

Article

Not peer-reviewed version

Tracing the evolution of reviews and research articles in the biomedical literature: a multi-dimensional analysis of abstracts

[Stefano Guizzardi](#) , [Maria Teresa Colangelo](#) , [Prisco Mirandola](#) , [Carlo Galli](#) *

Posted Date: 1 August 2023

doi: 10.20944/preprints202307.2149.v1

Keywords: Abstract; Narrativity; Scientific Publishing



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Tracing the Evolution of Reviews and Research Articles in the Biomedical Literature: A Multi-Dimensional Analysis of Abstracts

Stefano Guizzardi, Maria Teresa Colangelo, Prisco Mirandola and Carlo Galli *

Department of Medicine and Surgery, Histology and Embryology Lab, University of Parma, Via Volturmo 39, 43126 Parma, Italy; stefano.guizzardi@unipr.it; mariateresa.colangelo@studenti.unipr.it; prisco.mirandola@unipr.it

* Correspondence: carlo.galli@unipr.it;

Abstract: We have previously reported on the diachronical changes in the narrative structure of research articles (RAs) and review manuscripts, using corpora of abstracts from MEDLINE. The present study investigates 5 linguistic dimensions (D1-5) of the text, following Biber's well established Multidimensional Analysis, on the same corpora, to assess how writing practices have changed in these two biomedical literature sub-genres over the years. Our analysis encompassed a sample of more than 1.2 million substracts from manuscripts that were published over the course of more than 30 years. Both RAs and reviews have reinforced their informational, emotionally detached tone (D1), and have progressively refrained from the use of narrative devices (D2), while increasing their context-independent content (D3). Both RAs and reviews have displayed low levels of overt persuasion (D4), while steering away from abstract content to focus on a more marked author agency and identity adfirmation. When the linguistic features that underly these 5 dimensions are compared, it becomes apparent that RAs and review papers have often changed quite independently, both usually converging to standardized stylistic canons.

Keywords: abstract; narrativity; scientific publishing

1. Introduction

Biomedical publishing plays a crucial role in the field of biomedicine and has significant importance for the dissemination of scientific knowledge [1]. The publication of research articles allows for the systematic evaluation, synthesis, and analysis of data from multiple studies. This evidence base is vital for informing clinical guidelines, treatment protocols, public health policies, and healthcare interventions [2]. The publication of biomedical literature, however, also allows researchers and scientists to share their findings, discoveries, and innovations with the global scientific community. This dissemination is essential for advancing scientific understanding and fostering collaboration. Publishing research findings also promotes transparency and accountability within the scientific community, as researchers are subject to scrutiny and evaluation by their peers. This transparency encourages researchers to conduct rigorous, well-designed studies, and discourages unethical practices or misrepresentation of data [3].

However, biomedical publishing is also an important factor in career advancement for researchers, scientists, and healthcare professionals. Publications in reputable journals enhance researchers' visibility, credibility, and professional reputation. The number and impact of publications are often considered during academic promotions, grant applications, and funding decisions [4,5]. This means that there are strong incentives to publication, and their number has steadily grown in the last years, reaching levels that make it very difficult for researchers in any field to keep abreast of the published material [6], a problem made even more acute by the surge of predatory journals [7]. This abundance of literature has many consequences, such as the flourishing of review articles, which are, in their classic narrative form, summaries of the literature, which may

prove essential to many researchers that try to get an overview of an issue but also pose a challenge to readers because reviews are prone to redundancy by nature [8,9].

Being able to effectively communicate one's findings, thoughts and opinions then becomes of the utmost importance to ensure that one's voice be heard and stand out from the background noise of the thousands of papers that are published at any given moment. Communication practices evolve and keep adapting to the current situations and understanding what the parameters of communications are can be very relevant to optimize them and maximize the impact of one's scientific production. This holds especially true for abstracts, that one piece of text that readers go through after the title, to identify whether a manuscript is of interest. To better understand the linguistic features that characterize abstracts in research articles and reviews we resorted to Biber's multidimensional analysis.

Biber's multidimensional text analysis is a comprehensive framework used in corpus linguistics to examine and understand various dimensions of linguistic variation in written and spoken language [10]. Developed by Douglas Biber, this multidimensional analysis aims to identify and describe these variations across multiple linguistic dimensions, including lexical, syntactic, morphological, and discourse features [11]. Biber's multidimensional analysis involved extracting linguistic features from 481 spoken and written texts of contemporary British English, which were then used to compute 5 dimension scores, through a factor analysis of the co-occurrences of these features. The texts that Biber used were taken from the Lancaster-Oslo-Bergen Corpus [12] and the London-Lund Corpus [13], which were chosen because they represent over 20 major register categories, including academic writing in many fields, fiction, letters, conversations, etc.). Biber proceeded to catalogue his 67 features according to the relative factor loading on each dimension, as a measure of the strength of the association between a factor and a dimension. This allowed Biber to trace 5 main dimensions [14], which he interpreted as listed in Table 1.

Table 1. Outline of the 5 dimensions of Biber's Multidimensional Analysis that was used in the present paper [11].

Dimension	Feature
1	Involved vs. Informational discourse
2	Narrative vs. Non-Narrative Concerns
3	Context-Independent Discourse vs. Context Dependent Discourse
4	Overt Expression of Persuasion
5	Abstract and Non-Abstract Information

Biber's tool has been extensively used to investigate corpora of different kind and origin, such as abstracts published in different countries [15] or by writers of different origin [16], and has proved very useful, because it concisely provides an overview of the general linguistic and rhetorical stances of a text, in the broader context of the literature production in many fields and genres. Nini made this tool easily and freely available through a Multidimensional Analysis Tagger, which replicates Biber's 67 original features used to compute his dimension scores [17], and which we relied upon too for the present work.

As we previously characterized the narrative arcs in research articles and literature reviews in the biomedical field, by applying the LIWC 2022 analysis tool [18] to a corpus of abstracts from research articles and reviews obtained from MEDLINE over the course of the last 33 years [19], we moved on to apply Biber's analysis to this corpus to gain further insights into the linguistics changes in the 1989-2022 periods in these two popular genres and how these changes differed.

