the handbook tradition.

The handbook rules that would lead one to predict that plural is acceptable within oral usage are partially disconfirmed by the result shown in Figure 1. Yes, the spoken word uses the plural most often, a partial confirmation of those handbooks that allow plural construction (e.g. Hacker’s A Writer’s Reference and the eighth edition of Hacker and Sommers’ Rules for Writers). However, academic texts would seem to be the genre least likely to use the plural given that they infrequently quote the spoken word. Instead, novels and newspapers, genres that frequently quote spoken words, are the least frequent users of the plural, with magazines tending to have smaller odds of plural compared to the average over other written genres (p=0.010).

Meanwhile, of course, the COCA results are an almost entire disconfirmation of the prediction we might expect from works like Strunk and White’s Elements of Style, which states “Use a singular verb form after each, either, everyone, everybody, neither, nobody, someone” (2007: 18). The rule is hard and fast, and it has been broken, repeatedly, in actual usage.

Some people might respond to these results by arguing that Figure 1 illustrates the relative frequency of poor usage. Poor usage may be more
common in spoken discourse, after all. However, there were no significant changes in plural/singular usage across years from 1990 to 2012 based on p-values of 0.9563, 0.4273, 0.5628, and 0.3428 for year, (year x construction), (year x genre), and (year x construction x genre), respectively. In other words, the variability appears not to result from language changing or increasing in formality [a change in standards for correctness over the passing of time]. Usage remains stable, more or less, and so the frequencies reflect something more than the likelihood of poor usage.

The handbooks noting the arbitrary and idiosyncratic nature of rules [e.g. the editions of Williams updated by Columb and by Bizup] appear to be confirmed—at first glance. And if students were to follow Bizup’s advice, we would have mountains of flashcards to learn the rules “one-by-one.” However, the variations are not arbitrary and idiosyncratic exceptions to be memorized; they are variations explicable within the rules of Cognitive Linguistics.

8. Discussion: Replacing the rules of handbooks with the insights of cognitive grammar

Cognitive linguistics is a branch of linguistics that examines how language is grounded in social interaction and conceptualization (Langacker 2008: 8–9); people subconsciously use language that matches what they are thinking and want to express to those around them. Rules are never humanly constructed prescriptions. The work of cognitive linguists, then, is to discover the underlying rules of grammar so that linguists can better understand how the human mind works—and how the mind best communicates. In the instance of the indefinite pronoun, Cognitive Grammar studies how people conceptualize and want to express “Either of X plus Verb.”

To understand this conceptualization and expression, linguists turn to metonymy, incorporating synecdoche into metonymy—and using the term metonymy differently than do the fields of rhetoric and literature. Metonymy helps linguists recognize discrepancies between the coded pattern of language and the intended meaning of the writer, discrepancies that are fundamental to everyday language. To illustrate this phenomenon, let us consider the following
Rethinking the teaching of grammar from the perspective of corpus linguistics

sentences:

(5) The cigarette is in my mouth.
(6) James Joyce is hard to read.
(7) Time magazine needs to incorporate more neutral views to attract readers from diverse backgrounds.

All of these examples are metonymic. In the first example (5), what is in my mouth is only the tip of the cigarette, not the whole cigarette. Nevertheless, the writer chooses the cigarette to mean the tip of the cigarette in this grammatical context by utilizing the conceptual metonymy the whole for the part. Here, conceptual metonymy means a general metonymic pattern from which multiple expressions can be created, such as The swan is in the water or My grandma’s got a new set of wheels. The sentence James Joyce is hard to read (6) is slightly different from the cigarette example in that what James Joyce refers to is James Joyce’s work, not a part of James Joyce. This type of referential shift occurs based on the conceptual pattern person for his/her work. In the next example (7), Time magazine refers to the institution, not the magazine itself, motivated by the conceptual metonymy institution for its product. In all of these examples, the metonymic shift of the said words is induced by the grammatical contexts, i.e., the predicates. Due to the use of the predicate is hard to read, for instance, James Joyce undergoes a referential shift from the author to his work. The other two examples can be explained in a similar way.

