Using the ACS Journals Search To Validate Assumptions about Writing in Chemistry and Improve Chemistry Writing Instruction

Marin S. Robinson
Department of Chemistry and Biochemistry, Northern Arizona University, Flagstaff, AZ 86011-5698; marin.robinson@nau.edu

Fredricka L. Stoller and James K. Jones
Department of English, Northern Arizona University, Flagstaff, AZ 86011-5698

Chemistry educators face the ongoing challenge of helping students (native and non-native speakers alike) develop and improve their disciplinary writing skills. To this end, increasing numbers of institutions are integrating substantial writing components (1–7) or writing-intensive research-based experiences (8–12) into chemistry curricula. Faculty who teach these courses often rely on authoritative sources, such as The ACS Style Guide (13), to guide their instruction. However, many also rely on assumptions about chemistry writing, assumptions often formed as they learned to write in (and for) the discipline. How well-founded are these assumptions? Do they accurately represent current chemistry writing practices? Or are they outdated or biased toward a particular subdiscipline? As part of a larger chemistry-writing project (11, 14–16), we asked ourselves these questions and set out to validate our own assumptions about how chemists write.

We began by consulting experts in the field of corpus linguistics, a discipline that specializes in investigating language empirically through computer-based analyses of large collections of texts, known as corpora (17–19). With their guidance, we constructed a corpus of chemistry journal articles (referred to as the chemistry corpus), comprising 200 full-length refereed journal articles and 240 sections of refereed journal articles (i.e., 60 abstracts, introductions, methods, and results and discussions) from the Journal of the American Chemical Society, The Journal of Organic Chemistry, The Journal of Physical Chemistry A and B, Analytical Chemistry, and Biochemistry. The corpus was tagged and searched for common writing practices and linguistic patterns in chemistry writing.

Although it was initially useful, we soon realized that the chemistry corpus was difficult to search, hard to update, and too small (only 200 articles). These limitations led us to a second ready-made corpus, accessible through the ACS Journals Search at the ACS Web site (20). The ACS Journals Search uses three databases: the 1879–1995 Legacy Archives (comprising 23 journals and 464,233 articles), the 1996-to-Current Issue (to date comprising 33 journals and 263,000+ articles), and the ASAP-Articles (currently comprising 2,342+ articles). These databases (referred to in this article as the ACS corpus) are continuously updated and accessible to faculty and students. They can be searched collectively or individually for science content, as is well known (21), but also for words, phrases, grammatical constructions, and other common writing practices in chemistry.

In this article, we share results from both corpora, emphasizing results from the ACS corpus. Findings are presented and activities suggested that can raise students’ awareness of common writing practices in chemistry and promote independent use of the ACS Journals Search to facilitate students’ writing development. Sample exercises are included in the online supplement.

Corpus Analyses Findings

Nominalizations

Chemists have a reputation for using nominalizations in their writing; the chemistry corpus was used to examine this assumption. Recall that nominalizations are nouns formed from other parts of speech; the noun is formed by adding endings such as -tion, -sion, -ment, -ity to verbs and adjectives (e.g., distillation from the verb distill, solubility from the adjective soluble). As expected, nominalizations are abundant (List 1); a likely reason for their popularity is that they often make writing more concise (a hallmark of writing in chemistry).

To encourage students to use nominalizations in their own writing, we introduce them to the examples in List 1 and then ask them to (i) identify common nominalizations in excerpts taken from the primary literature, (ii) rewrite excerpts, substituting nominalizations with other words, to witness first hand the rapid loss of conciseness, and (iii) improve passages (often their own) by replacing wordy passages with nominalizations.

Active and Passive Voice

Educators often express conflicting opinions about the use of active and passive voice in writing. Some argue that passive voice is weak and should be avoided; others believe that passive voice is objective and should be used exclusively. Students may be confused by these mixed messages. Which voice is preferred in
chemistry writing? An analysis of the chemistry corpus revealed that passive verbs are used roughly 10–20 times every 500 words, varying with the section of the article (Figure 1). The frequency of passive verbs nearly doubles in the methods section. A closer examination of methods sections indicated that this section is written almost exclusively in passive voice. To emphasize this point, we show students Figure 1 and share with them illustrative excerpts from published methods sections with passive voice (e.g., “A mixture of X and Y was heated” rather than “We heated a mixture of X and Y”).