2. Materials and Methods

The datasets that we used for the present study was composed by two independent corpora of abstracts of scientific manuscripts obtained from MEDLINE and previously published [20]. Briefly,

we used the python 3.9 litter-getter library [21] to search and retrieve the abstracts from Medline using its Pubmed API. We relied on the following search terms

```
#1 'year[dp] NOT Review[pt]';
```

```
#2 'year[dp] AND Review[pt]';
```

where the 'year' parameter was set to iterate from 1989 through 2022 and retrieved two lists of Pubmed IDs (PMIDs). Search #1 generated a list of abstracts PMIDs excluding the 'Review' type and search #2 generated a list of PMIDs exclusively constituted of 'Review' abstracts in the same time interval. The reason for this distinction is that Reviews are a genre of scientific article that comprises several peculiar sub-types including 'Narrative reviews' or "Systematic reviews", each with its distinctive purpose and structure [22], and we hypothesized that reviews may display different linguistic features than research articles, in agreement with the differences in narrativity, highlighted by our previous LIWC analysis [19].

As previously explained [19], to balance our corpora, we randomly sorted 20000 PMIDs out of the total number of retrieved PMIDs for each year, and proceeded to retrieve the data from the PMID list, thus obtaining 2 independent corpora as follows:

```
#1 Abstracts from Research articles (excluding Reviews), published between 1989-2022 (n=680000)
```

```
#2 Abstracts from Review articles, published between 1989-2022 (n=680000)
```

To obtain the abstracts, Litter-getter downloaded an XML file based on its PMID from MEDLINE, and we then created a Pandas Dataframe [23] using the BeautifulSoup library [24] for data extraction.

The abstracts were lowercased and passed into the Multidimensional Analysis Tagger v 1.3.3 [17]. This tagger is based on Biber's (1988) Variation across Speech and Writing tagger for the multidimensional functional analysis of English texts. This program is based on the Stanford Tagger [17] and generates both a grammatically annotated version of the text as well as the statistics following Biber [10]. The output of this tool is a series of scores for the 5 dimensions outlined by Biber, plus scores for each of the underlying 67 linguistic features, all expressed as Z scores. A Z score is, simply put, a measure of the distance of a score for a given sample from the mean of that score for a whole population [25], expressed as number of standard deviations from the mean. So, in our case, the MAT software contains the means for each score for a vast corpus of texts from various genres, including conversations, speeches, personal letters, broadcasts and academic writing [17]. As an example, a Z score of 2 for any linguistic feature means that this score is 2 standard deviations above the mean of that mixed corpus, which is representative of a general literature production.

Matplotlib [26] and Seaborn [27] libraries were then used to plot the data. All the analyses were conducted on Jupyter notebooks [28].

3. Results and Discussion

The two corpora comprised 680000 abstracts each, without overlap, because of the way they were selected. The selection criteria for corpus #1, however, had a consequence, i.e. that this corpus contained not only RAs, but also a small number of different genres. A post hoc analysis on the corpus showed that 611450 abstracts out of the total 680000 in #1 corpus belonged to research articles, and 43567 abstracts (7.1%) belonged to the comment, letter and editorial categories, which do not fall within our area of interest, while the remaining 24983 could be classified as less frequent manuscript types, e.g. news, or historical articles [19].

3.1. Dimension 1

Dimension 1 (D1) refers to informational vs involved discourse. The positive pole of this dimension would be most typically associated with dialogues, which are rich in a language that focuses on interaction and expressing an affective content, rather than just delivering information [11]. The negative pole of this dimension, on the other hand, is associated to information-rich and highly edited text, as it would be typically expected from academic pieces [29]. Both research articles (RA) and reviews in our corpus have extremely low scores (Figure 1A), which have remained

relatively constant over time for RA, but have been slightly declining for reviews. To better understand how dimensions changed in our two article groups, we resorted to scatterplots that show the value of different items for RA and reviews over time. The scatterplot in Figure 1B shows that RA with different Dimension 1 scores can be found for all publication dates, but older reviews tend to cluster around comparatively higher values that more recent ones, indicating a slightly more involved style for older reviews. To make the dimension score a bit more transparent and gain some further insight in the phenomena that have occurred in the literature, we decided to analyze the underlying linguistic features associated to D1.

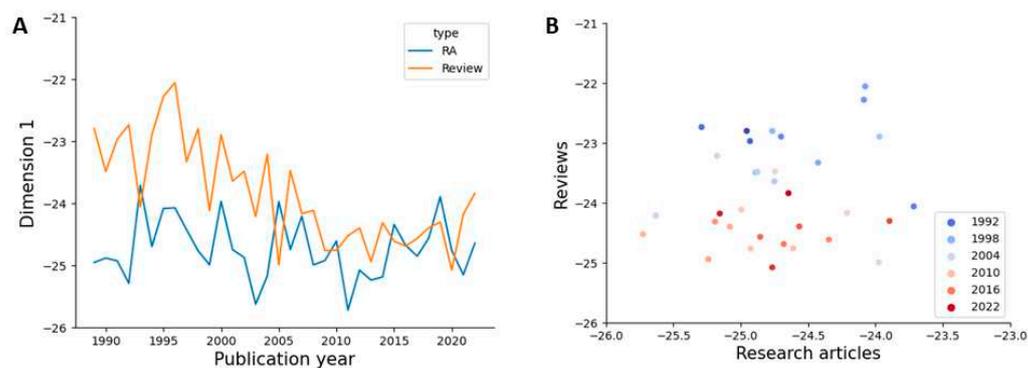


Figure 1. A) Line plot of Dimension 1 (D1) score over the years for the Research article (RA) corpus and the review corpus, in blue and orange respectively; B) scatter plot of D1 score for RAs and reviews.

A frequent use of **nouns** and **long words** is, unsurprisingly, a characteristic feature of texts with negative scores for Dimension 1, as they require planning in production and are less suited for improvised speeches and dialogues. Consistently with this, our analysis reveals that, over the years, both RA and reviews have further increased their Z scores for these features in a similar way (Figure 2A, B), although no trend of any sort can really be detected for the **type/token ratio** or the frequency of **attributive adjectives** (Figure S1), which are also typical for negative D1 scores.

