

JOURNAL OF COMMUNICATION, LANGUAGE AND CULTURE

Small Words, Big Patterns: Cross-Journal, Stylistic, and Diachronic Perspectives on Function Word Use in Astrophysical Research Abstracts

David Israel Méndez Alcaraz^{1*}, María Ángeles Alcaraz Ariza²

¹Department of Physics, System Engineering and Signal Theory, Polytechnic University College, University of Alicante, Alicante, Spain

²Department of English Studies, Faculty of Arts, University of Alicante, Alicante, Spain

Corresponding author: david.mendez@ua.es; ORCID iD: <https://orcid.org/0000-0002-7754-559X>

ABSTRACT

In this study, we examine the use of function words in a specialised corpus of 220 research paper abstracts published between 2011 and 2021 in four leading astrophysical journals written in English. Although often overlooked in linguistic analyses, function words play a crucial role in ensuring textual cohesion, clarity, and rhetorical structure in scientific writing. However, they pose significant challenges for non-native English speakers who publish in international journals. Despite extensive research on academic writing, the use of function words in specialised scientific disciplines such as astrophysics remains unexplored. Our pilot analysis focuses on four categories of function words—prepositions, determiners, conjunctions, and pronouns—using quantitative corpus-based frequency analyses and cross-journal comparisons. The results show that prepositions are the most frequently used function words, followed by determiners, conjunctions, and pronouns. The differences observed among journals could be attributed to differences in authorial backgrounds (e.g. authors who acquired English as a first language versus those who learned it as a second or third language) and the divergent stylistic conventions between British and American authors. Some variations may also be linked to the journal's scope, with more generalist and experimental outlets exhibiting distinct linguistic tendencies. These findings contribute to ongoing corpus-based research on disciplinary, stylistic, and diachronic patterns in scientific communication and may inform academic writing pedagogy across diverse scholarly and professional domains.

Keywords: function words, research paper abstracts, astrophysics, corpus-based analysis, cross-journal comparison, diachronic variation

Received: 16 October 2025, **Accepted:** 8 February 2026, **Published:** 30 July 2026

1.0 Introduction

As scientific research continues to expand rapidly, research papers (RPs) remain the primary medium for disseminating new findings. Within these papers, abstracts play a crucial role in summarising key aspects of a study and enabling editors, reviewers, and readers to assess its relevance quickly (Huckin, 2001; Swales, 2004). With the growing volume of publications and the rise of interdisciplinary research, abstracts serve as effective filters for navigating the research landscape (Hartley & Benjamin, 1998).

Consequently, scholarly attention to research paper abstracts (RPAs) has increased across disciplines, with studies focusing on their rhetorical structure, stylistic features, and the challenges encountered by non-native English-speaking researchers seeking to meet international publishing standards (Hyland, 2000; Swales, 1990).

Content and function words are core linguistic features of abstracts and have been widely examined in academic discourse. This distinction, first proposed by Fries (1952), is based on their respective roles in conveying meaning and structuring a text. Subsequent linguistic theories have emphasised that sentence comprehension relies on both the lexical meanings of words and the grammatical framework that organises them (Crystal & Robins, 2025). Content words—nouns, main verbs, adjectives, and adverbs—carry semantic weight and belong to an open class that evolves over time (Harnad, 2023). In contrast, function words constitute a closed class (Aarts, 2001); although limited in number and lacking concrete meaning (Nordquist, 2024), they encode relationships between content words, enhance textual cohesion, and are essential for comprehension (Beare, 2025; Pennebaker, 2011).

Nevertheless, a watertight distinction between content and function words does not always hold in English from either a semantic or syntactic perspective, as some words with identical graphic forms display properties of both categories (e.g., *that*). For instance, “behind” may function as an adverb (content word) or as a preposition (function word), depending on its syntactic position, while “many” can operate as a noun or adjective (content word) or as a pronoun (function word). Therefore, accurate classification requires consideration of both semantic and syntactic criteria.

Moreover, the content/function distinction primarily applies to analytic languages, such as Chinese, English, French, German, and Spanish, which rely on function words and word order to encode semantic relationships. In contrast, agglutinative languages, including Finnish, Hungarian, Japanese, Korean, and Turkish, largely express grammatical and semantic relations through affixation. Accordingly, Corver and van Riemsdijk (2013) argue that content and function words are better understood as forming a continuum rather than two discrete categories. This view is echoed in metaphorical accounts: paraphrasing Biber *et al.* (1999), Tang (2020) describes lexical words as the “bricks” of a text and function words as the “mortar” that binds them, a characterisation similarly articulated by Chung and Pennebaker (2007), who refer to function words as the “cement” holding content words together.

Given their central role in linguistic structure and meaning, function words have been investigated across a wide range of disciplines, including computational linguistics, child language acquisition, language disorders, neurolinguistics, psycholinguistics, psychology, sociolinguistics, stylometry, and language transfer. Their importance is further reflected in the abundance of resources devoted to their use, from grammar books and educational websites to creative writing blogs and online forums, as well as in digital tools such as grammar checkers, writing assistants, and large language models (LLMs), which increasingly incorporate function word analysis into academic, technical, and formal writing.

