The passive voice in scholarly writing:
A diachronic look at science and history

Ping Alvin Leong
Nanyang Technological University

Abstract

While much has been written about the passive voice in scientific writing, similar interest involving humanities writing in general has been more modest. A paucity of diachronic studies also raises the need to understand more about how passive use has changed over time and what such changes imply for the norms in academic writing. This study investigated the use of the passive voice in science and history, with the latter serving as a specific case for the humanities. Eighty articles from the *English Historical Review* (*EHR*) from the late 19th century to the present were analyzed for passive use, and the statistics were compared with those reported in a recent study involving 80 articles from the journal *Science* over the same time span. The results revealed that passive use had fallen in *Science* and *EHR*, and that fewer passives were found in the *EHR* articles. The use of non-finite passives, however, remained stable over time in both *Science* and *EHR*. The semantic functions of passives also differed between *Science* and *EHR*. The study highlights the importance of considering the use of the passive voice in discipline-specific terms. More work involving other disciplines in the humanities and the social sciences is recommended.

Keywords: passive voice, academic writing, science, history, diachronic study

1   Introduction

Rachul (2008), in her master's thesis comparing the features of academic writing in history and biology, made an interesting statement that might well echo how some of us think about the humanities and sciences in general. She wrote:

> While writing in Biology is, as presented by the participants, a record of the work that is accomplished, writing in History is the work. This work involves crafting a story as constructed from the primary sources the historians have collected. (Rachul 2008: 86–87)

Scientific writing, that is to say, records a completed study of the researcher, reporting what was done and found. In history – and arguably the humanities in general – the act of writing is instead regarded as the work *itself*. The focus is on telling the story, on creating an account or an argument from primary and secondary sources.

Such a characterization of academic writing in these two discipline groups naturally suggests marked differences in the way writing is crafted and structured in each group. One would thus reasonably expect these differences to have been investigated to some extent by scholars. The reality, however, is that while there has been extensive work done on scientific writing, research interest in humanities writing has been much more modest. Indeed, the research literature on scientific writing is voluminous, revealing valuable insights into its rhetorical structure (Swales 1981; 1990; Swales & Feak 2004; 2009), thematic structure (Leong 2015; 2016; Leong et al. 2018), and linguistic features (Bazerman 1988; Ding 1998; 2002; Leong 2014; Banks 2017), among others. By contrast, articles investigating humanities writing tend to be more advisory in nature (e.g. Grech 2019). Guidebooks on humanities writing are harder to come by. In fact, Hayot (2014: 7), who published one such book, noted wryly: "Why write a book on scholarly writing for graduate students and faculty in the humanities? Partly because no such book exists".

Admittedly, comparative work of this nature is not easy. Given the wide specializations categorized under each discipline group, differences are bound to be many and varied, blurring any broad conclusions that one may make for a particular aspect of investigation. Even the basic AIMRaD (abstract, introduction, methodology, results and discussion) structure, which appears to apply to scientific writing but not humanities writing, is not robust. Cargill & O'Connor (2009), for instance, have noticed differences in this rhetorical structure not only across various scientific disciplines, but also within the same discipline.

In this article, I propose that we examine a far more fundamental grammatical feature that may shed some light on the general way in which each discipline group approaches its area of study. I have in mind the passive voice. Ding (2002), for instance, argues that the passive voice actually *embodies* the practice of science – it draws attention to the work (rather than the person), and so invites fellow scientists to either replicate or verify it. On the other hand, as we have seen earlier, history writing appears to be the work itself (Rachul 2008), focusing on the account, and thus the writer's perspective

and interpretation. It stands to reason, therefore, that we would expect to see more passives in scientific writing than in history writing.

In order for this grammatical feature to accurately mirror the general approach in each discipline group, we would also expect that this hypothesized difference in passive use to be true over time. This is complicated by the fact that the use of the passive in scientific writing has fluctuated. Recent studies have suggested a shift towards a greater use of the active voice in scientific writing (Leong 2014; 2020; Banks 2017). In particular, Leong's (2020) diachronic study of articles from the journal *Science* showed that the proportion of passive clauses in scientific articles dropped from approximately one third to one fifth of all clauses between 1880 and 2017. He suggests that this shift implies a move towards making scientific writing more accessible to both specialists and non-specialists, given the current interdisciplinary trend in scholarly research. This in turn implies a reconceptualization of scientific writing – that it is not merely a record of past work done (because the passive is still frequently used, especially in the methodology section), but a record that aims to be as accessible as possible.

Is this trend in scientific writing mirrored in history writing as well? There is suggestive indication that written English, in general, has seen a fall in passive use (Mair & Leech 2006; Hou & Smith 2018), ranging between 12% and 20%; in academic prose in general, Biber et al. (1999: 476) report that about 25% of all finite verbs are in the passive voice. Whether this decline is also true of history writing requires investigation. This study addressed this research gap by analyzing the use of the passive voice in history research articles and comparing the trend with the statistics for scientific writing as reported in Leong (2020).

This paper is organized as follows. § 2 presents a review of the related literature on studies done on scholarly writing, with a particular focus on comparative and diachronic studies. The corpus for this study, including the selection of history research articles, and the analytical procedures are detailed in § 3. This is followed by the findings in § 4, which discusses the general trend concerning the use of passives, as well as the semantics of the verbs typically chosen as passives. The final section summarizes the main conclusions and implications of the study, and recommends areas for further work.

2   Review of related literature

Much has been written about the development and linguistic features of scientific writing (Bazerman 1988; Gross 1996; Montgomery 1996). We may attribute the style of scientific discourse to the recommendations of Francis Bacon (1561–1626). Scientific writing started off as being plain and direct; writers were encouraged to avoid ornate language so as to make science "accountable to its readers" (Montgomery 1996: 93). In the 19th century, this direct style of reporting became more abstract. Science became more impersonal and thing-centered, what Gross et al. (2002: 231) refer to as "an objective enterprise". The passive voice became a hallmark of scientific writing. Diachronic studies involving articles from the *Philosophical Transactions of the Royal Society of London*, the oldest journal devoted to science, reported a marked increase in the use of the passive voice from the 17th century to the 20th century. Various measures were used to depict the increase; these included instances per thousand words (from 22.3 to 36.6; Atkinson 1996), and the percentage of finite verbs used as passives (from 21.2% to 32.8%; Banks 2008).

Beyond the 20th century, available studies suggest that scientific writing has now become less reliant on the passive voice. The work of Seoane & Loureiro-Porto (2005) found a fall in passive use in articles in the sciences and medicine, from 64.2–66.4% in the early 20th century to 46.4–58.7% by the end of the century, reflecting the growing need for scientific discourse to be more direct and accessible to readers (Seoane 2006; 2013).