The role of grammatical contexts is not just limited to the subject-predicate relationship and has bearing on our previous indefinite pronoun analysis. Consider examples (8–10) that contain unitary nouns such as flock and lot:

(8) A flock of geese always fly in a V-formation.
(9) As I write this, a couple of people are strolling through the park near my bench.
(10) These days, a lot of people use their smart phone in lieu of a GPS device.

The noun phrases in bold are similar to the [neither of NounPL] construction in
that all of them exhibit the [Y of X] pattern. In addition, the first nouns of these phrases, schematically notated as X here, are all indefinite, just like the indefinite pronoun *neither*. More importantly, all of these phrases agree with plural verbs: *fly, are, use*. The two noun phrases, a couple of people and a lot of people, even reinforce the plural agreement. The grammatical contexts given in these examples also induce a metonymic shift.

The reason why these noun phrases also show a metonymic pattern may be best accounted for within the Cognitive Grammar framework (Langacker 1987, 1991, 2008). Figure 3 illustrates the conceptual semantic structure of *a flock of geese* (8). To readers who are not familiar with Cognitive Grammar, the figure might look daunting, but what this diagram does is to show the process of combination for *flock* and *of geese*. The portion of this diagram relevant to our discussion is the top two rectangles. The top left box contains a bold circle, which is the representation of *flock*. The inside circle represents a plurality of geese; each goose is indicated by the circle containing the letter *g* and the equal symbol connecting the two circles represents the preposition *of*. In this diagram, the writer’s attention falls on the outer circle—*a flock*—that is indicated by the bold circle. Because the focus is place on *a flock*, as opposed to the plural entity *geese*, the phrase *a flock of geese* may be treated as singular.

However, this is not the only possible construal. The writer has the ability to construe the phrase in a different way by shifting the focus to the inner circle, which is depicted on the top right box. Without doubt, this shift is metonymic because the smaller portion—the inner circle—now represents the whole without losing its intended meaning.
What we are demonstrating here is that the [neither of Noun PL] construction is another case of metonymic shift. Instead of focusing the writer’s attention on neither, she can shift the focus to X<sub>PL</sub>, yielding the plural interpretation of this construction. Earlier, we briefly mentioned that expressions like a lot of people require the plural interpretation. This is due to the loss of the original meaning of lot. Historically, lot meant “an allowance of corn paid as part of a fee to a thresher” equaling either 1/30 or 1/32 of a pound in European countries (OED 2017). This original meaning, however, is now obsolete, and the loss forces the observed metonymic shift; placing the writer’s attention to lot is impossible due to its lack of concrete meaning.

The indefinite pronoun neither seems to undergo a similar historical change. The original meaning of neither was “not of two” (OED 2017). But through contraction accompanied by semantic change, it became an indefinite pronoun. Because the original meaning has been obliterated, its meaning becomes abstract. This newly acquired meaning of neither makes the metonymic shift available to the writer. Different from the case of lot, neither does not force the metonymic shift, because neither still maintains its meaning in some sense: it retains the concepts of negation and duality. As a result, two construals—singular or plural—are accessible to writers, which was confirmed by our corpus analysis.
What we want to emphasize here is that the plural use of the [neither of NounPl] construction is neither ungrammatical nor less-desirable. The variation is just a result of natural language change facilitated by the cognitive ability to process diverse construals, in this case via metonymic shift. When a person states “A flock of geese always fly in a V-formation” as quoted above (8), the person is noting that individual geese adhere to this structure. Changing the sentence to “A flock of geese always flies in a V-formation” shifts the emphasis from the individual geese to the group, an emphasis that seems appropriate in appreciating how geese work together to create the formation.