Despite broad scholarly interest, function words remain understudied in astrophysics RPAs, representing an underexplored intersection between linguistics and discipline-specific research. Astrophysics, a field that combines astronomy and physics, has rarely been examined from a linguistic perspective, partly because scientists and linguists have traditionally operated in separate communities (Kragh, 2023; Tarone *et al.*, 1981, 1998). Addressing this gap, the present study investigates the distribution and communicative roles of function words in astrophysics RPAs, aiming to contribute to a deeper understanding of disciplinary writing practices and support more effective scientific communication.

2.0 Literature Review

Prepositions are among the most extensively studied function words because of the challenges they pose for non-native speakers, often resulting from first-language transfer, substitution, or omission errors (Budiharto *et al.*, 2022; Du & Du, 2024; Kang *et al.*, 2024; Rocha, 2022). Across these studies, L2/L3 learners (i.e., non-native English speakers) exhibit systematic difficulties that reflect both cross-linguistic influence and developmental patterns in second-language acquisition. Pedagogical interventions, particularly those combining explicit instruction with input enhancement, have been shown to improve learner accuracy, thereby highlighting the importance of targeted teaching strategies.

Determiners, especially articles, also present challenges for EFL learners, as they are absent or realised differently in many languages. Studies have indicated that learners whose first language lacks articles, such as Chinese and Russian, tend to omit or misuse them more frequently than learners whose L1 includes an article system (Antes, 2019; Derkach & Alexopoulou, 2024; Snape, 2019). Research further suggests that determiners require contextual interpretation and implicit retrieval from discourse, which adds to the complexity of learning (Divjak *et al.*, 2023). These findings underscore the interaction between linguistic structure, cross-linguistic influence, and exposure, and have clear implications for teaching English academic writing.

Conjunctions also present difficulties for EFL learners, particularly in academic writing where they contribute to textual cohesion and coherence. Studies have shown that learners frequently overuse, underuse, or misuse conjunctions, often relying on a limited range of familiar connectors while struggling to express more complex logical relationships (Alawerdy & Alalwi, 2022; Sanosi, 2024). These difficulties have been linked to both first-language transfer and limited exposure to academic discourse. Research further suggests that explicit instruction can improve learners' use of conjunctions and enhance the overall quality of their writing.

Pronouns likewise constitute a challenging category of function words for EFL learners. Difficulties commonly involve pronoun reference, antecedent agreement, and the selection of appropriate personal or relative pronouns, often reflecting cross-linguistic influence from the learner's first language (Li & Akram, 2024). Because accurate pronoun use requires the integration of grammatical, semantic, and discourse-level information, learners may continue to experience difficulties even at advanced proficiency levels. These findings further highlight the role of linguistic structure and cross-linguistic influence in the acquisition of English academic writing.

3.0 Methodology

The purpose of the present investigation is to expand upon our previous studies on different linguistic features of RPAs in astrophysics (Méndez & Alcaraz, 2020; 2025a; 2025b; 2025c), with a focus on function words. The objectives of the present study are:

- a. To quantify the frequency and distribution of function words in astrophysical research abstracts.
- b. To identify the most common variants within each function word category.
- c. To compare usage patterns across four leading astrophysical journals.
- d. To examine diachronic changes in function word usage between 2011 and 2021.
- e. To interpret variations in function word usage in relation to authorial backgrounds and journal characteristics.

To this end, we employ two analytical approaches: (1) an extrinsic approach that examines the overall frequencies and percentages of function words and their variants, and (2) an intrinsic approach that calculates proportional relationships between function word categories and other linguistic units, thereby enabling standardised cross-journal and diachronic comparisons.

Our corpus consists of 220 RPAs published between 2011 and 2021 in four leading astrophysical journals: *Astronomy & Astrophysics* (A&A), *Astronomy Journal* (AJ), *The Astrophysical Journal* (ApJ), and *Monthly Notices of the Royal Astronomical Society* (MNRAS).

A&A, originally a European-based journal distributed on behalf of the European Southern Observatory, currently belongs to China Science Publishing & Media and publishes research in all areas, such as theoretical, observational, and instrumental astronomy and astrophysics, including exoplanets. AJ emphasises scientific results derived from observations and describes data capture, surveys, analysis techniques, astronomical interpretation, instrumentation, software, and computing topics. The ApJ is devoted to recent developments, discoveries, and theories in astronomy and astrophysics. Publications in ApJ constitute research directly relevant to astrophysical applications, whether based on observational results, theoretical insights, or modelling. Both AJ and ApJ are US-based and owned by the American Astronomical Society (AAS). MNRAS, which is often preferred by astronomers from the United Kingdom and the Commonwealth, covers original research in positional and dynamical astronomy, astrophysics, radio astronomy, cosmology, space research, and astronomical instrument design.

The specific questions guiding our research on function words are the following:

- a. What is the total number of function words and their variants relative to the total number of words?
- b. What are the most common function words?
- c. What is the proportion of prepositions relative to the total number of words, nouns, and pronouns?
- d. What is the proportion of determiners relative to the total number of words and nouns?
- e. What is the proportion of conjunctions relative to the total number of words and main verbs?
- f. What is the proportion of pronouns relative to the total number of words?