Figure 1 also indicates that passive voice is used less frequently in the abstract, introduction, and results and discussion sections. By analyzing individual articles, we learned that active voice is used in some instances in these sections (e.g., see the discussion of we below) but that passive voice still predominates. One relevant example is in sentences that refer to others’ works. In passive voice, the science (rather than the scientists) is the subject of the sentence, bringing the science to the forefront. Passive voice is also more concise because the scientists’ names are omitted:

Passive voice: Hexavalent chromium compounds have been shown to be carcinogenic in vivo and mutagenic in vitro. (adapted from ref 22)

Active voice: Bridgewater et al. & Condee et al. & Cunningham et al. & have shown that hexavalent chromium compounds are carcinogenic in vivo and mutagenic in vitro. (adapted from ref 22)

To raise students’ consciousness about voice, we ask them to (i) read and analyze excerpts from various sections of journal articles for appropriate uses of active and passive voice and (ii) convert sentences adapted from the literature that are written in active voice to passive voice and vice versa. Students are encouraged to use passive voice exclusively when writing methods sections and frequently (but not exclusively) elsewhere in their writing.

Personal Pronouns

Most chemists are taught to avoid personal pronouns (e.g., I, we, our) in their writing. The ACS Journals Search was used to determine how closely this practice is followed. Three ACS journals were selected and searched for the use of we over three time periods (Figure 2). Each search returned the number of documents (articles, book reviews, corrections, etc.) that used we at least once. This number was divided by the total number of documents in that time period, estimated by searching for the word the. As shown in Figure 2, the use of we increased in all three journals but particularly in Analytical Chemistry. Today, we appears at least once in more than 85% of the documents published in these journals.

A closer examination of individual articles suggests that we occurs only a few times per article and is used in quite specific contexts. For example, we is commonly used in the introduction, in a sentence that transitions to the work at hand (e.g., In this work, we ...). Additional phrases used to signal this transition are shown in Table 1, along with a list of verbs that commonly follow we (e.g., In this study, we report ...). We is also used in results and discussion sections to highlight decisions, offer interpretations, and summarize accomplishments (e.g., We have adopted a similar approach ...; We reasoned that ...; We have defined a set of compounds that ...; Therefore, we propose that ...; In summary, we have demonstrated ...).

![Figure 1. The number of passive verbs per every 500 words in four sections of the journal article (identified by searching 60 articles, by section, in the chemistry corpus).](image)

![Figure 2. The number of documents using we at least once [relative to the number using the] over three time periods, determined using the ACS Journals Search. (Note: J. Phys. Chem. includes J. Phys. Chem. A and B after 1996.)](image)

| Table 1. Common Phrases Used To Introduce Current Work and Verbs That Follow We in Introductions |
|-----------------------------------------------|-----------------------------------------------|
| Common Transitional Phrases (typically followed by we) | Verbs that Follow We (in order of frequency) |
| in the present study | report | study |
| in the present work | describe | determine |
| in this context | present | assume |
| in this investigation | find | need |
| in this paper | investigate | solve |
| in this study | use | calculate |
| in this work | show | chose |
| herein | focus | propose |
| carry out | employ |

Note: Sixty introductions in the chemistry corpus were searched.