While many of these metonymic shifts occur on a subconscious level, learning about them can help writers make conscious decisions that will better convey the writers’ meanings. Just as people unconsciously accept how a computer works, but conscious knowledge can assist a person in becoming more effective, students’ conscious knowledge of how language works can make writing more effective. In his uncommon grammar handbook, Frank Cioffi attempts to do just that. Noting, as we do, the frequent publication of the indefinite pronoun “noun” taking the “grammatically incorrect” plural verb in many published texts, Cioffi cites a sentence from the New Yorker describing the photographs in an exhibit: “None are boring” (2015: 25). Though he states he prefers the singular verb, his reasoning is not for matters of correctness:

Looking at the sentence carefully might get you thinking a bit more about the grammatical principle, about what’s at stake here. What does that none signify? If it’s plural, as the are verb implies, what does that say—about the photographic exhibit—that differs from a sentence in which none is seen as singular. A plural verb invokes the whole group, the entirety of photos in the exhibition. For me, the entirety that the plural summons seems to work against the claim that the photos making it up are not boring. The profusion puts me off. Thus I prefer the singular, “none is,” because using a singular verb emphasizes how not a single, solitary one of these photographs is boring.” (Cioffi 2015: 26)

Cioffi’s explanation of why he prefers the singular verb is similar to why we prefer the plural in our geese example. Cioffi stresses that the indefinite pronoun
works to observe the “entirety” of the exhibit that is not boring; we stress that our example works to observe the individuality of the geese in the flock knowing how to make the V-formation. Cioffi, too, is similar to us in his reasoning for citing this indefinite pronoun: the example to make his point that writers “need to go beyond memorizing ‘rules’: [writers] need to internalize how sentences work and deal with each one individually” (2015: 28). So, too, do we.

Such processes are widely observed throughout human language and are by no means limited to this particular construction. Grammar textbooks have not recognized these cognitive and rhetorical complexities. Instead, they hold on to the notion that verbs agree with subjects, and subjects cannot be objects of prepositions, and so writers must match the indefinite pronoun to a singular verb. However, these notions are fictions that do not map onto the way our minds process language. These “rules” undermine the complex cognitive processes going on behind the scenes when good writers write. Without fully examining how our mind works, the textbook tradition is not just rigidly inflexible; it is teaching students to write in a way that decreases their communicative effectiveness. Our “incorrect” usages often reflect how our thinking works.

We agree with Columb’s and Bizup’s revisions of Williams—and we disagree. We agree that there are different types of rules, and we agree that we need to know more about the rules we unconsciously know and about how these rules are shaped by human interaction and conceptualization. These are rules that linguists study, and these rules describe how language works. However, we do not agree that people need to “learn the rules one-by-one” (Williams 2016: 16), as Bizup’s revision of Williams argues, because these rules disagree, not only with what COCA illustrates is used in various publications, but with each other. A deeper understanding of how the mind makes meaning in language, through cognitive grammar, helps the writer make the best choice.

9. Conclusion

As a linguist and two rhetoricians, we are concerned with the elegance Williams’ texts encourage. We also want rhetorical effectiveness. Memorizing
rules one-by-one is not the means to rhetorical effectiveness; understanding how language works is. This examination of handbooks has been dizzying, with rules contradicting each other and evolving. Attempting to learn prescriptive, arbitrary, and idiosyncratic rules is equally dizzying—and limiting. When we continue to teach prescriptive “rules” that people need to learn “one-by-one,” whether folklore, elegant, real or otherwise, we are replacing thinking and understanding with memorization. For as long as teachers work with long mercurial lists for students to memorize, and as long as teachers are unable to teach grammar based on a rich knowledge of genuine linguistic data and theory, they will not succeed in teaching students to become better writers. The fiction of correctness should not be the goal of grammar instruction in the classroom. Instead, following Fraser and Hodson (1978: 53), we believe that “The overriding aim of any good grammar program should be … to enhance the students’ control of language. The essence of style is choice, and choice entails a working knowledge of the available alternatives”. When we dictate rules of correctness (of whatever sort), we not only prescribe truths not found in common usage, we prescribe a standard counterintuitive to our best understandings of language and the mind.