Our research procedure is summarised as follows. First, we collected our samples and performed word counting. Because of the considerable variation in writing, we chose to perform a manual word count, as training a Natural Language Processing (NLP) tool would have been time-consuming and labour-intensive. In other words, our study can be characterised as a pilot experiment that provides the groundwork for future applications in NLP training and large-scale corpus analysis. We therefore defined the concept of “word” as any unit occurring between spaces, including blanks, dashes, or hyphens. Initialisms (abbreviations formed from the initial letters of constituent words and pronounced separately) were recorded according to their individual semantic components, whereas acronyms, which are strings of letters with a syllabic structure, were considered single words. The different variants are illustrated below:

- a. “High-column-density environment” (A&A, 12/2011) = 4 words
- b. “Double-lined eclipsing binaries” (A&A, 02/2012) = 4 words
- c. “Gaia EDR3” [$<$ Gaia (Global Astrometric Interferometer for Astrophysics) Early Data Release 3] (AJ, 12/2021) = 5 words
“near-infrared-resolved H 2 emission” (ApJ, 03/2011) = 6 words

To ensure accuracy, both authors independently performed manual counting, and the results were compared before any discussion. Our counting methodology yielded 57,972 words, distributed as follows: A&A (17,076 words), ApJ (14,341 words), MNRAS (13,517 words), and AJ (13,038 words). In the third stage, we examined the function words themselves and calculated several extrinsic variables: the total number of function words and their different variants, namely prepositions, determiners, conjunctions, pronouns, and helping verbs.

Before presenting the results of the analysis, it is necessary to clarify the criteria used to identify function words in this study. Only items serving a purely grammatical function were included. Homographic forms functioning as lexical or content words were excluded. Disambiguation was conducted manually by the authors on a case-by-case basis. The following examples illustrate forms that were excluded because they function lexically in context:

1. “The active region has (main verb) a size $<1 R_{\odot}$, while the pulses produced by the ECMI mechanism are (main verb) from a much more compact region, for example $\sim 0.007 R_{\odot}$.” (A&A, 01/2011)
2. “B+HRC-S and XMM-Newton RGS spectra of H-like C and N lines formed in the corona of the primary star of the RS CVn-type binary λ And, a mildly metal-poor G8 III–IV first-ascent giant that completed dredge-up (noun) ~ 50 Myr ago, have been used to make a precise measurement of its surface C/N ratio.” (AJ, 11/2011)
3. “Using CanariCam at the 10.4 m Gran Telescopio Canarias, we have mapped out (particle) the mid-infrared polarization of the protoplanetary disk around the Herbig Ae star AB Aur.” (ApJ, 11/2016)
4. “We compute the luminosity and volume distributions of the superstructures, finding that about 10 per cent of the luminosity mass will end up (phrasal verb) in future virialized structures.” (MNRAS, 07/2011)

In the fourth step, we computed a series of intrinsic variables: the mean number of prepositions per number of words, nouns, and pronouns; the mean number of determiners per number of words and nouns; the mean number of conjunctions per number of words and main verbs; and the mean number of pronouns per number of words. Since this study serves as a pilot, exploratory analysis of the four journals involved, and considering the unequal variances and small sample sizes, we opted to apply Welch’s two-sample t-tests for all pairwise journal tests, six in total. These tests were exploratory rather than confirmatory, and the results should be viewed as hypothesis-generating. Consequently, p-values are presented to suggest potential differences, but they should not be interpreted as definitive evidence of these differences. An alpha value of 0.05 was used.

In summary, a mixed-method approach was employed, combining quantitative corpus analysis and qualitative examination to uncover frequency patterns and contextual functions. All the referred variables were examined from both a cross-journal and a diachronic perspective to identify any potential differences over time and across the four journals.

4.0 Results and Discussion

4.1 Extrinsic Approach

Table 1

Breakdown of Function Words

Journal	Words	Function Words	Prepositions	Determiners	Conjunctions	Pronouns	Helping Verbs
A&A	17,076	5,796	2,396	1,969	627	420	384
ApJ	14,341	4,568	1,843	1,566	551	328	280
MNRAS	13,517	4,373	1,731	1,499	545	351	247
AJ	13,038	4,193	1,667	1,374	560	320	272
TOTAL	57,972	18,930	7,637	6,408	2,283	1,419	1,183

Table 1 displays a cross-journal breakdown of function words in our sample. The most striking observation is the predominance of prepositions, which account for 40.34% of all function words. They are followed by determiners (33.85%), conjunctions (12.06%), pronouns (7.50%), and helping verbs, which appear in the last position (6.25%). Helping verb frequencies are reported for completeness only; the helping verb category was addressed in detail in Méndez & Alcaraz (2025b) and is therefore not analysed in the present study. The dominance of prepositions is expected, given their wide range of grammatical and semantic functions, including establishing relationships between nouns by indicating category, content, material, position, quality, size, type, value, manner, place, or time. In addition to describing the means, modes, or instruments used, they also indicate purpose, link ideas, introduce the different stages of research, and provide conclusions.

Determiners play a vital role in communication by specifying which nouns are being referred to, thereby making sentences more precise and focused, an essential quality in written scientific discourse, where objectivity, precision, and formality are paramount. The prototypical determiners are the definite article “the” and the indefinite articles “a” and “an.” “A” and “an” refer to non-specific or not previously mentioned nouns, while “the” refers to specific or previously mentioned nouns. In other words, the choice of articles depends on the knowledge shared by the writer and the reader.