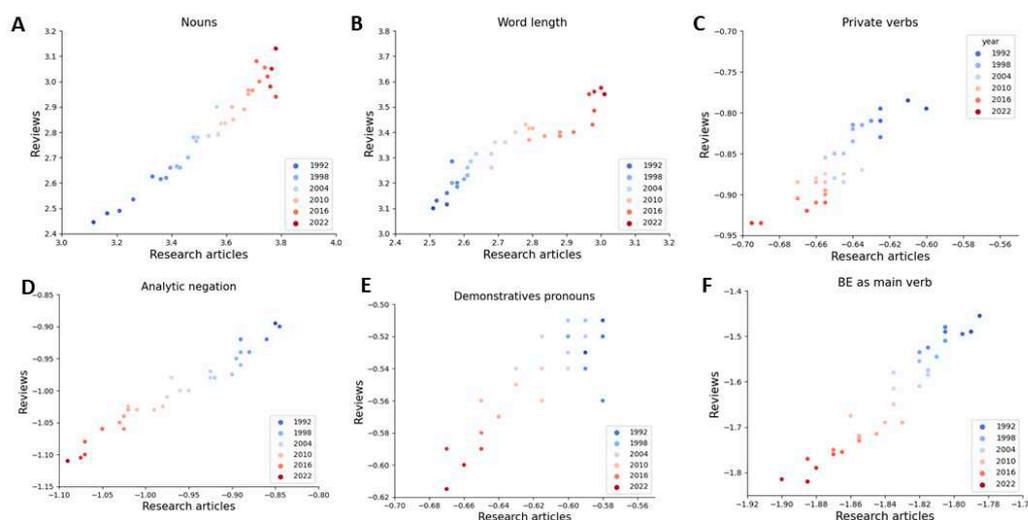


Figure 2. Scatter plots of linguistic features of Dimension 1 in RA and review corpora by publication years. These linguistic features change similarly in the 2 corpora.

Similarly, a movement toward an even more information-rich writing style can be detected based on the decrease in other typical features of involved texts, such as the **analytic negation** (Figure 2D), the use of **demonstrative pronouns**, which can often have a deictic use, typical of spoken

language and interaction (Figure 2E), **private verbs** (i.e. that express internal, cognitive and thus private processing, e.g. to think, to feel, to perceive etc.) and **be as the main verb** (Figure 2F). Interestingly, the use of **prepositions**, which is typically high in texts with strongly negative scores of D1, has been constantly decreasing in both corpora (Figure 3A), while **non-phrasal coordination**, which is associated to involved writing, increased in both corpora (Figure 3B).

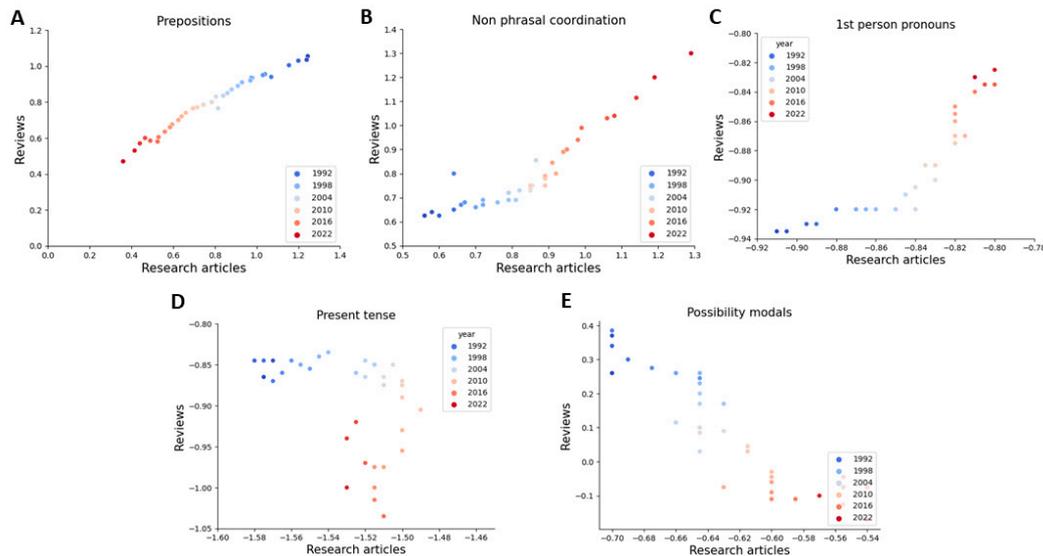


Figure 3. Scatter plots of the linguistic features of Dimension 1 in RA and review corpora by publication years. These features change differently in the 2 corpora.

RA and reviews, however, behaved differently in at least three features. The Z score for **1st person pronouns** (a characteristic of dialogue, and, more generally, involved writing) is negative in both corpora, as expected, although there are a few -striking - examples of use as in:

It was my second clinical placement and I was working on a surgical ward when I was asked to accompany a patient to theatre [30].

which is quite frankly an unusual style for academic prose, yet is found in our corpus.

However, the frequency of 1st person pronouns increased over the course of the 90s in RA, but remained quite constant in reviews, and only in the first years of the new century it started to increase in both text types. (Figure 3C) The most likely explanation for this behavior is that, although passive verbs have been used abundantly in academic writing as a rhetorical device to highlight the detachment of the narrator from the events contained in the text and as a sign of objective observation [31], the use of active verbs and 1st person pronouns has been advocated in more recent times for the sake of clarity [32] and has been observed to be on the rise in academic writing [33]. It may be assumed that RAs were more prone to the use of 1st person pronouns, as they often reported on the experimental activity of a research group, as opposed to reviews, which typically summarize the findings of other research groups, and thus this increased occurred earlier.

The use of **present tense** verbs is strongly associated to the positive pole of Dimension 1 too (as it is very frequent in interactions between speakers) and, though generally low in both corpora, our data indicate that it is higher in review papers (Figure 3D). A possible explanation for this discrepancy is that reviews often summarize the current knowledge in a certain area, using descriptive discourse, and may use the literature to draw conclusions that may be proposed as general rules, as in the following:

Primary care clinicians treat patients with cancer and cancer pain. It is essential that physicians know how to effectively manage pain including assessment and pharmacologic and nonpharmacologic treatment modalities [34].

where the use of the present tense is appropriate to convey the sense of lasting value that these conclusions have, while the purpose of RAs is usually to report on one or more experiments, which are situated in time and place, and are thus often described using past tenses as in:

During 8 observation days (with time delay of 10-14 days between each observation day), all adult patients hospitalized at an internal medicine ward of 4 Belgian participating hospitals were screened for AB use. Patients receiving AB on the observation day were included in the study and screened for signs and symptoms of AAD using a period prevalence methodology [35].

The use of present tense increased slightly in RAs in the 90s, and remained stable afterwards, while it started to drop in reviews around the same time. Any explanation is purely speculative at this stage and may be related to the increase of systematic reviews, which may be more grounded in the research articles they are based on, or to a purely stylistic change. Noticeably, the Z score for RA remains significantly lower than for reviews (Figure 3D).

The use of **possibility modals** is associated to positive Dimension 1 scores too, as these are often utilized to express subjectivity, a guess, which is a common situation in a dialogue context. However, they can also be found quite regularly in academic writing [36], usually to express a hypothesis, as in:

Administration of thioredoxin may have a good potential for anti-aging and anti-stress effects [37].