To help students use we appropriately in their writing, we ask them to consider the most compelling reasons for using we in, for example, a results and discussion section and then decide whether we is used appropriately in sample passages. We also ask that they search for uses of we in a journal of their choice, using the ACS Journals Search, to ascertain how and where we is used in that journal.
Words To Avoid

Experienced chemists know almost intuitively what words to avoid in their writing (e.g., researcher, research, truth, fact, dramatically, very, really, and prove). Corpus findings support these intuitions. When analyzing the 60 introductions in the chemistry corpus, we found no occurrences of the word researcher(s). We found a few instances of research, though not as a verb (e.g., to research a problem) as students might use the term. Rather, the term was used in a more generic sense to refer to an area of investigation (e.g., genetics research, pharmaceutical research, macromolecular research). Similar to, in an ACS Journals Search, only 485 documents contained the word truth (compared to 265,783 containing the word the). The word was used in phrases such as “a picture closer to the truth” or “it has been a tightly held truth” rather than as students might use the term (e.g., “We wanted to discover the truth.”).

Words to avoid also include some abbreviations. For example, students often ask whether it is appropriate to introduce abbreviations for room temperature (rt), deionized water (DI), or rotary evaporator (rotovap) in their writing. An ACS Journals Search clearly answers this question; room temperature, deionized water, and rotary evaporator were used without abbreviations greater than 97% of the time in ACS journals published between 1995 and 2006.

We instruct students to search for these and other objectionable words and phrases (e.g., very, good, really, to see if, to find out, to look into, to get rid of) using the ACS Journals Search. In each case, only a few documents are found, sending a clear message to students that these words and phrases should be avoided. At the same time, we give students a list of more appropriate words (e.g., analyze, determine, eliminate, examine, investigate, measure, monitor, reduce). In a follow-up activity, we ask them to replace inappropriate words (in sentences adapted from the literature) with these more appropriate alternatives.

Words To Use

Chemists also know, intuitively, what words to include in their writing. For example, hedging words (e.g., data suggest, findings indicate) are commonly used to temper conclusions, demonstrate restraint, and let the science speak for itself. To raise students’ awareness of hedging words, we introduce Table 2, guide students in identifying hedging words in the literature, ask them to rewrite passages from the literature in which all findings are demonstrated, and let the science speak for itself. To help students add lexical variety to their writing, we compiled a list of words and phrases used to create linkages in chemistry writing and organized them by function (List 2). This list allows students to view the many choices available to them, thereby encouraging variation. The words and phrases can also be searched using the ACS Journals Search to determine whether, indeed, they serve the functions listed or alternative ones. Moreover, students can browse the literature to add new phrases and functions to the list.

We also searched the chemistry corpus for commonly used multiword combinations, or “bundles” (22). Not surprisingly, the most common four-word bundle was “as shown in Figure”. Its popularity was confirmed in an ACS Journals Search; as of this writing, the phrase has been used in over 68,500 documents. In class, we ask students to search both for “as shown in Figure” and the word “figure”. In this way, they learn how common the four-word phrase is and other ways to call out figures in their texts (not to mention how to capitalize and format figures correctly).

Confusing Word Pairs

Even experienced writers often confuse word pairs such as affect and effect, comprise and compose, fewer and less, farther and further, precede and proceed, and principal and principal. The ACS Journals Search can be used in several ways to resolve this confusion. One way is to use the Search to generate a list of words that accompany one word in the confusing pair (e.g.,

List 2. Common Phrases Used To Create Linkages and Their Functions

<table>
<thead>
<tr>
<th>To Show Contrast:</th>
<th>To Provide Additional Information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversely</td>
<td>Additionally</td>
</tr>
<tr>
<td>However</td>
<td>Furthermore</td>
</tr>
<tr>
<td>In contrast</td>
<td>In addition</td>
</tr>
<tr>
<td>Nevertheless</td>
<td>Moreover</td>
</tr>
<tr>
<td>On the other hand</td>
<td>Namely</td>
</tr>
<tr>
<td>Unfortunately</td>
<td>To Give Examples:</td>
</tr>
<tr>
<td>To Add Emphasis or Clarify:</td>
<td>For example</td>
</tr>
<tr>
<td>In particular</td>
<td>For instance</td>
</tr>
<tr>
<td>More specifically</td>
<td>Specifically</td>
</tr>
<tr>
<td>Specifically</td>
<td>To Signal Time:</td>
</tr>
<tr>
<td>To Show Cause and Effect:</td>
<td>Ultimately</td>
</tr>
<tr>
<td>Accordingly</td>
<td>Initially</td>
</tr>
<tr>
<td>As a consequence</td>
<td>To Refer to Something</td>
</tr>
<tr>
<td>As a result</td>
<td>Previously Stated:</td>
</tr>
<tr>
<td>Consequently</td>
<td>As mentioned/described above</td>
</tr>
<tr>
<td>Hence</td>
<td>In the latter case</td>
</tr>
<tr>
<td>Therefore</td>
<td>In this/these/those cases(s)</td>
</tr>
<tr>
<td>Thus</td>
<td>In this context</td>
</tr>
<tr>
<td>To this end</td>
<td>In this respect</td>
</tr>
</tbody>
</table>