The diachronic evidence provided by Wanner (2009), as part of a larger work, appears to counter this. She examined texts from seven corpora of scientific writing from 1650 to 1990, but restricted the investigation to only eight reporting verbs (e.g. *argue*, *demonstrate*, *indicate*). Her results showed that the passive versions of these reporting verbs increased and peaked at 78 tokens per 2,000 words in the 1900–1949 period, but this was followed by a marked decline to 45 tokens per 2,000 words in the 1950–1990 period. However, as Wanner's (2009) work focused on only selected verbs, these results offer but a restricted view of scientific writing. By her own admission, "[i]t remains to be seen if the lower numbers of passives in reporting events are really a trend that specifically affects the passive" (Wanner 2009: 189).

As it turns out, the views of the scientific community at large appear to have turned against the overuse of the passive. In the 20th century, early published accounts on the need for a more direct style of writing took the

form of personal viewpoints (Bridgman 1955; Ormes 1957; Robinson 1957). There were, as might be expected, opposing opinions. For instance, in a letter to the journal *Nature*, Leather (1996: 467) argued that the use of the passive should be seen as a virtue, and that "[t]he active voice encourages carelessness, partisanship and, as used by many of its adherents, does no favours to the English language or science". This led to pushback from a number of scientists (Baskin 1996; Goodman 1996; Jolly 1996). The reply that perhaps best captured the discrepancy in Leather's own writing is the following from Perlman:

> I consider Leather's letter an outrageous display of scientific hypocrisy. He makes dogmatic pronouncements on a subject he knows nothing about […] Leather should practise what he preaches: of the 18 transitive sentences in his letter, only four were in the passive. (Perlman 1996: 108)

More recent studies involving texts published in the 21st century suggest a further drop in passive use in scientific articles. In his analysis of scientific articles from 1985 to 2015, Banks (2017) found that the use of the passive generally decreased during this period, with the fall being more marked in the physical sciences than the biological sciences. This is echoed in Leong's (2020) larger study involving scientific articles from the journal *Science*. The proportion of passive clauses to all clauses fell from a high of 35.82% in 1980 to 23.21% in 2017.

How this trend in scientific writing compares to that in humanities writing, unfortunately, is uncertain. To the best of my knowledge, there has not been a similar diachronic study attempted for humanities writing with a distinct focus on the grammatical voice. While non-scientific writing has been analyzed for grammatical features, such studies have tended not to include the passive voice. Hyland (2002), for instance, investigated the use of personal pronouns in the sciences, social sciences, and humanities, and MacDonald (2010) focused on knowledge creation in three subfields of history, but neither discussed passive use to any extent. In their extensive study involving academic texts (among others) from 1700 to 2005, Biber & Gray (2016) acknowledge that the use of the passive has decreased over time in scientific writing to about 25% of all finite verbs. However, their diachronic comparison of scientific writing and humanities writing (represented by history writing) laid emphasis on other grammatical features. In particular, Biber & Gray (2016) contrasted the relative stability of history

writing in its dense use of nouns and heavy reliance on relative clauses with the move in scientific articles towards a more compressed style of writing.

Where comparative studies specifically involving the passive voice are concerned, the works of Iddings (2007) and Rachul (2008) are noteworthy. In the textual-analysis part of his study, Iddings (2007) selected six texts, but only two of them – one each from literature and biochemistry – were analyzed in-depth for their grammatical features. His analysis revealed that the biochemistry paper used about four times as many passives as the literature paper – the passive was found in about 46% of all clauses in the former but just 10% in the latter. A similar finding was reported by Rachul (2008), who compared the use of the passive voice in four biology articles and two history articles. Using the T-unit, a single main clause together with any subordinate clauses attached to it (Hunt 1977), she found that biology papers used two to three times more passives than history papers. The passives per T-unit for the biology papers ranged from 0.317 to 0.681, as compared to the 0.115–0.185 range for the history papers. She also found that the biology papers used more verbs describing the methods and tools in the research process (i.e. epistemic verbs) as compared to the history papers, which relied more on verbs focusing on the topic (i.e. phenomenal verbs).

These studies, then, suggest that scientific writing contains more passives than humanities writing does. However, while the efforts of Iddings (2007) and Rachul (2008) are insightful, their results are compromised by at least two issues. First, the sample sizes in both studies are very small. Iddings (2007) analyzed only two articles in-depth, and Rachul (2008), just six. Any generalizations arising from such a small sample size are hardly likely to be representative. Second, the works of Iddings (2007) and Rachul (2008) are *synchronic*, not diachronic. It is therefore unclear whether humanities writing has always been less reliant on the passive voice as compared to scientific writing, or whether this is a relatively recent development.

Addressing this research gap, though, is not a straightforward task, chiefly because of the range of disciplines grouped under the broad label "humanities". Accessing a ready corpus of articles through the ages is another challenge, since such articles need to be carefully selected for a fair comparison to be made. Details about the selection of the articles for comparison and the methodological procedures are outlined in the next section, to which we now turn.

3 Methodology

3.1 Corpus

In this study, I used the statistics reported by Leong (2020) for scientific writing, as his study is the most recent record of this trend. In his work, Leong sourced 80 research articles from the journal *Science*, a highly prestigious journal launched in 1880 (Scimago 2020). *Science* was selected, as its long history allowed for the trend in scientific writing over the years to be readily noted. It is also multidisciplinary and publishes articles from any scientific discipline and topic. The articles were taken from four time periods – 1880, 1930, 1980, and 2017. The first 20 articles each for 1880, 1930, and 1980, and the final 20 articles for 2017 were selected for investigation. That is to say, at the time of analysis, Leong's corpus included articles from the very first issue of *Science* to the final 20 articles published in 2017, spanning more than 130 years.

Unlike the sciences, unfortunately, there is no single multidisciplinary journal for the humanities. Attempting to locate scholarly articles from different humanities disciplines dating back to the 19th century would also be cumbersome and unproductive. To ensure that the comparison with the scientific articles in Leong's study was as fair as possible, I therefore selected a journal that was similar to *Science* in terms of its heritage and continuity. This journal was *English Historical Review* (hereafter *EHR*). Like *Science*, *EHR* also has a long history. Established in 1886, it is, in fact, the oldest English-language journal in the discipline of history. It is also a respected journal, ranked 142nd out of 1,387 journals in the "history" category in *Scimago's Journal and Country Rank* (2020). Similar to Leong's study, 80 history research articles in *EHR* were selected for analysis – the first 20 articles each for 1886, 1930, and 1980, and the final 20 for 2018.