All these features make both prepositions and determiners ideal linguistic resources for use in abstracts, which are summaries of longer articles that, as recommended in most journal rules, cannot exceed a given number of words, usually around 200–250 words. This implies that short and precise independent sentences expressing a complete thought, with verbs conjugated in simple tenses (present and past) in the active voice (Méndez & Alcaraz, 2025b), are preferred over more complex and dependent structures. The combination of these linguistic strategies reflects a direct and assertive tone that helps highlight the research carried out and the achievements obtained to convince editors to publish the paper and readers to read the entire content.

Table 2 shows the five most common variants for each function-word category, with helping verbs excluded from the analysis.

Table 2

Most Common Function Words

Function words	Most Common Variants	Percentage (%)
Prepositions	of	29.67
	in	13.30
	to	12.41
	with	8.27
	for	7.25
Determiners	the	65.16
	a/an	18.74
	this	3.20
	our	2.38
	these	2.06
Conjunctions	and	60.00
	that	15.73
	or	5.17
	as	4.40
	but	2.67
Pronouns	we	53.05
	that	13.31
	which	11.10
	it	5.83
	this	2.91

4.1.1 Prepositions

As can be seen, “of”, “in”, and “to” are the most used prepositions. The predominance of the preposition “of” over the remaining prepositions had already been pointed out by Mindt and Weber (1989) in their analysis of the distribution of prepositions in the *Brown University Corpus of American English* and the *LOB (Lancaster-Oslo-Bergen) Corpus of British English*. Moreover, our results partially coincide with Yuan (2014), who found that “to” ranks first, “of” second and “in” third in his study on the transfer of Chinese language characteristics to English writings by Chinese students. Similarly, Nghi and Phuc (2022) discovered that “of”, “in” and “to” are the three most used prepositions in their comparison between CAEC (*Cambridge Academic English Corpus*) and a self-built corpus of third-year English major academic writing from Ho Chi Minh City University of Food Industry during the 2019-2020 academic year. Finally, Li (2023), analyzing the Corpus of Contemporary American English (COCA), a sub-corpus of the *Chinese Learner English Corpus (CLEC)*, and a self-compiled corpus of 12th-grade EFL learners’ writing samples, also found “of”, “in”, and “to” as the three most frequent English prepositions, supporting the results of the present study. The fourth and fifth positions were occupied by “for” and “with”, with “with” preceding “for” in our sample. This ordering may indicate a stronger emphasis on methodological description and relational specification (e.g. “with respect to”, “with the use of”) over purpose-oriented constructions typically associated with “for”.

From a cross-journal perspective, the percentage patterns of preposition use generally mirror the overall trend across journals, except in A&A, where “to” overcomes “in”. This difference is likely due to the higher frequency of infinitives in this journal (Méndez & Alcaraz, 2025b).

4.1.2 Determiners

Regarding determiners, Table 2 shows that articles are the most frequent type (Parisse & Maillart, 2019). This further indicates that the use of the definite article “the”, the most frequently occurring individual word in the English language (*British Council Learn English*, 2025), prevails by far over the four remaining words included in the list. The predominance of “the”, both in written and spoken English, was previously observed by Leech *et al.* (2001) in their analysis of the British National Corpus (BNC) and by Master (2013) in his comparison between research papers published in the domain of science and technology and *Teaching English to Speakers of Other Languages (TESOL)* courses. Nevertheless, Master (2013) added that the “zero” determiner, which is not detectable by word-counting programs, is actually the most frequent free morpheme in English.

Shifting focus to cross-journal patterns, the highest frequency of use of the determiner “the” appears in A&A (66.88%) and the lowest in AJ (57.05%), while ApJ (65.26%) and MNRAS (62.68%) fall in between them. In contrast, the determiners “a/an” appear most often in AJ (20.73%), followed by ApJ (19.32%), A&A (18.59%), and MNRAS (17.73%). This suggests that indefinite determiners are used more frequently in US-based journals.

In the same line as the definite article “the”, the demonstrative determiners singular (“this”) and plural (“these”) refer to specific things that have been just mentioned, that is, they cannot be understood without context. They can also be used to discuss things and ideas in the present. From a cross-journal perspective, the highest percentages of both determiners corresponded to AJ (4.77% for “this” and 2.71% for “these”) and the lowest ones to A&A (1.71% for “this” and 1.11% for “these”).

When looking at demonstratives (“this”, “these”, “that”, and “those”), they make up 5.69% of all determiners. Their highest overall percentage appeared in AJ (7.69%), while A&A (3.02%), ApJ (7.05%), and MNRAS (5.91%) occupied the middle positions. These figures indicate that demonstrative determiners such as indefinite articles are more commonly employed in US-based journals.

To conclude the section on determiners, the first-person plural possessive “our” is used before nouns to indicate ownership of the noun. Across journals, its highest frequency of appearance is found in MNRAS (6.57%), followed by AJ (6.41%), A&A (5.58%), and ApJ (1.51%). Considering all possessives together (“our”, “their”, “its”, “whose”, and “his”), they account for 5.52% of all determiners overall. Possessive usage is highest in MNRAS (6.57%) and lowest in ApJ (3.65%), with AJ (6.41%) and A&A (5.58%) in between. In other words, researchers publishing in MNRAS tend to

use possessives more frequently, along with authors in the more experimental AJ, than in the more generalist ApJ.