Admittedly, the room for hypothesis, although a common and actually quite essential practice in the scientific method [38], is quite limited in academic literature, given the need for evidence-grounded reasoning, hence the low z scores. Interestingly, the use of possibility modals has been moving in the opposite directions in the two corpora we analyzed: the Z score for this feature was slightly positive in reviews, as it could be expected in a text genre that used to be quite prone to drawing conclusions based on the reviewed data, but negative in RAs, underlying that assumptions and hypothesis were likely confined to few sentences in such texts. However, this index steadily decreased in the review group, reaching negative values in the last decade, maybe in association to the increase in systematic reviews, where the sometimes massive use of statistical tools might make guessing more rare, while it increased by almost 30% in the RA corpus, possibly in association to the use of a bolder, or more personal style, as previously noted (Figure 3E) [33].

3.2. Dimension 2

Dimension 2 (D2) is associated to narrative discourse, so a positive score for this dimension indicates that the text has a narrative, active, event-oriented nature rather than a more descriptive or static quality [11].

Our corpora of RAs and reviews have a negative Z score for D2, with reviews having a lower score than RAs (Figure 4A). This is not unexpected, as RAs more likely report, by definition, on the execution of one or more experimental procedures, which are usually associated with some sort of activity, as in:

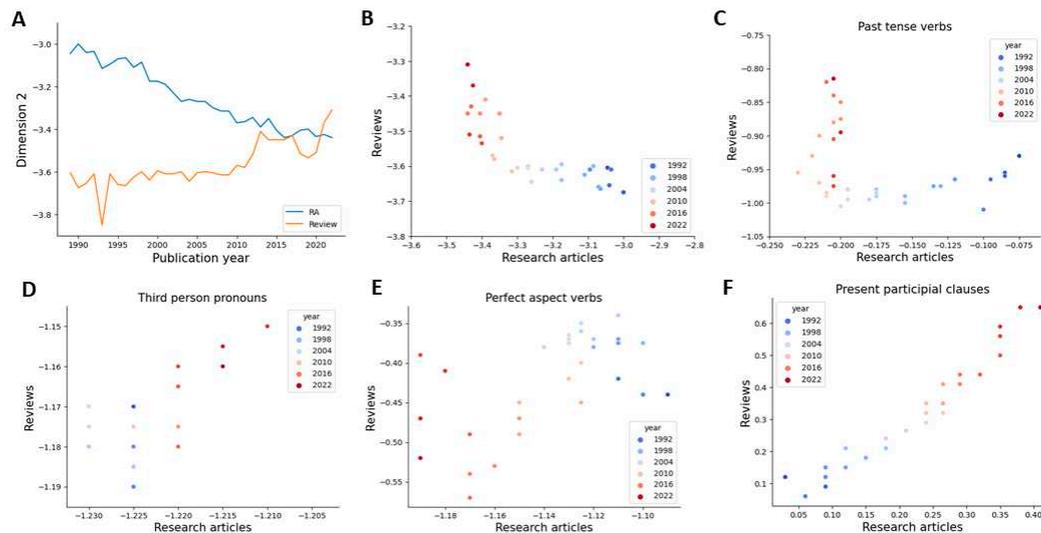


Figure 4. A) Line plot of Dimension 2 (D2) score over the years for the Research article (RA) corpus and the review corpus, in blue and orange respectively; B) scatter plot of D2 score for RAs and reviews. C-F) Scatter plots of the linguistic features of D2 in RA and review corpora by publication years.

We *investigated* expression of the five ssts in various adrenal tumors and in normal adrenal gland. Tissue *was obtained* from ten pheochromocytomas (PHEOs)...[39]

The text is all about action, about doing, selecting, analyzing and other similar activities, which are historically situated and hence require a narration to go through them.

Interestingly, the Z score for RAs tended to progressively decrease and become more negative with time, while the D2 score for reviews remained constant, and even increased – becoming less negative - in the last 5 years (Figure 4A). This is also reflected in Figure 4B, which shows that RAs' D2 score dropped in the 90s and early 2000s and reviews' score started to increase quite independently from RAs in the first decade of the years 2000. This trend is possibly justified by the change in Z score for the use of **past tense** verbs, which has the highest bearing on D2 [10]. This score, which was and has remained negative in both corpora for the whole timeframe (Figure 4C), decreased in RAs until the first decade of the XXI century and it was followed by an increase in this score for review articles in the last two decades.

This means that an abstract from a review article in 1989 could more easily contain a passage like:

Several lines of evidence *indicate* that platelet-activating factor (PAF-acether) *is implicated* in hypersensitivity reactions. Indeed, PAF-acether *reproduces* the features of asthma *in vivo* and *in vitro*, since it *induces* bronchoconstriction, hypotension, and hemoconcentration and activates platelets and leukocytes [40].

which is rich in present tense verbs that are used to convey general principles about a phenomenon, e.g. a disease or a condition, while a more recent review text more easily incorporates some past tenses, such as in the case of:

Mammalian neonates *have been simultaneously described* as having particularly poor memory, as evidenced by infantile amnesia, and as being particularly excellent learners [41].

This phenomenon could be hypothesized to indicate that, since the early 2000s, review articles have been tending to circumscribe their conclusions to the research papers they use as sources, contextualizing them, and possibly being more wary of generalizations.

Other important linguistic features associated with D2 underwent similar changes in both corpora: the use of **third person pronouns** increased for both text types (Figure 4D), as did the use of **present participial clauses** (Figure 4F), while the frequency of **perfect aspect verbs** decreased in both

RAs and reviews, although the scores for this feature remained significantly lower in RAs than in reviews (Figure 4E).