3.2 Analytical procedure

Following the methodology in Leong (2020), I first divided up the history articles into clauses, both main and subordinate, the latter of which included embedded clauses (such as noun clauses and relative clauses). Each clause was then manually analyzed for the passive voice. The passive is a well-known grammatical feature in English; descriptions about its structure can be found in any authoritative grammar book (e.g. Greenbaum et al. 1985;

Biber et al. 1999; Carter & McCarthy 2006) and will not be further elaborated here. In this study, all forms of the passive were considered, both finite and non-finite. These comprised not merely the typical forms containing the *be* auxiliary verb, but also bare passives, which are a reduced form of the passive and so do not have any auxiliary verb (Puckica 2009: 215). Including non-finite passives, particularly bare passives, is crucial for a more complete account of passive use in scholarly writing, as past studies have tended to omit non-finite passives from consideration (e.g. Banks 2008; 2017).

Examples (1–3) below – involving different types of clauses and passive forms – illustrate the analysis undertaken in this study. The number assigned to each clause is represented in square brackets, and the year and reference number of the article are indicated at the end of each example. All passives are in bold typeface, and recovered words are enclosed in angled brackets.

(1)   [82]   […] so it **was determined** by Theodore and his councillors
      [83]   to change their abode to the hill-set village of Cervione,
      [84]   where the coronation could take place,
      [85]   and whence operations **could be** more advantageouly [sic] **commenced** against the Genoese. [two finite passives; 1886, 11]

(2)   [41]   It was the normal administrative practice for drafts of legislative enactments
      [42]   **prepared** by ministries
      [43]   **to be scrutinized** in detail by members of the Council. [bare passive followed by non-finite passive; 1980, 7]

(3)   [1377]   The knowledge of history – the careful perusal of the past – allowed for lessons
      [1378]   **to be drawn**
      [1379]   and <**to be**> **disseminated** […] [two non-finite passives; 2018, 20]

3.3   Statistical analysis

The statistical analysis was facilitated by the Real Statistics Resource Pack (Zaiontz 2018), an add-in which extends Microsoft Excel's capability to perform statistical calculations. The one-way analysis of variance (ANOVA) test was used to investigate if the observed differences in the means between the *Science* and *EHR* articles across the time periods were statistically

significant. The Tukey Honest Significant Difference (HSD) post-hoc test was used for significant ANOVA results. The significance level for all statistical tests was $\alpha = 0.05$. In this paper, significant differences are indicated using the asterisk – a single asterisk for a significant result ($p < 0.05$), and double asterisks for a highly significant result ($p < 0.01$).

4   Findings and discussion

4.1   Summary statistics of the *Science* and *EHR* articles

The summary statistics of the *Science* articles, taken from Leong (2020: 473), and the *EHR* articles are presented in Table 1. In general, the *EHR* articles are about three times as long as the *Science* articles, and also contain about three times as many clauses, whether main or subordinate. Of particular interest is the number of non-finite clauses used in each sub-corpus; these clauses account for more than a quarter of all clauses used in *Science* articles (27.19%) and slightly less than a third in *EHR* articles (33.17%). As alluded to in § 3.2, not all studies in the past have paid much attention to non-finite passives in their analyses. The statistics in Table 1 suggest that omitting non-finite clauses (which may carry such passives) effectively removes between a quarter to a third of all available clauses from consideration, thereby compromising the completeness of the analysis. We shall return to the issue of non-finite passives, particularly bare passives, in § 4.3.

4.2   General trend in passive use

The mean percentages of all clauses containing passives are illustrated in Figure 1. The trend lines highlight two important details. The first is that in all years, the *EHR* articles used fewer passives than the *Science* articles. The observed differences between *Science* and *EHR* for all four time periods were statistically significant. In three of the four periods – i.e. 1880/1886, 1980, and 2017/2018 – the differences were in fact highly significant (i.e. $p < 0.01$).

This lends weight to what some scholars have noted regarding the use of the passive in scientific and humanities writing (Iddings 2007; Rachul 2008; Baratta 2009; Hundt et al. 2016). Grech (2019: 97), for instance, notes that "[i]n scientific papers, the passive voice is used as this underscores objectivity, whereas the humanities utilise the active voice, emphasizing subjective and

Table 1. Summary statistics of the *Science* and *EHR* articles

|		*Science*	*EHR*
Words	1880 (*Science*), 1886 (*EHR*)	23,713	163,071
	1930	47,503	143,582
	1980	46,496	171,544
	2017 (*Science*), 2018 (*EHR*)	56,891	208,303
	Total words	174,603	523,429
All clauses	1880 (*Science*), 1886 (*EHR*)	2,331	17,845
	1930	4,338	15,320
	1980	4,336	18,046
	2017 (*Science*), 2018 (*EHR*)	5,009	21,549
	Total clauses	16,014	54,915
Main clauses	1880 (*Science*), 1886 (*EHR*)	994	7,314
	1930	1,938	6,663
	1980	2,124	8,030
	2017 (*Science*), 2018 (*EHR*)	2,342	9,399
	Total main clauses	7,398	24,092
Total subordinate clauses (including non-finite clauses and relative clauses)	1880 (*Science*), 1886 (*EHR*)	1,337	10,531
	1930	2,400	8,657
	1980	2,212	10,016
	2017 (*Science*), 2018 (*EHR*)	2,667	12,150
	Total subordinate clauses	8,616	30,823
Non-finite clauses	1880 (*Science*), 1886 (*EHR*)	636	4,217
	1930	1,119	3,615
	1980	1,114	4,486
	2017 (*Science*), 2018 (*EHR*)	1,486	5,899
	Total non-finite clauses	4,355	18,217
Finite relative clauses	1880 (*Science*), 1886 (*EHR*)	76	437
	1930	95	274
	1980	95	261
	2017 (*Science*), 2018 (*EHR*)	90	279
	Total finite relative clauses	356	1251
Non-finite relative clauses	1880 (*Science*), 1886 (*EHR*)	128	711
	1930	236	601
	1980	361	817
	2017 (*Science*), 2018 (*EHR*)	304	971
	Total non-finite relative clauses	1029	3100
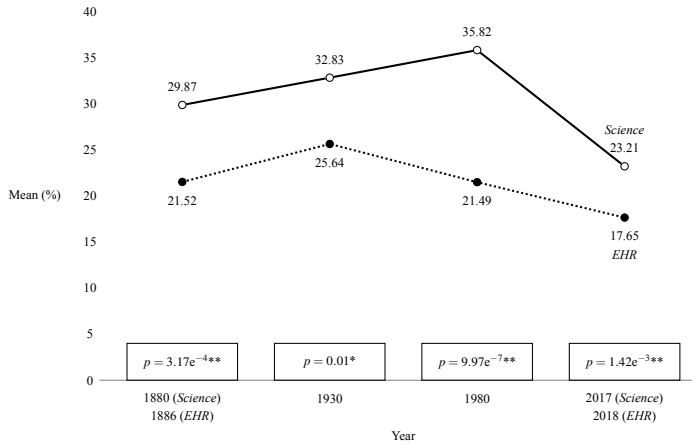
Figure 1. Mean percentages of clauses containing passives in *Science* and *EHR* articles across the four time periods