Instead, a reach across the divide of linguistics and writing studies—as Kolln and Hancock suggest—is what we need; we need more of an understanding why a plural verb following an indefinite pronoun might be more appropriate in one rhetorical situation—and why it would not in another. We need to understand how human conceptions of entities like a flock of geese work metonymically. We do not need to memorize when a flock of geese is plural and when it is singular.

This article argues that teaching writing without a knowledge of how language works, without understandings of why language appears arbitrary and idiosyncratic, will fail -- because language is neither arbitrary nor idiosyncratic. There are reasons people will better understand or be persuaded by “neither of the actors is” or by “neither of the actors are,” and our writing classes need to teach students those reasons. Instead, as we hope we have demonstrated, the teaching of rules “one-by-one” creates a dependence on faulty handbooks and leads to teachers imparting partial and flawed knowledge to students, perhaps limiting their communicative resources. Writing teachers will only be able to guide students through this kind of critical pedagogy if they, themselves,
Rethinking the teaching of grammar from the perspective of corpus linguistics

understand the limitations of their tools. A critical education in linguistics is essential for tomorrow’s writing teachers to enact this work and to help their students become rich and powerful communicators.

At the undergraduate level and graduate level, training to teach writing needs to be about understanding the writing process, including how language works. The subject matter knowledge that teachers need is about the functioning of the mind as it processes language. We recommend that every teacher education program expand its offerings to include more than sociolinguistic approaches to language (which assume that language use is learned and marked by race, class, gender), more than literary critical approaches to reading, and more than process-based and rhetorical approaches to writing.

Effective writing depends on a grasp of human cognition that the study of the science of linguistics opens to writing teachers and so to their students.

References


Rethinking the teaching of grammar from the perspective of corpus linguistics


Appendix A:
Handbooks from a Stratified Sample of Colleges and Universities

<table>
<thead>
<tr>
<th>Style Guide</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Pocket Style Manual</td>
<td>22</td>
</tr>
<tr>
<td>The Elements of Style</td>
<td>11</td>
</tr>
<tr>
<td>Style: Lessons in Clarity and Grace</td>
<td>9</td>
</tr>
<tr>
<td>The Everyday Writer</td>
<td>9</td>
</tr>
<tr>
<td>A Writer’s Reference</td>
<td>8</td>
</tr>
<tr>
<td>They Say I Say: The Moves That Matter in Academic Writing</td>
<td>8</td>
</tr>
<tr>
<td>The Associated Press Style Book</td>
<td>6</td>
</tr>
<tr>
<td>The Little Brown Compact Handbook</td>
<td>5</td>
</tr>
<tr>
<td>The St Martin’s Handbook</td>
<td>5</td>
</tr>
<tr>
<td>APA Manual</td>
<td>4</td>
</tr>
<tr>
<td>MLA Handbook</td>
<td>4</td>
</tr>
<tr>
<td>Rhetorical Grammar: Grammatical Choices, Rhetorical Effects</td>
<td>4</td>
</tr>
<tr>
<td>Rules for Writers</td>
<td>4</td>
</tr>
</tbody>
</table>

Chongwon Park
Professor of Linguistics
Department of English, Linguistics, and Writing Studies
University of Minnesota Duluth
1201 Ordean Court, Duluth, MN 55812
E-mail: cpark2@d.umn.edu

Elizabethada Wright
Professor of Rhetoric, Writing Program Administrator
Department of English, Linguistics, and Writing Studies
University of Minnesota Duluth
1201 Ordean Court, Duluth, MN 55812
E-mail: eawright@d.umn.edu

David Beard
Associate Professor of Rhetoric
Department of English, Linguistics, and Writing Studies
University of Minnesota Duluth
1201 Ordean Court, Duluth, MN 55812
E-mail: dbeard@d.umn.edu
Ron Regal
Professor Emeritus of Mathematics and Statistics
Department of Mathematics and Statistics
University of Minnesota Duluth
1117 University Drive, Duluth, MN 55812
E-mail: rregal@d.umn.edu

Received: 2018. 10. 12.
Revised: 2019. 02. 20.
Accepted: 2019. 03. 04.