Noticeably, these findings are also apparently in contrast with what we reported on the same corpora using LIWC 2022 [19]. In particular, we reported a higher **Narrativity Overall** score for reviews. That score was calculated based on the adherence to a peculiar metrics, i.e. the three fundamental narrative curves that were measured in each abstract, namely Staging, Plot Progression and Cognitive tension [42]. The theory behind these measures is that a narrative trajectory can be traced in a text, which follows Freytag's dramatic arc: first the stage for the action is set, characters and referents are introduced and presented; the action then begins and as the text progresses it intensifies, as the narrator described events and activities; and cognitive tension refers to the struggles and conflicts that ensue in the story and that reach a culmination point with the resolution of the crisis that leads to the end of the narration [43]. To get an automated measure of these features, Pennebaker et al. decided to rely on grammatical words, which admittedly form a small set of words in English (and any language) [44]. In particular Boyd et al. proposed to measure the frequency of articles and prepositions as proxies for the staging score, because they can be assumed to be more abundant when new referents are introduced in the text (via articles) and their relations are explained (possibly also through the use of prepositions), while auxiliary verbs and anaphoric pronouns are taken as proxy measures of plot progression, because they can be expected to be used when describing an action. Cognitive tension is measured on the abundance of verbs in a special dictionary created ad hoc and that includes words as 'think', or 'believe' (which would be classified as 'private verbs' in Biber's Multidimensional analysis. Boyd et al. recommend splitting the texts in at least 5 segments, to monitor how these scores vary as the text progresses. It is therefore apparent that LIWC 2022 and Biber's narrativity scores are based on different features and the readers should appreciate the characteristics of the text these tools are actually measuring rather than getting hung up on the 'narrativity' label.

3.3. Dimension 3

A high score for Dimension 3 (D3) is associated to explicit and context-independent reference, as opposed to nonspecific, context-dependent content [10]. This means that referents in the text are mentioned and described explicitly, so that there cannot be any doubt about their identity. According to our data, reviews have a higher D3 score than RAs, and both their scores have been progressively increasing over time (Figure 5A, B). Among the features that affect D3, **nominalization** appears to have followed this trend and may be responsible for the visible changes in D3 over time.

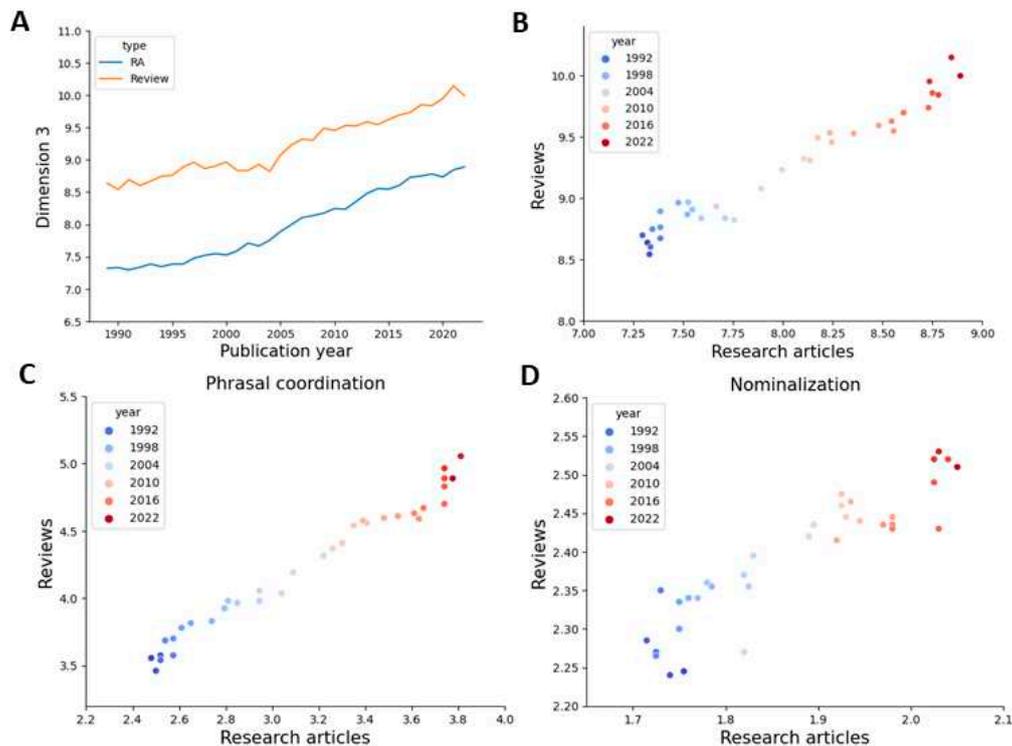


Figure 5. A) Line plot of Dimension 3 (D3) score over the years for the Research article (RA) corpus and the review corpus, in blue and orange respectively; B) scatter plot of D3 score for RAs and reviews. C, D) Scatter plots of the linguistic features of D3 in RA and review corpora by publication years.

Nominalization [45] indicates the replacement of a verb with a noun that denotes the same action, and is a common feature of technical language [46], which is often used to convey a more impersonal tone, because a noun, by describing an action as an entity, detaches it from the agent and confers it a higher independence [47]. The use of nominalization, albeit often deemed undesirable [48], has been growing in academic writing [49]. An example of nominalization in our corpus could be the following:

Pancreatic cancer (PC) is characterized by high tumor invasiveness, distant metastasis, and insensitivity to traditional chemotherapeutic drugs...[50].

Phrasal coordination is also positively associated to D3, as it may be associated to a higher degree of descriptivity and more thorough explanation of textual referents, and, similarly to nominalization, displays a similar trend. An example of phrasal coordination in a manuscript with a high score for this feature is:

.. the specific mechanisms are blurry, especially the involved immunological pathways, and the roles of beneficial flora have usually been ignored [50].

3.4. Dimension 4

Dimension 4 is associated to overt expression of persuasion [11], not only referring to the writer's opinion, but also the quality of texts to prompt readers toward a certain course of action. Both our corpora have a negative score (Figure 6A), which indicates that both RAs and reviews from our corpus tend to be non-persuasive, which is in line with the declared function of biomedical literature, as previously stated elsewhere [29]. Unsurprisingly reviews tend to be less negative than RAs in regard to D3 score. This is easily explained by the fact that reviews, by nature, provide readers with an overview of facts and knowledge that can be used to trace recommendations or guidelines. However, the D3 score changed over time, and while RAs have been mostly stable over the years, displaying a slight trend for D3 to increase by about 10% over the course of the last 30 years, reviews

have further decreased this score by the same amount in the last decade (Figure 6B), signaling a movement toward a more impartial stance in review papers. Among the factors that may have affected these changes, the use of **infinitives** has been increasing in both corpora in a similar way (Figure 6C), such as in:

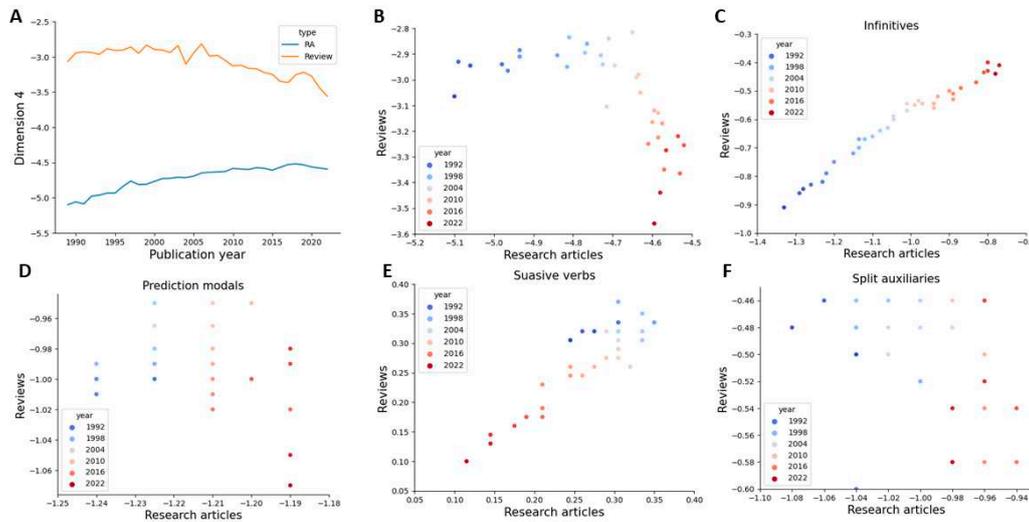


Figure 6. A) Line plot of Dimension 4 (D4) score over the years for the Research article (RA) corpus and the review corpus, in blue and orange respectively; B) scatter plot of D4 score for RAs and reviews. C-F) Scatter plots of the linguistic features of D4 in RA and review corpora by publication years.

Understanding the age-dependent neuromuscular mechanisms underlying force reductions ... allows researchers to investigate new interventions to mitigate these reductions [51].

Suasive verbs are, understandably, another hallmark of overt persuasion, as in:

...an ad hoc committee of the American Venous Forum, working with an international liaison committee, has recommended a number of practical changes [52].

Their frequency, quite similar in both manuscript types, has however been decreasing steadily over the years (Figure 6E), consistently with that more neutral stance we mentioned above. However, **prediction modals**, which have quite a high bearing on this dimension, though displaying quite a high variability in our corpora, have mostly changed for RAs (Figure 6D) and a slight increase can be observed, while the use of **split auxiliaries** has changed for reviews only in the last decade (Figure 6F).

Prediction modals include forms like will, should, or must, which indicate the future directions that research or practice should take, as in:

*The data suggest that treatment of *H. pylori* infection should be considered in children with concomitant GERD* [53].

3.5. Dimension 5

Dimension 5 refers to the abstract or non-abstract nature of the information contained in the texts [11]. As already reported, academic texts, including those from the biomedical area, tend to have high scores for D5, as they tend to contain technical, abstract concepts.

In our corpora, review papers score higher than RAs regardless of the publication date (Figure 7A), although the D5 score decreased for both text types over the year, and the gap between the two groups vanished by the mid second decade of the 2000s (Figure 7A). In the last 5 years D5 score appeared to increase again in reviews only (Figure 7B). The frequent use of **passives** is a hallmark of abstract style, as it typically mitigates the action of an agent (even more so if the passive is **agentless**). These two indices – passives with a “by” agent and agentless passives – have been decreasing in both

text types (Figure 7D, E), presumably driving the trend of the overall D5 score. The use of **conjunctions**, however, has increased both in reviews and RAs, and this increase has been quite sudden in the last 5 years for reviews, which might explain the surge in D5 score in that timeframe.

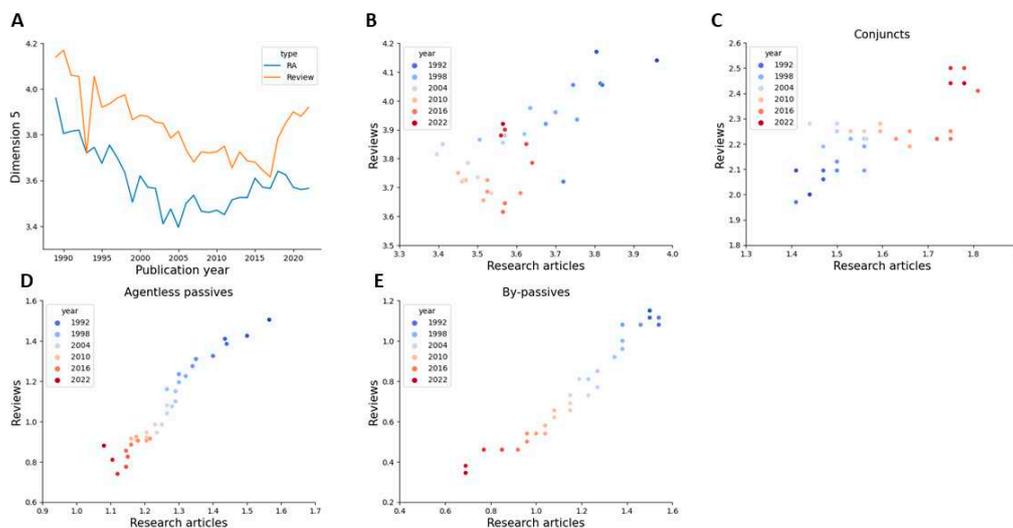


Figure 7. A) Line plot of Dimension 5 (D5) score over the years for the Research article (RA) corpus and the review corpus, in blue and orange respectively; B) scatter plot of D5 score for RAs and reviews. C-E) Scatter plots of the linguistic features of D5 in RA and review corpora by publication years.

Our analysis indicates that, when considering a sample of more than 1.2 million abstracts from the biomedical literature, published in MEDLINE over the last 30 years, we can notice a consolidation of the informational tone of the texts (D1), which occurs in both RAs and reviews. This is combined with a decrease in the use of narrative devices (D2), a change that is most marked in the RA corpus, and a parallel increase in context-independent stances (D3) of both RAs and reviews. The relative lack of overt persuasion (D4) in the academic texts we examined has remained relatively stable over the years, while the degree of abstractness has been decreasing, concomitantly with a decrease in the use of passives. When RAs are compared to reviews, it is apparent that RAs used to rely on narration more heavily than reviews but have toned down the use of this stylistic devices, to a similar level to reviews., while this latter manuscript type used to have a higher degree of content-independency, overt persuasion, and abstractness, which has maintained over the years.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1: Scatter plots of additional D1 features.

Author Contributions: Conceptualization, C.G. and S.G.; methodology, C.G. and M.T.C.; formal analysis, C.G.; resources, S.G. and P.M.; writing—original draft preparation, C.G.; writing—review and editing, P.M.; visualization, C.G.; supervision, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The Data are available on request.