persuasive rhetoric". The statistics presented here, using history as a case in point, allow us to understand this difference a little better. Through the years, the difference in passive use between *Science* and *EHR* is in the region of 5–15%, with 2017/2018, the final time period, showing the least difference between the two.

The second area of interest is that the shape of the trend lines for both *Science* and *EHR* seem, at first glance, similar. There is a peak in 1930 and 1980 for *EHR* and *Science*, respectively, with a decline in passive use thereafter. Leong (2020) reports that the mean for the 2017 *Science* articles was significantly different from the means for all the other time periods. The decrease in the mean values between 1980 and 2017 was most marked ($p = 1.08e^{-4}$).

The situation in the case of the *EHR* articles, though, is slightly different. Following up from a significant ANOVA result ($p = 7.49e^{-4}$) for the *EHR* means across the four time periods, a Tukey HSD post-hoc test was run. The test revealed that only the 1930 and 1980 means were significantly different ($p = 2.57e^{-4}$). All other differences were not statistically significant. This suggests that the fall in passive use may be leveling out at about 17% for *EHR*. Further, given that the difference in means between *Science* and *EHR* is at its

narrowest in 2017/2018, could a further fall in passive use in scientific writing in subsequent years result in its mean nearing that for the *EHR* articles?

   While this is a tantalizing possibility, which may give credence to the observation of Mair & Leech (2006: 331) that "in the course of the past century written English has moved closer to the norms of spoken usage", the evidence presented here suggests that scientific writing will continue to use more passives than history writing in the foreseeable future. The remarks by Ding (2002) about the role of the passive in scientific writing have relevance here. He argues that the use of the passive voice is socially conditioned by the scientific community and that their thing-centered approach to scientific work is embodied in the passive voice (Ding 2002: 138). Hence, while scientific writing may appear to be relying less on the passive voice now, it is still customary to describe the methodology using the passive. The following example, taken from Leong (2020), is a case in point:

(4)   For this purpose $(129 \times AKR)F_1$ hybrid mice **were** each **inoculated** subcutaneously with IO^6 AKR SL2 cells [a dose 2 logs greater than the 100 percent lethal dose (LD_{100})]. The mice **were** then **separated** into four groups for treatment, which **was initiated** 1 to 2 hours after tumor inoculation. Mice in one test group **were** each **treated** with 19-A10 ascites fluid; mice in a second test group **were** each **injected** with 19-A10 ascites fluid and rabbit serum as a source of exogenous C. Two additional groups in this experiment **were used** as controls; [...] (Leong 2020: 482)

It is entirely possible to re-word (4) in the active voice. But this is rarely done in scientific writing; indeed, in the re-written version (5) below, the repeated use of the first-person pronoun *we* is not just stylistically awkward but draws attention away from the mice.

(5)   For this purpose $(129 \times AKR)F_1$ we inoculated hybrid mice subcutaneously with IO^6 AKR SL2 cells [a dose 2 logs greater than the 100 percent lethal dose (LD_{100})]. We separated the mice into four groups for treatment, which we initiated 1 to 2 hours after tumor inoculation. We treated each mouse in one test group with 19-A10 ascites fluid; we injected mice in a second test group with 19-A10 ascites fluid and rabbit serum as a source of exogenous C. We used two additional groups in this experiment as controls; [...]

By contrast, the passives in the *EHR* articles do not appear to be bound by such constraints. Instead, they are employed in the typical ways described in grammar books, such as when the actor is either non-specific (6) or simply too numerous to list (7).

(6)   [9]   Most of the more important or interesting works of Italian humanist historians of the fifteenth century **must** still **be read** in the original Renaissance editions

        [10]   or **have to be studied** in widely scattered manuscripts. [1980, 11]

(7)   [5]   European and colonial themes **are** usually **pursued** by different scholars with varying purposes, […] [2018, 3]

The use of the passive voice in *EHR*, that is to say, is perhaps closer to its congruent use in ordinary written language. The passive is used to fit the description of the historical figure or phenomenon, and less to promote thing-centeredness so highly valued in science.

4.3   Passives in non-finite clauses and relative clauses, and bare passives

We turn now to the issue of passives in non-finite clauses, as this class of passives is rarely investigated in past studies. The mean proportions of non-finite passives in relation to total clauses are shown in Figure 2.

Across the four time periods, the mean values for each journal were not found to be statistically different from each other. In other words, the use of non-finite passives in the *Science* and *EHR* articles remained stable over time. Hence, while there has been a noticeable decline in the overall use of the passive in both *Science* and *EHR* (see § 4.2), this has not affected the use of non-finite passives in either journal, with such passives ranging between 8–10% and 5–7%, respectively, of all clauses.

Comparing the journals, though, reveals a marked difference in the use of non-finite passives. The trend lines in Figure 2 clearly show that for each time period, more non-finite passives were used in the *Science* articles than in the *EHR* articles. The observed difference in each time period between the *Science* and *EHR* articles was highly significant, as shown by the small *p* values in Figure 2. The result here, then, suggests a consistency in the use of passives between *Science* and *EHR* – apart from one possible exception, scientific articles simply use more passives, whether finite or non-finite.
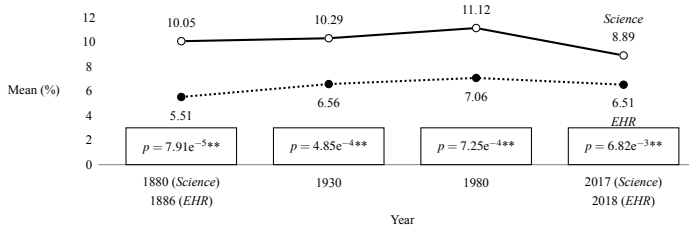
Figure 2. Mean percentages of non-finite passives in *Science* and *EHR* articles across the four time periods

This exception concerns the use of passive finite relative clauses. Biber & Gray (2016: 207) suggest that there has been a drift in the grammatical structure of the NP towards an "economy of expression resulting in structural compression to convey the maximum amount of information in the fewest words possible". If this is true, we would expect to find more instances of non-finite relative clauses than finite ones, the latter of which are "complete clauses, including subjects and full verb phrases marked for tense, aspect and modality" (Gray 2015: 126). Leong (2020) cites an example, reproduced in (8) below, to show the compactness that is achieved through the use of a non-finite relative clause.