4.1.3 Conjunctions

According to Table 2, the conjunction “and” is by far the most frequently used among all conjunctions, accounting for 60% of the total. This predominance has been well documented in previous research (Hertwig *et al.*, 2008; Leech and Svartvik, 2013; Leung, 2005; Quirk and Greenbaum (2016 [1973])). Following “and” in frequency are the conjunctions “that”, “or”, “as”, and “but”, each appearing at significantly lower percentages. This sharp drop indicates a strong preference for additive coordination over other types of logical relationships in academic writing, at least in this corpus. Interestingly, our findings differ slightly from those of previous studies. While “and” consistently ranks first across the literature, the studies cited above typically identify “or” and “but”, rather than “that” and “as” as the next most common conjunctions. This variation may reflect differences in corpus types, disciplinary focus, or stylistic conventions across journals.

Some discrepancies also emerge when comparing our findings with those of Purba *et al.* (2017), who analyzed the types of conjunctions used in the “Results” and “Discussion” sections of a series of research articles written by graduate students at the Faculty of Education, University of Bengkulu (Indonesia), as well as in two articles published in the *Teaching of English as a Foreign Language in Indonesia (TEFLIN)* journal. According to their study, the most frequent type of conjunctions in student-written research articles are coordinating conjunctions (56.11%), with “and” alone comprising 42.98% of the total instances. The *TEFLIN* journal article, however, shows a marked preference for subordinating conjunctions (61.74%), with “that” being the most frequently used (36.66%). This contrast underscores the potential differences in writing style, academic maturity, and publication standards, with student writers leaning towards additive coordination and professional publications favoring more complex subordination to express nuanced relationships.

The predominance of coordinating conjunctions, particularly “and”, was also noticed by Rahayati *et al.* (2021), who observed an overuse of “and” in abstracts of international journals focused on EFL teaching. Likewise, Riadi *et al.* (2023) found that “and” prevails in research articles written by undergraduate students from higher educational institutions in West Borneo (Indonesia) as part of their bachelor’s degree requirements, these findings further support the idea that “and” serves as a default connector in various academic contexts, especially among less experienced writers or in specific disciplinary and regional conventions.

From a cross-journal perspective, the usage patterns of coordinating and subordinating conjunctions reveal notable variations. The coordinating conjunction “and” appears most frequently in A&A (65.04%), followed by ApJ (59.57%), MNRAS (57.58%), and AJ (57.19%), suggesting a stronger preference for additive coordination in A&A. In contrast, the analysis of the subordinating conjunction “that” shows that ApJ records the highest percentage (19.50%), followed closely by MNRAS (19.49%), AJ (13.81%), and finally A&A (10.87%). The relatively lower use of “that” in A&A can be attributed to a higher frequency of asyndetic structures, which frequently replace clauses introduced by “that” and conjugated reporting verbs (Méndez & Alcaraz, 2025b), thereby lowering the frequency of subordinating conjunctions.

Turning to the coordinating conjunction “or”, A&A once again ranks highest in terms of usage (6.30%), closely followed by AJ (6.29%). MNRAS displays a slightly lower percentage (5.05%), and ApJ has an even smaller proportion (2.86%). Focusing on the subordinating conjunction “as”, AJ holds the highest percentage (4.90%), followed closely by MNRAS (4.87%) and ApJ (4.83%), with A&A ranking last (3.15%). This lower figure in A&A may be attributed to the linguistic backgrounds of its contributors, who are predominantly L2/L3 users. This observation aligns with the findings of Pangaribuan *et al.* (2018), who suggested that difficulties in using conjunctions among L2/L3 learners are often due to interlingual differences and the numerous meanings of “as”.

Regarding the coordinating conjunction “but”, A&A again records the highest frequency at 4.25%. The other journals show significantly lower percentages: AJ (2.62%), MNRAS (1.99%), and ApJ (1.61%).

The higher occurrence of “but” in A&A could be interpreted in terms of its overuse by L2/L3 writers, who form most of the contributors to this journal. In contrast, L1 authors (more prevalent in the other journals) tend to employ a wider and more nuanced range of connectors. In this context, our results agree with Yakonthova (2020), who highlighted the stylistic challenges L2/L3 researchers face when attempting to publish in English-language indexed journals.

In summary, the higher frequency of coordinating conjunctions in A&A RPAs, which typically serve to link clauses or phrases rather than entire sentences, correlates with a lower presence of subordinating conjunctions. This pattern reflects a preference for simpler syntactic structures in A&A, as opposed to the more complex sentence constructions found in the other journals. Moreover, the shorter average sentence length in A&A RPAs is probably influenced by the journal’s required five-part structure (context, aims, methods, results, and conclusions), which encourages concise, self-contained statements. This format naturally favors coordination over subordination, reinforcing the tendency towards brevity and syntactic simplicity in A&A.

Taken together, these patterns suggest that even subtle function word usage can influence textual cohesion, clarity, and rhetorical style, with implications for both authors navigating publication and teaching disciplinary writing practices.

4.1.4 Pronouns

The first-person plural pronoun “we” dominates, accounting for over half of all pronouns, highlighting collaborative engagement and authorial presence (Hyland, 2003; Haggan, 2004; Harwood, 2005; Weinberger *et al.*, 2015; Hyland & Jing, 2018). Following “we”, the most used pronouns are “that” and “which”, although their frequency is substantially lower (approximately four to five times less). The third-person singular personal pronoun “it” ranks fourth, occurring roughly nine times less frequently than “we”. The demonstrative singular pronoun “this” appears the least often, accounting for fewer than 3% of all pronoun instances.