Acknowledgements: The authors would like to thank dr. Silvana Belletti for her advice on corpora.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The following is the list of linguistic features that Biber followed for his Multidimensional analysis, after factor analysis, grouped by dimension and sorted by factor loading, from the highest, modified from [11].

Dimension 1: Involved versus informational production

Positive features (involved production)

Private verbs
that-deletions
 Contractions
 Present tense verbs
do as pro-verb
 Analytic negation
 Demonstrative pronouns
 General emphatics
 First-person pronouns
 Pronoun *it*
 Causative subordination
 Discourse particles
 Indefinite pronouns
 General hedges
 Amplifiers
 Sentence relatives
wh- questions
 Possibility modals
 Nonphrasal coordination
wh- clauses
 Final prepositions

Negative features (informational production)

Nouns
 Word length
 Prepositions
 Type/token ration
 Attributive adjectives

Dimension 2: Narrative versus nonnarrative discourse*Positive features (narrative discourse)*

Past tense verbs
 Third-person pronouns
 Perfect aspect verbs
 Public verbs
 Synthetic negation
 Present participial clauses

Dimension 3: Situation-dependent versus elaborated reference*Positive features (situation-dependent reference)*

Time adverbials
 Place adverbials
 Adverbs

Negative features (elaborated reference)

wh- relative clauses in object positions
 Pied piping constructions
wh- relative clauses in subject positions
 Phrasal coordination
 Nominalizations

Dimension 4: Overt expression of persuasion*Positive features (overt expression of persuasion)*

Infinitives
 Prediction modals
 Suasive verbs
 Conditional subordination

Necessity modals
 Split auxiliaries
 (Possibility modals)

Dimension 5: Nonimpersonal versus impersonal style

Negative features (impersonal style)

Conjuncts
 Agentless passives
 Past participial adverbial clauses
 By passives
 Past participial postnominal clauses
 Other adverbial subordinators

References

- Narin, F.; Pinski, G.; Gee, H.H. Structure of the Biomedical Literature. *Journal of the American society for Information Science* **1976**, *27*, 25–45.
- Cartabellotta, A.; Montalto, G.; Notarbartolo, A. Evidence-Based Medicine. How to Use Biomedical Literature to Solve Clinical Problems. Italian Group on Evidence-Based Medicine-GIMBE. *Minerva Med* **1998**, *89*, 105–115.
- Hrynaszkiewicz, I. The Need and Drive for Open Data in Biomedical Publishing. *Serials* **2011**, *24*.
- Sanberg, P.R.; Gharib, M.; Harker, P.T.; Kaler, E.W.; Marchase, R.B.; Sands, T.D.; Arshadi, N.; Sarkar, S. Changing the Academic Culture: Valuing Patents and Commercialization toward Tenure and Career Advancement. *Proceedings of the National Academy of Sciences* **2014**, *111*, 6542–6547.
- Rice, D.B.; Raffoul, H.; Ioannidis, J.P.A.; Moher, D. Academic Criteria for Promotion and Tenure in Biomedical Sciences Faculties: Cross Sectional Analysis of International Sample of Universities. *Bmj* **2020**, *369*.
- Landhuis, E. Scientific Literature: Information Overload. *Nature* **2016**, *535*, 457–458.
- Sharma, H.; Verma, S. Predatory Journals: The Rise of Worthless Biomedical Science. *J Postgrad Med* **2018**, *64*, 226.
- Ioannidis, J.P.A. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q* **2016**, *94*, 485–514.
- Pieper, D.; Antoine, S.-L.; Mathes, T.; Neugebauer, E.A.M.; Eikermann, M. Systematic Review Finds Overlapping Reviews Were Not Mentioned in Every Other Overview. *J Clin Epidemiol* **2014**, *67*, 368–375.
- Biber, D. On the Complexity of Discourse Complexity: A Multidimensional Analysis. *Discourse Process* **1992**, *15*, 133–163.
- Biber, D. *Variation across Speech and Writing*; Cambridge University Press, 1991; ISBN 0521425565.
- Stig, J.; Leech, G.N.; Goodluck, H. Manual of Information to Accompany the Lancaster-Oslo: Bergen Corpus of British English, for Use with Digital Computers. (*No Title*) **1978**.
- Pöldvere, N.; Johansson, V.; Paradis, C. On the London–Lund Corpus 2: Design, Challenges and Innovations. *English Language & Linguistics* **2021**, *25*, 459–483.
- Biber, D.; Conrad, S.; Reppen, R.; Byrd, P.; Helt, M. Speaking and Writing in the University: A Multidimensional Comparison. *TESOL quarterly* **2002**, *36*, 9–48.
- Friginal, E.; Mustafa, S.S. A Comparison of US-Based and Iraqi English Research Article Abstracts Using Corpora. *J Engl Acad Purp* **2017**, *25*, 45–57.
- Cao, Y.; Xiao, R. A Multi-Dimensional Contrastive Study of English Abstracts by Native and Non-Native Writers. *Corpora* **2013**, *8*, 209–234.
- Nini, A. The Multi-Dimensional Analysis Tagger. *Multi-dimensional analysis: Research methods and current issues* **2019**, 67–94.
- Tausczik, Y.R.; Pennebaker, J.W. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *J Lang Soc Psychol* **2010**, *29*, 24–54.
- Guizzardi, S.; Colangelo, M.T.; Mirandola, P.; Galli, C. The Evolution of Narrativity in Abstracts of the Biomedical Literature between 1989 and 2022. *Publications* **2023**, *11*, 26, doi:10.3390/publications11020026.
- Guizzardi, S.; Colangelo, M.T.; Mirandola, P.; Galli, C. The Evolution of Narrativity in Abstracts of the Biomedical Literature between 1989 and 2022. *Publications* **2023**, *11*, 26.