(8)   [176]   However, the supernova features <…> are absent in SSS17a, […]
        [177]   <**seen** in GRB980425> (Leong 2020: 479)

As he points out, converting the verb in (8) to the active voice would both lengthen the word count and compromise the elegance of the message flow (see 9). Using a finite passive would not improve the situation either, as it would increase the word count even further (see 10).

(9)   However, the supernova features which the scientists saw in GRB980425 are absent in SSS17a. (Leong 2020: 479)

(10)   However, the supernova features which **were seen** by the scientists in GRB980425 are absent in SSS17a.

As seen in Table 1 earlier, passive non-finite relative clauses were used more often than passive finite clauses in all time periods. In the 2017/2018 time
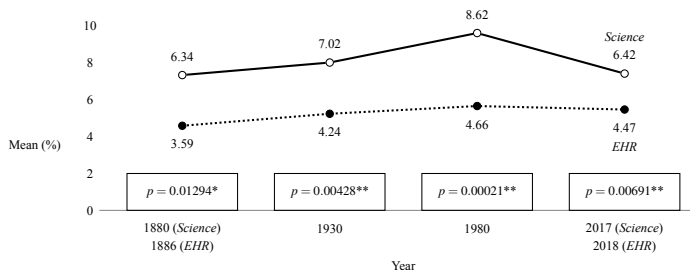
Figure 3. Mean percentages of passive non-finite relative clauses in *Science* and *EHR* articles across the four time periods
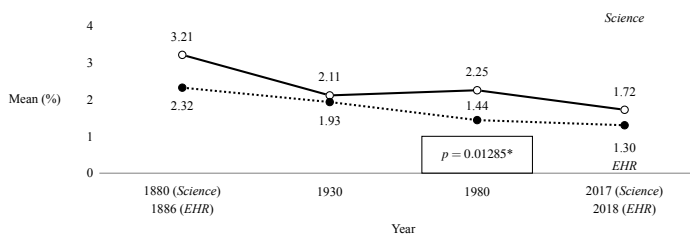

Figure 4. Mean percentages of passive finite relative clauses in *Science* and *EHR* articles across the four time periods

period, for instance, passive non-finite clauses in both journals appeared almost 3.5 times as often as finite relative clauses. This preference for passive non-finite relative clauses was more marked in scientific writing than history writing, as shown in Figure 3.

Where passive finite relative clauses are concerned, though, the results are mixed. As shown in Figure 4, even though the trend line for *Science* is above that for *EHR*, the observed differences in three of the four time periods were found to be statistically insignificant. This suggests a shift away from the use of finite relative clauses in both scientific and history writing, and mirrors what Gray (2015) and Biber & Gray (2016) have observed about the changes in academic prose over time. Specifically, "the most notable change […] occurred in academic research writing, which has undergone a dramatic reduction in the use of relative clauses over the last fifty years" (Biber & Gray 2016: 150).

The greater reliance on non-finite passives in *Science* may be further understood by turning our attention to the use of bare passives. This is because non-finite passives in both journals were largely made up of bare passives. Bare passives formed 80–95% and 81–90% of all non-finite passives in *Science* and *EHR*, respectively; this works out to approximately 8–10% of all clauses in *Science*, and 4–6% of all clauses in *EHR*. In other words, roughly one in ten clauses and one in twenty clauses in *Science* and *EHR*, respectively, contains a bare passive. Bare passives, in fact, were far more extensively used in *Science*; in 27 of the 80 scientific articles, all the non-finite passives were bare passives, as compared to only one such article in *EHR*.

As noted earlier, the extensive use of bare passives in *Science* reflects the compressed nature of scientific writing in general (Biber & Gray 2016). Referring to bare passives as *participial passives*, Wanner (2009: 85) observes that since "compact expressions are highly valued, it does not come as a surprise that the participial relative clause is not a marginal phenomenon in that particular register". Writing guides, such as one from the Royal Literary Fund, also share the same sentiment. In relation to scientific writing, they advocate the following:

> All writing guides – including this one – give similar advice: no unnecessary words, make every word count, keep it concise. Every sentence needs to be toned for high performance: plenty of muscle and no excess fat. This is good advice for writing in general but lean writing is especially important in scientific writing because scientific writing places its emphasis on gathering and reviewing evidence; and on conveying quantitative information. (Royal Literary Fund 2021)

This "lean writing" is facilitated by the bare passive, as it contains only the past participle verb and, as a non-finite verb, typically occurs without the grammatical subject, as seen in (8) earlier.

In the *EHR* articles, bare passives were also used in the same way to achieve economy in the writing and enhance the message flow, as exemplified in (11–12):

(11)   [30]   Of these Wietersheim, **edited** by Dahn,
        [31]   is perhaps the most helpful; […] [1886, 3]

(12)   [2]   It was central to the maintenance of a post-war political
                settlement **defined** by support for full employment, high
                welfare spending, and an active Cold War foreign policy,
                […] [2018, 10]

Where the motivation to use bare passives is concerned, then, there does not appear to be any difference between *Science* and *EHR*; in both, bare passives are used to achieve a level of compactness in the writing. The only obvious difference lies in the extent of its use. As we have seen, bare passives occur more frequently in the *Science* articles – almost one in every ten clauses contains a bare passive.

4.4 Semantics of passive verbs

We turn lastly to the semantics of the passive verbs used in *Science* and *EHR*. In Leong's (2020) study, the passive verbs in the *Science* articles were coded into one of three categories – *productive*, *indicative*, and *derivative*. A productive verb describes any physical action performed by the writer(s) (e.g. *One grub was placed in each of eight sterile tin salve cans*), an indicative verb draws the reader's attention to an external source or segment within the text, including the basis of that information (e.g. *As shown in Table 3*), and a derivative verb describes the development or production of a thing or an idea from something else (e.g. *The AKR SL2 leukemia was derived from a spontaneous thymoma*). Leong noted that the majority of the finite verbs in the *Science* articles were productive in nature, but that the bare passives shifted from productive verbs in 1880 to a mixture of indicative and derivative verbs by 2017.