From a cross-journal perspective, the highest occurrence of “we” is found in MNRAS (55.18%), followed by ApJ (53.75%), A&A (52.24%), and AJ (51.07%). These results may be due to two factors: 1) all but two RPAs, one published in ApJ by a Chinese author who uses the first-person singular pronoun “I” and another published in A&A by a Spanish author who employs the first-person plural pronoun “we” (a case of “Royal we” or “pluralis majestatis”, probably to convey a sense of collective representation), are multi-authored, which clearly reflects the collaborative nature of contemporary scientific research; 2) most of the verbs following “we” are conjugated in the active voice (Méndez & Alcaraz, 2025b), which not only enhances author visibility within the scientific community but also signals a higher degree of authorial engagement in the research process.

Concerning the relative pronoun “that”, the highest percentage is found in ApJ (15.02%), followed by A&A (13.88%), AJ (13.15%), and MNRAS (11.20%). The discrepancy in the use of “that” between the ApJ and MNRAS (two journals with similarly generalist scopes) could be interpreted from two perspectives: (1) differing stylistic preferences among contributing researchers and (2) distinct geographic contexts, with the ApJ being a US-based journal and MNRAS a British one.

Regarding the relative pronoun “which”, MNRAS records the highest frequency (11.76%), followed by AJ (11.31%), ApJ (10.81%), and A&A (10.59%). The first position reached in MNRAS compared to A&A could be tentatively interpreted in two ways: 1) authors publishing in MNRAS are more likely to be L1 speakers, whereas A&A tends to attract contributions from L2/L3 users; 2) L2/L3 writers may find it more challenging to construct complex syntactic structures involving non-restrictive clauses typically introduced by “which”.

The pronoun “it” is more common in US-based journals, with ApJ (7.81%) and AJ (6.42%) leading, followed by A&A (4.94%) and MNRAS (4.48%). This suggests a greater presence of impersonal constructions in US-based scientific writing, possibly reflecting stylistic conventions that favour objectivity and detachment. By contrast, “this” is more frequently used in A&A (3.53%) and MNRAS (3.36%), while appearing less frequently in ApJ (2.70%) and AJ (1.83%). Overall, the data suggest an

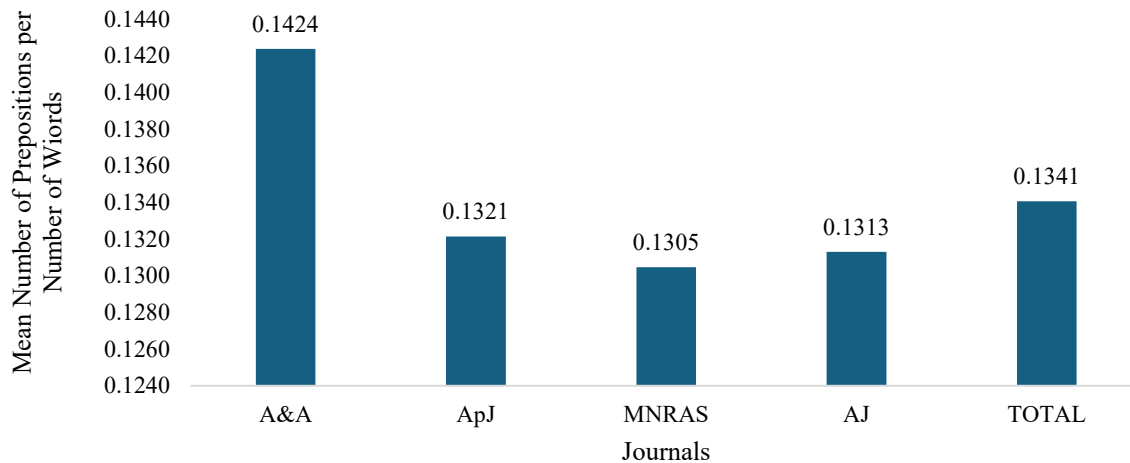
inverse usage pattern: “it” dominates in US-based journals, whereas “this” is more common in British and European publications, pointing to stylistic and linguistic influences that vary across regions.

4.2 Intrinsic Approach

4.2.1 Prepositions

Figure 1

Mean Number of Prepositions per Number of Words



Turning to intrinsic variables, Figure 1 shows a cross-journal comparison of the mean number of prepositions per number of words.

As shown, A&A registers the highest value for this parameter, whereas MNRAS displays the lowest. Statistically significant differences are found between A&A and the other three journals ($p=0.0058$ for ApJ, $p=0.0120$ for AJ, and $p=0.0019$ for MNRAS). This result reinforces the conclusions of a previous study (Méndez & Alcaraz, 2025a), which reported lower semantic informativity and complexity in A&A RPAs. The higher frequency of prepositions in A&A could reflect, in part, a two-fold scenario: 1) a geographic context whereby the majority of published authors are L2/L3 writers who may not yet possess full mastery of academic English compared to L1 authors (Ruan, 2018; Yakonthova, 2020; Xue & Ge, 2021); and 2) as a consequence, these authors may transfer grammatical structures from their native language into English (Mahmoodzadeh, 2012; Loke *et al.*, 2013; Lorincz & Gordon, 2012; Alonso, 2017; Kirmizi & Karci, 2017; Sanjaya & Bram, 2021; Mahardika & Bram, 2022, among others who examined the challenges faced by L2/L3 learners in acquiring English prepositions).