21. Shapiro, A. Littler-Getter.
22. Greenhalgh, T.; Thorne, S.; Malterud, K. Time to Challenge the Spurious Hierarchy of Systematic over Narrative Reviews? *Eur J Clin Invest* **2018**, *48*, e12931, doi:10.1111/eci.12931.
23. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the Proceedings of the 9th Python in Science Conference; van der Walt, S., Millman, J., Eds.; 2010; pp. 51–56.
24. Richardson, L. Beautiful Soup Documentation.
25. Curtis, A.; Smith, T.; Ziganshin, B.; Elefteriades, J. The Mystery of the Z-Score. *AORTA* **2016**, *04*, 124–130, doi:10.12945/j.aorta.2016.16.014.
26. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **2007**, *9*, doi:10.1109/MCSE.2007.55.
27. Waskom, M. Seaborn: Statistical Data Visualization. *J Open Source Softw* **2021**, *6*, doi:10.21105/joss.03021.
28. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S.; et al. Jupyter Notebooks—a Publishing Format for Reproducible Computational Workflows. In Proceedings of the Positioning and Power in Academic Publishing: Players, Agents and Agendas - Proceedings of the 20th International Conference on Electronic Publishing, ELPUB 2016; IOS Press BV, 2016; pp. 87–90.
29. Liu, J.; Xiao, L. A Multi-Dimensional Analysis of Conclusions in Research Articles: Variation across Disciplines. *English for Specific Purposes* **2022**, *67*, 46–61, doi:10.1016/j.esp.2022.04.003.
30. Andrews, R. My Cheerful Attitude Upset an Anxious Pre-Op Patient. *Nursing Standard* **2009**, *24*, 27–28.
31. Hyland, K. Authority and Invisibility. *J Pragmat* **2002**, *34*, 1091–1112, doi:10.1016/S0378-2166(02)00035-8.
32. Hyland, K. Options of Identity in Academic Writing. *ELT Journal* **2002**, *56*, 351–358, doi:10.1093/elt/56.4.351.
33. Hyland, K.; Jiang, F. (Kevin) Is Academic Writing Becoming More Informal? *English for Specific Purposes* **2017**, *45*, 40–51, doi:10.1016/j.esp.2016.09.001.
34. Pathak, S.K.; Salunke, A.A.; Chawla, J.S.; Sharma, A.; Ratna, H.V.K.; Gautam, R.K. Bilateral Radial Head Fracture Secondary to Weighted Push-Up Exercise: Case Report and Review of Literature of a Rare Injury. *Indian J Orthop* **2021**, 1–6.
35. Elseviers, M.M.; Van Camp, Y.; Nayaert, S.; Duré, K.; Annemans, L.; Tanghe, A.; Vermeersch, S. Prevalence and Management of Antibiotic Associated Diarrhea in General Hospitals. *BMC Infect Dis* **2015**, *15*, 1–9.
36. Carrió Pastor, M. Cross-Cultural Variation in the Use of Modal Verbs in Academic English. *SKY Journal of Linguistics* **2014**, *27*, 153–166.
37. Nakamura, H. Experimental and Clinical Aspects of Oxidative Stress and Redox Regulation. *Rinsho Byori* **2003**, *51*, 109–114.
38. Harris, E.E. *Hypothesis and Perception: The Roots of Scientific Method*; Routledge, 2014; ISBN 1317851609.
39. Ueberberg, B.; Tourne, H.; Redman, A.; Walz, M.K.; Schmid, K.W.; Mann, K.; Petersenn, S. Differential Expression of the Human Somatostatin Receptor Subtypes Sst1 to Sst5 in Various Adrenal Tumors and Normal Adrenal Gland. *Hormone and Metabolic Research* **2005**, *37*, 722–728.
40. Pretolani, M.; Lellouch-Tubiana, A.; Lefort, J.; Bachelet, M.; Vargaftig, B.B. PAF-Acether and Experimental Anaphylaxis as a Model for Asthma. *Int Arch Allergy Immunol* **1989**, *88*, 149–153.
41. Wilson, D.A.; Sullivan, R.M. Neurobiology of Associative Learning in the Neonate: Early Olfactory Learning. *Behav Neural Biol* **1994**, *61*, 1–18.
42. Boyd, R.L.; Blackburn, K.G.; Pennebaker, J.W. The Narrative Arc: Revealing Core Narrative Structures through Text Analysis. *Sci Adv* **2020**, *6*, doi:10.1126/sciadv.aba2196.
43. Freytag, G. *Freytag's Technique of the Drama*; Scott, Foresman, 1894;
44. Corver, N.; van Riemsdijk, H. *Semi-Lexical Categories: The Function of Content Words and the Content of Function Words*; Walter de Gruyter, 2013; Vol. 59; ISBN 3110874008.
45. Alexiadou, A. Nominalizations: A Probe into the Architecture of Grammar Part I: The Nominalization Puzzle. *Lang Linguist Compass* **2010**, *4*, 496–511, doi:10.1111/j.1749-818X.2010.00209.x.
46. Khamesian, M. On Nominalization, A Rhetorical Device in Academic Writing. *Armenian Folia Anglistika* **2015**, *11*, 42–48, doi:10.46991/AFA/2015.11.2.042.
47. Baratta, A.M. Nominalization Development across an Undergraduate Academic Degree Program. *J Pragmat* **2010**, *42*, 1017–1036, doi:10.1016/j.pragma.2009.08.007.
48. Biber, D.; Gray, B. Challenging Stereotypes about Academic Writing: Complexity, Elaboration, Explicitness. *J Engl Acad Purp* **2010**, *9*, 2–20, doi:10.1016/j.jeap.2010.01.001.
49. Biber, D.; Gray, B. Nominalizing the Verb Phrase in Academic Science Writing. In *The Register-Functional Approach to Grammatical Complexity*; Routledge, 2021; pp. 176–198.

50. Wei, X.; Mei, C.; Li, X.; Xie, Y. The Unique Microbiome and Immunity in Pancreatic Cancer. *Pancreas* **2021**, *50*, 119–129.
51. Orssatto, L.B. da R.; Wiest, M.J.; Diefenthaler, F. Neural and Musculotendinous Mechanisms Underpinning Age-Related Force Reductions. *Mech Ageing Dev* **2018**, *175*, 17–23, doi:10.1016/j.mad.2018.06.005.
52. Eklöf, B.; Rutherford, R.B.; Bergan, J.J.; Carpentier, P.H.; Gloviczki, P.; Kistner, R.L.; Meissner, M.H.; Moneta, G.L.; Myers, K.; Padberg, F.T.; et al. Revision of the CEAP Classification for Chronic Venous Disorders: Consensus Statement. *J Vasc Surg* **2004**, *40*, 1248–1252, doi:10.1016/j.jvs.2004.09.027.
53. Pollet, S.; Gottrand, F.; Vincent, P.; Kalach, N.; Michaud, L.; Guimber, D.; Turck, D. Gastroesophageal Reflux Disease and Helicobacter Pylori Infection in Neurologically Impaired Children: Inter-Relations and Therapeutic Implications. *J Pediatr Gastroenterol Nutr* **2004**, *38*.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.