In the analysis of the *EHR* articles, the *productive* category was amended, since, by and large, historians do not perform research procedures in the way scientists do. Instead, historians recount what others did, and make sense of their actions and outcomes. As a consequence, the *productive* label for the *EHR* articles was amended to refer to any actions performed by historical figures (rather than the researchers) in the text. The other two categories – *indicative* and *derivative* – were left unchanged. Examples of productive, indicative, and derivative passives used in the *EHR* articles are given in (13–15).

(13) [52] He speaks of the year 1361
     [53] when the Chancery rolls **were removed** from the White Tower to the Wakefield Tower. [productive finite passive; 1930, 10]

(14) [627] An early example of a second council created by charter
     [628] **is found** at Colchester. [indicative finite passive; 1930, 17]

(15) [139] Welsh dower, <…>, was also consistent with Welsh legal norms.

[140]   <while **influenced** by English common law> [derivative bare passive; 2018, 1]

The top three finite and bare passive verbs used in *Science* and *EHR* for each time period are listed in Table 2. The labels *productive*, *indicative*, and *derivative* are shortened to *Pro*, *Ind*, and *Der*, respectively.

Two differences between the *Science* and *EHR* articles are immediately obvious. First, as noted earlier, the majority of the finite passives in *Science* were productive verbs. By contrast, in the *EHR* articles, there was a roughly even distribution of productive and indicative finite passives. Second, whereas the *Science* articles witnessed a shift from a greater use of productive bare passives to a mix of indicative/derivative bare passives, there was no change as regards the *EHR* articles – the bare-passive verbs remained largely productive verbs across the four time periods.

Leong (2020) observed that the difference in the use of finite and bare passives in *Science* may reflect a "division of labor":

> This may well signal a 'division of labor', as it were, between finite and bare passives in scientific writing, where the productive function, among others, is served chiefly by finite passives, and the more stative functions by bare passives. (Leong 2020: 480)

In the *EHR* articles, by contrast, it is the finite passives that express the dynamic and stative functions; the bare passives serve primarily the productive function. The use of productive and indicative passives in *EHR* is only to be expected, since the analysis of historical events necessitates the historical actions to be first conveyed (through productive verbs) before a particular aspect is singled out for study (through indicative verbs). Indeed, the use of productive and indicative passives mirrors the general rhetorical move in history writing. This move, of course, is largely facilitated by active-voice verbs, given that passive verbs form only about one fifth of the total verbs in *EHR* (see Figure 1). But evidence of this move involving passive verbs is seen in the corpus as well. Consider (16) below, taken from the opening few paragraphs of an article about the recruitment of princely armies in the late medieval low countries.

(16)   [1]   In 1338–9, Duke Jan III of Brabant (r. 1312–55) raised an army of more than 1,500 knights and squires
        [2]   to help Edward III of England (r. 1327–77) in his struggle
        [3]   to win the French Crown.

Table 2. Top three finite passives and bare passives used over the four time periods in *Science* and *EHR*

Finite passives[a]	*Science*			*EHR*		
	Verb	%	Label	Verb	%	Label
1880/1886	*made*	7.02	Pro	*said*	2.18	Ind
	placed	3.40	Pro	*found*	1.42	Ind
	published	1.91	Ind	*told*	1.32	Ind
1930	*found*	7.00	Ind	*made*	3.10	Pro
	made	4.42	Pro	*found*	2.04	Ind
	given	3.07	Pro	*given*	1.55	Pro
1980	*used*	5.38	Pro	*made*	2.37	Pro
	shown	2.43	Ind	*known*	1.72	Ind
	found	2.32	Ind	*given*	1.50	Pro
2017/2018	*expected*	4.47	Der	*made*	2.20	Pro
	observed	4.34	Ind	*seen*	1.52	Ind
	used	3.29	Pro	*used*	1.40	Pro
Bare passives[b]						
1880/1886	*made*	5.95	Pro	*made*	2.22	Pro
	used	4.76	Pro	*published*	1.55	Ind
	observed	4.17	Ind	*written*	1.55	Pro
1930	*based on*	5.48	Der	*made*	4.20	Pro
	involved	3.96	Ind	*given*	3.44	Pro
	obtained	3.66	Der	*copied*	3.31	Pro
1980	*used*	4.79	Pro	*written*	3.72	Pro
	calculated	3.92	Der	*offered*	3.07	Pro
	obtained	3.70	Der	*given*	2.98	Pro
2017/2018	*shown*	4.91	Ind	*given*	3.76	Pro
	compared	4.18	Ind	*made*	2.19	Pro
	observed	3.69	Ind	*written*	2.04	Pro

[a] The figures are represented as a percentage of the total number of finite passives used.

[b] The figures are represented as a percentage of the total number of bare passives used.

[4]   The king paid him an enormous sum of money for his help;

[5]   in the autumn of 1339, he owed the duke a staggering 307,000 Florentine florins.

[6]   Edward's first serious military campaign against Philip VI of France (r. 1328–50) **was launched** in September 1339 from the Brabantine town of Leuven,

[7]   where the English king had established his headquarters.

[8]   Duke Jan's men, <…>, **were** primarily **recruited** from among the duke's 'own' Brabantine vassals,

[9]   <who supported King Edward>

[10]   although there were also men from neighbouring principalities, such as the counties of Loon, Mark, Namur and Hainaut. […]

[45]   We focus in particular on the formal ties between princes and nobles in relation to the build-up of princely hosts, and on the ways

[46]   in which these ties developed

[47]   and changed.

[48]   These relationships, both feudal and non-feudal, **have** previously **been studied** only in part (in the case of fiefrentes, for example) or in the context of particular military conflicts.

[49]   In looking at the full gamut of these relationships and at their evolution across an entire century,

[50]   this article also contributes to the wider debate about the military position of the nobility in the Low Countries and the nature of their relationship with the princes.

[51]   Whereas the role of the nobility in the growing Burgundian state in the fifteenth century **has been examined** in depth,

[52]   noble–prince relations in the fourteenth century remain relatively little studied. [2018, 11]

The extract in (16) contains four passive verb phrases. The description starts with the actions of Duke Jan III and Edward III using a range of productive verbs, including the verbs *launched* and *recruited* in the first two passives. The description then narrows towards the relationships between princes and nobles, with the indicative verbs *studied* and *examined* in the last two passives.