From a diachronic point of view, the pattern appears clearly erratic, although there is a slight overall decrease from 2011 to 2021. The highest mean number of prepositions per number of words is recorded in 2018 (0.1424), while the lowest occurs in 2020 (0.1258), with a statistically significant difference between both years ($p=0.0183$).

Figure 2

Mean Number of Prepositions per Number of Nouns and Pronouns

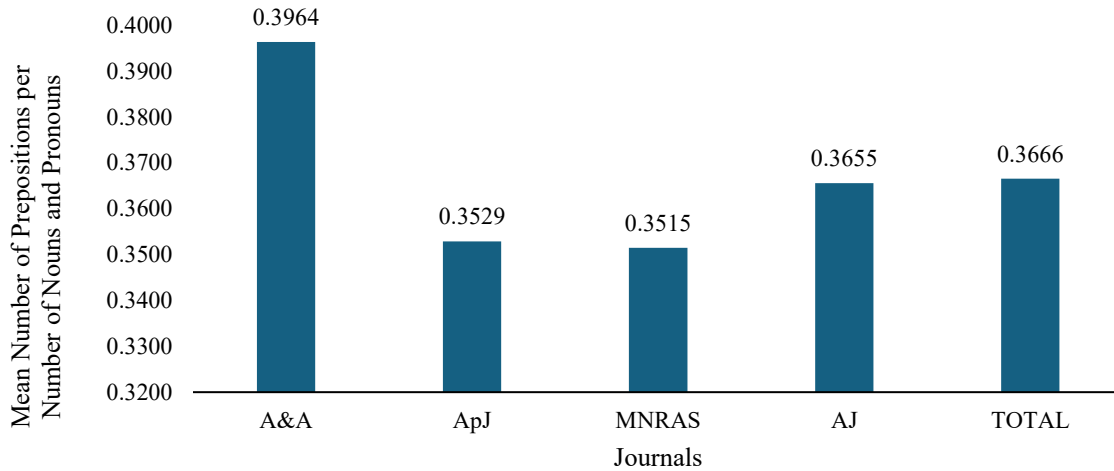


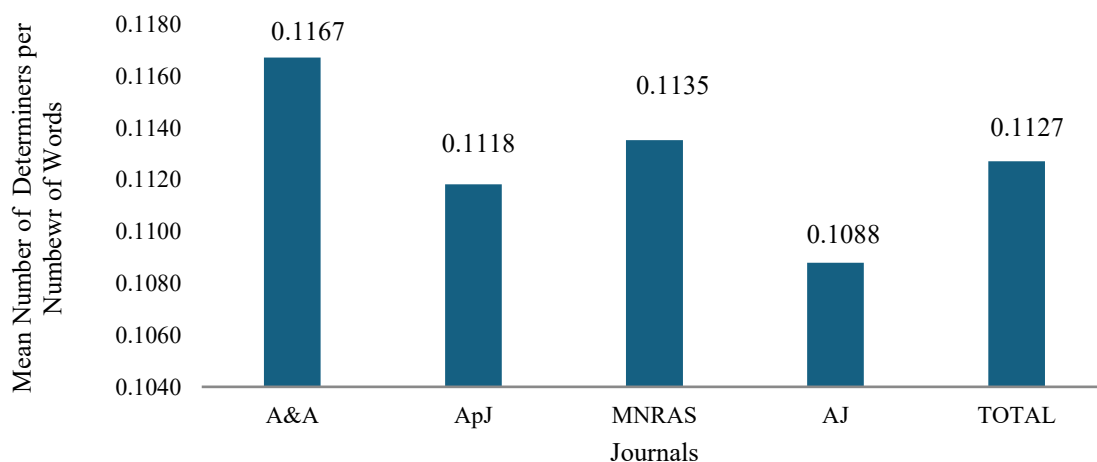
Figure 2 displays a cross-journal study of the mean number of prepositions per number of nouns and pronouns. Consistent with previous findings, A&A ranks highest on this scale, presenting statistically significant differences compared to ApJ ($p=0.0013$) and AJ ($p=0.0347$). Meanwhile, MNRAS again shows the lowest value among the journals. Notably, ApJ and AJ switch positions relative to their rankings based on the mean number of prepositions per number of words. This reversal can be attributed to the mean number of nouns per number of words found in ApJ compared to AJ (Méndez & Alcaraz, 2025a). In addition, the pattern illustrated in Figure 2 is precisely the inverse of that observed for the mean number of compound nouns per number of words (Méndez & Alcaraz, 2025a), thereby again suggesting the dichotomy between L1 and L2/L3 authorship.

From a diachronic perspective, the trend appears irregular, with a slight decrease overall. The highest value corresponds to 2013 (0.3995) and the lowest to 2020 (0.3403), with a statistically significant difference ($p=0.0226$).

4.2.2 Determiners

Figure 3

Mean Number of Determiners per Number of Words



With respect to the mean number of determiners per number of words, Figure 3 shows that the highest value occurs in A&A, while the lowest is found in AJ, with a statistically significant difference between them ($p=0.0254$). It is worth noting, however, that AJ also has the highest percentage of demonstratives and the second-highest percentage of possessives, just after MNRAS.