Note: Sixty introductions in the chemistry corpus were searched.
principal in List 3). Another way is to compare frequencies of use (Table 3). Knowing, for example, that further occurs in 240,766 documents, but farther in only 3,942, signals that further is more likely the correct choice. Information of this sort captures students’ attention and increases their awareness of commonly confused words. For additional practice, we ask students to select or fill in the correct word in sentences adapted from the literature (not from everyday conversational English).

Two additional confusing word pairs are since—because and while—although. Indeed, these word pairs are confusing even for experienced writers; hence, correct usages will not always be found in the literature. According to The ACS Style Guide (13), since and while should have strong connotations of time (e.g., since the last decade; while the mixture cooled); because, on the other hand, generally suggests a cause–effect relationship and although signals a contrast of some sort (e.g., because the rate increased; although methanol was the preferred solvent). It is instructive to send students to the literature, via the ACS Journals Search, to determine how many authors actually use these terms correctly. Students feel quite accomplished when they discover misuses, which are not uncommon.

**Confusing Plurals**

Tricky plurals abound in chemistry—spectra, appendices, and minima—to name only a few. But perhaps the most confusing plural noun is data, in part, because data can be used correctly as both a singular and plural noun. Its use as a plural noun, however, far exceeds its use as a singular noun. For example, an ACS Journals Search resulted in 58,015 occurrences of data are and 18,054 occurrences of data is (some of which were incorrect). In some instances of data is, the verb accurately agrees with a different singular noun in the sentence (e.g., “A feature of the data is ...”). It is useful to send students to the literature to search for the word data and learn for themselves how it is used (e.g., Table 4). For additional practice, we ask students to select the correct verb in sentences containing data (e.g., the data show or shows) compiled from the literature.

**Directions for Using the ACS Journals Search**

Students can be easily trained to use the ACS Journals Search. From the ACS Publications home page (20), select Advanced Article Search. Go to the full-text search option. Enter the word or phrase to be investigated by the box labeled “anywhere in article.” Use quotation marks around multword phrases (e.g., “data are”) so that the phrase, not the individual words, is searched. The search may be delimited by journal, timeframe (ASAP articles, current issues, legacy issues), or date range. To estimate the current number of documents in the database, search the word “the.” A search returns the number of documents matching the search item and displays citations for those documents. (A site license is required to open and review the documents containing the searched item. For information on current site-license rates and policies, select the Institutional Subscription Information link on the ACS publications home page.)

**Conclusions**

To test our assumptions about how chemists write, two corpora of chemistry journal articles were analyzed for linguistic patterns and words that prove troublesome for novice writers. The findings were incorporated into classroom activities designed to improve students’ chemistry-specific writing and train students to conduct similar analyses on their own using the ACS Journals Search. In this way, we equip students with a
In the Classroom

tool for ongoing writing improvement that uses the literature as the primary source for identifying common chemistry writing practices.

Acknowledgments

The authors gratefully acknowledge the National Science Foundation for support of the Write Like a Chemist project (DUE 0087570 and 0230913). Any opinions, findings, conclusions, and recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Literature Cited


Supporting JCE Online Material

Abstract and keywords

Full text (PDF)

Links to cited URLs and JCE articles

Supplement
Sample exercises involving nominalizations, voice, personal pronouns (we), words to avoid and to use, and confusing word pairs and plurals