Where the stative verbs in *EHR* are concerned, the indicative function is far more prevalent than the derivative function; in fact, as can be seen in Table 2, none of the frequently used verbs in *EHR* belong to the *derivative* category. This might appear surprising, since history, like science, is also focused on outcomes and consequences. The reason is likely to be a more straightforward one, attributed to the way such outcomes and consequences are worded in writing – they are simply more commonly expressed in the active voice rather than the passive voice. That is to say, it is far more common to write *X led to Y* or *X resulted in Y*, neither of which has a passive version.

5 Conclusion

This study investigated the trend in the use of passives over four time periods (1880/1886, 1930, 1980, and 2017/2018) in *Science* and *EHR* articles. The following are the broad findings of the study:

1. The use of the passive voice declined in both *Science* and *EHR*. In *Science*, the fall was most marked between 1980 and 2017; the percentage of all clauses containing passives decreased from 35.82% in 1980 to 23.21% in 2017. In *EHR*, the decline was more gradual; the percentage of all clauses containing passives fell from 25.64% in 1930 to 17.65% in 2018.

2. Over the four time periods, fewer passives were used in the *EHR* articles than in the *Science* articles.

3. The use of bare passives was stable across the four time periods in both *Science* and *EHR*, but fewer bare passives were used in *EHR* than in *Science*. Bare passives constituted 8–11% of all clauses in *Science*, and 5–7% of all clauses in *EHR*.

4. The finite passives in *Science* largely served the productive function, whereas the same in *EHR* served both the productive and indicative functions.

5. The bare passives in *Science* shifted from the productive function in 1880 to a mix of the indicative and derivative functions by 2017. The bare passives in *EHR* remained largely productive verbs across the four time periods.

The findings in this study therefore affirm anecdotal accounts of writing in the humanities (represented here using history) and the sciences. We are often told that the two types of writing are different in their use of the passive voice (e.g. Grech 2019), but it has not always been clear what the extent of that difference is. What I have sought to do in this study is to provide the statistics to show this difference.

More crucially, though, what are the implications for writing in the sciences and humanities? I have suggested that the passive in scientific writing appears to be socially conditioned by the scientific community to reflect the thing-centeredness in their work. I have also noted the marked fall in passive use from 1980 to 2017, and this is most likely due to an increased awareness of the changing sentiments against the *overuse* of the passive voice. It is this overuse that has led to numerous recommendations in favor of the active voice. In her survey of academic style manuals, for instance, Bennett (2009) lists a number of authors who advocate the use of the active voice instead of the passive. To my mind, such a swing of the pendulum from the passive to the active is extreme and perhaps unwise. The passive cannot be totally avoided, nor can it be easily replaced in all contexts. The passive is the only option to use if the actor is unknown; it also appears in formulaic expressions that would look decidedly odd in the active version (e.g. *As **compared** to Y ~ As we compared X to Y*). Although there is presently less reliance on the use of the passive in scientific writing, it is likely to continue to be used in various parts of the scientific paper, particularly the methodology section.

By contrast, the passive in history writing is not constrained by the need for thing-centeredness. There is no methodology section in a typical history article, and the focus is on an event, figure, or place of some historical significance. The historian recounts and interprets, but does not act to effect change; there is therefore little pressure on the historian to turn to the passive voice to background himself or herself as the actor. Therein lies the need for writers to be aware of the conventional norms in their disciplines. These norms include not merely matters to do with the rhetorical structure of the paper, the citation format, or argumentation, but also grammatical features such as the passive voice.

Much more remains to be done. As regards the decline in the use of the passive voice, an obvious question that is raised is whether this implies an increased presence of the author as the agent of the action. Seoane & Hundt (2018: 13) note that "it is extremely uncommon to have the author as subject of an active clause in scientific discourse". However, a smaller-scale study

on the thematic structure of scientific research articles has shown that the first-person pronoun (e.g. *we*) often serves an anchoring thematic function, which is "the text-level version of the constant [thematic progression], and highlights the inclination of writers to rely on certain points of departure to frame their writing" (Leong et al. 2018: 294). This is an interesting puzzle, and a further look at how passive use influences the thematic structure of academic texts is thus needed across disciplines.

Lastly, an obvious limitation of this study is that it compared articles from only science and history. Writing in the humanities, however, includes other diverse disciplines that need a closer investigation as well. Does the writing in philosophy and literature, for instance, differ to any extent from history? Future work should include as many of such disciplines, and others in the social sciences (e.g. psychology, sociology), for a broader, but better, understanding of academic writing as a whole. Diachronic studies, wherever possible, are recommended as trend lines can often be highly insightful to show whether any observed changes are likely to stabilize or change even further in the future.

References

Atkinson, Dwight. 1996. The *Philosophical Transactions of the Royal Society of London*, 1675–1975: A sociohistorical discourse analysis. *Language in Society* 25(3). 333–371.

Banks, David. 2008. *The development of scientific writing: Linguistic features and historical context*. London: Equinox.

— 2017. The extent to which the passive voice is used in the scientific journal article, 1985–2015. *Functional Linguistics* 4(12). DOI: 10.1186/s40554-017-0045-5.

Baratta, Alexander M. 2009. Revealing stance through passive voice. *Journal of Pragmatics* 41(7). 1406–1421. DOI: 10.1016/j.pragma.2008.09.010.

Baskin, Tobias I. 1996. Hazards of the passive voice. *Nature* 381(6585). 730. DOI: 10.1038/381730c0.

Bazerman, Charles. 1988. *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, WI: University of Wisconsin Press.

Bennett, Karen. 2009. English academic style manuals: A survey. *Journal of English for Academic Purposes* 8(1). 43–54. DOI: 10.1016/j.jeap.2008.12.003.

Biber, Douglas & Gray, Bethany. 2016. *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press.

Biber, Douglas & Johansson, Stig & Leech, Geoffrey & Conrad, Susan & Finegan, Edward. 1999. *The Longman grammar of spoken and written English*. Harlow: Pearson.

Bridgman, Percy Williams. 1955. *Reflections of a physicist*. New York: Philosophical Library.

Cargill, Margaret & O'Connor, Patrick. 2009. *Writing scientific research articles: Strategies and steps*. Chichester: Wiley-Blackwell.

Carter, Ronald & McCarthy, Michael. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Ding, Daniel D. 1998. Rationality reborn: Historical roots of the passive voice in scientific discourse. In Battalio, John T. (ed.), *Essays in the study of scientific discourse: Methods, practice, and pedagogy*, 117–135. Stamford, CT: Ablex.

— 2002. The passive voice and social values in science. *Journal of Technical Writing and Communication* 32(2). 138–154. DOI: 10.2190/EFMR-BJF3-CE41-84KK.

Goodman, Neville W. 1996. Hazards of the passive voice. *Nature* 381(6585). 730. DOI: 101038/381730c0.

Gray, Bethany. 2015. *Linguistic variation in research articles: When discipline tells only part of the story*. Amsterdam: John Benjamins.

Grech, Victor. 2019. Writing a scientific paper (WASP): An overview of differences in styles between the sciences and the humanities. *Early Human Development* 129. 96–99. DOI: 10.1016/j.earlhumdev.2018.12.008.

Greenbaum, Sidney & Quirk, Randolph & Leech, Geoffrey & Svartvik, Jan. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

Gross, Alan G. 1996. *The rhetoric of science*. Cambridge, MA: Harvard University Press.

Gross, Alan G. & Harmon, Joseph E. & Reidy, Michael. 2002. *Communicating science: The scientific article from the 17th century to the present*. Oxford: Oxford University Press.

Hayot, Eric. 2014. *The elements of academic style: Writing for the humanities*. New York, NY: Columbia University Press.

Hou, Liwen & Smith, David A. 2018. Modeling the decline in English passivization. *Proceedings of the Society for Computation in Linguistics (SCiL)* 1. 34–43. DOI: 10.7275/R5ZC812C.

Hundt, Marianne & Schneider, Gerold & Seoane, Elena. 2016. The use of the *be*-passive in academic Englishes: Local *versus* global usage in an international language. *Corpora* 11(1). 29–61. DOI: 10.3366/cor.2016.0084.

Hunt, Kellogg W. 1977. Early blooming and late blooming syntactic structures. In Cooper, Charles R. & Odell, Lee (eds.), *Evaluating writing: Describing, measuring, judging*, 91–104. Urbana, IL: National Council of Teachers of English.

Hyland, Ken. 2002. Options of identity in academic writing. *ELT Journal* 56(4). 351–358. DOI: 10.1093/elt/56.4.351.

Iddings, Joshua Glenn. 2007. *A functional analysis of English humanities and biochemistry writing with respect to teaching university composition*. Huntington, WV: Marshall University. (Master's thesis).

Jolly, Clifford. 1996. Hazards of the passive voice. *Nature* 381(6585). 730. DOI: 10.1038/381730c0.

Leather, Simon R. 1996. The case for the passive voice. *Nature* 381(6582). 467. DOI: 10.1038/381467a0.

Leong, Ping Alvin. 2014. The passive voice in scientific writing: The current norm in science journals. *Journal of Science Communication* 13(1). DOI: 10.22323/2.13010203.

— 2015. Topical themes and thematic progression: The 'picture' of research articles. *Text & Talk* 35(3). 289–315. DOI: 10.1515/text-2015-0001.

— 2016. Thematic density of research-article abstracts: A systemic-functional account. *Word* 62(4). 209–227. DOI: 10.1080/00437956.2016.1248668.

— 2020. The passive voice in scientific writing through the ages: A diachronic study. *Text & Talk* 40(4). 467–489. DOI: 10.1515/text-2020-2066.

Leong, Ping Alvin & Toh, Audrey Lin Lin & Chin, Soo Fun. 2018. Examining structure in scientific research articles: A study of thematic progression and thematic density. *Written Communication* 35(3). 286–314. DOI: 10.1177/0741088318767378.

MacDonald, Susan Peck. 2010. *Professional academic writing in the humanities and social sciences*. Carbondale, IL: Southern Illinois University Press.

Mair, Christian & Leech, Geoffrey. 2006. Current changes in English syntax. In Aarts, Bas & McMahon, April (eds.), *The handbook of English linguistics*, 318–343. Malden, MA: Blackwell.

Montgomery, Scott L. 1996. *The scientific voice*. New York: Guilford Publications.

Ormes, Robert V. 1957. Passive voice. *Science* 125(3247). 529. DOI: 10.1126/science.125.3247.529.

Perlman, Alan M. 1996. Passive voice. *Nature* 382(6587). 108. DOI: 10.1038/382108b0.

Puckica, Jérôme. 2009. Passive constructions in Present-Day English. *Groninger Arbeiten zur Germanistischen Linguistik* 49. 215–235.

Rachul, Christen M. 2008. *Comparing apples and oranges: A socio-rhetorical analysis of writing in biology and history*. Ottawa: Carleton University. (Master's thesis).

Robinson, Jane T. 1957. Passive voice and personal pronouns. *Science* 125(3258). 1160. DOI: 10.1126/science.125.3258.1160.

Royal Literary Fund. 2021. *Scientific writing*. (https://www.rlf.org.uk/resources/scientific-writing/). (Accessed 2021-12-22).

Scimago. 2020. *Scimago journal & country rank*. (http://www.scimagojr.com/). (Accessed 2021-06-09).

Seoane, Elena. 2006. Changing styles: On the recent evolution of scientific British and American English. In Dalton-Puffer, Christiane & Kastovsky, Dieter & Ritt, Nikolaus & Schendl, Herbert (eds.), *Syntax, style and grammatical norms: English from 1500–2000*, 191–211. Bern: Peter Lang.

— 2013. On the conventionalisation and loss of pragmatic function of the passive in Late Modern English scientific discourse. *Journal of Historical Pragmatics* 14(1). 70–99. DOI: 10.1075/jhp.14.1.03seo.

Seoane, Elena & Hundt, Marianne. 2018. Voice alternation and authorial presence: Variation across disciplinary areas in academic English. *Journal of English Linguistics* 46(1). 3–22. DOI: 10.1177/0075424217740938.

Seoane, Elena & Loureiro-Porto, Lucía. 2005. On the colloquialization of scientific British and American English. *ESP Across Cultures* 2. 106–118.

Swales, John M. 1981. *Aspects of article introductions*. Birmingham: Aston University.

— 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, John M. & Feak, Christine B. 2004. *Academic writing for graduate students: Essential tasks and skills*. 2nd edn. Ann Arbor, MI: University of Michigan Press.

— 2009. *Abstracts and the writing of abstracts*. Ann Arbor, MI: University of Michigan Press.

Wanner, Anja. 2009. *Deconstructing the English passive*. Berlin: Mouton de Gruyter.

Zaiontz, Charles. 2018. *Real Statistics Resource Pack*. (http://www.real-statistics.com/). (Accessed 2020-12-14).

Contact information:

Ping Alvin Leong
Language & Communication Centre
School of Humanities
Nanyang Technological University
48 Nanyang Avenue, Singapore 639818
e-mail: alvin(dot)leong(at)ntu(dot)edu(dot)sg