From a diachronic perspective, there is an irregular pattern with a slight overall increase. The maximum value is observed in 2018 (0.1262) and the minimum in 2016 (0.1033), with a statistically significant difference between both ($p=0.0087$).

Figure 4

Mean Number of Determiners per Number of Nouns

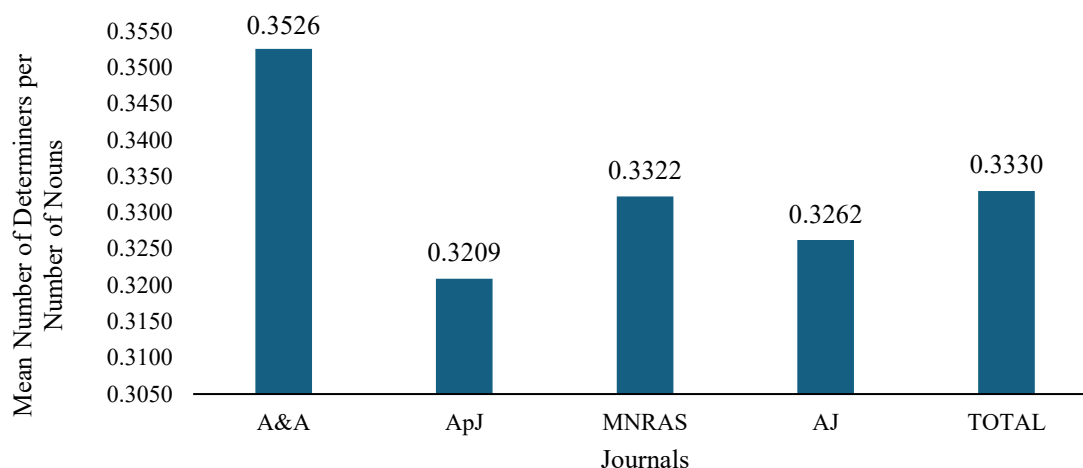


Figure 4 shows the mean number of determiners per number of nouns in the three groups. As observed, A&A once again exhibits the highest mean number of determiners per number of nouns, with MNRAS ranking second. Statistically significant differences are found between A&A and ApJ ($p = 0.0245$) and between A&A and AJ ($p = 0.0464$). Interestingly, the positions of AJ and ApJ are reversed in Figure 4 compared to those in Figure 3. This shift is not surprising, given that ApJ shows the highest mean number of nouns per number of words, whereas AJ has the second lowest, just above A&A (Méndez & Alcaraz, 2025a). Furthermore, the two US-based journals display the lowest mean number of determiners both per number of words and per number of nouns. This pattern likely reflects stylistic differences, as previously noted by Biber *et al.* (1999) in their comparison of the COCA and BNC corpora. The higher use of determiners per number of words and nouns found in A&A may also be influenced by linguistic transfer from other languages into English (Liu & Gleason, 2002), a phenomenon similarly observed with prepositions (see Figures 1 and 2) and determiners (see Figure 3). According to Kojima and Popiel (2023), L2/L3 authors often face linguistic challenges when choosing the correct articles, prepositions, and determiners in English.

Diachronically, the data exhibits an erratic pattern with a slight increase overall. The minimum value occurs in 2016 (0.3012) and the maximum in 2018 (0.3765), with a statistically significant difference between both ($p= 0.0228$).

4.2.3 Conjunctions

Figure 5

Mean Number of Conjunctions per Number of Words

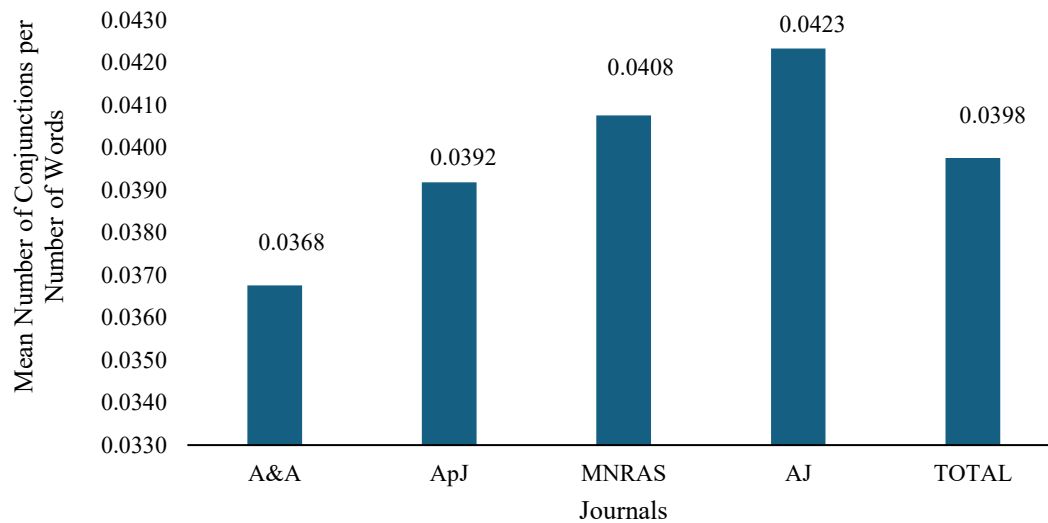


Figure 5 shows the mean number of conjunctions per number of words. A&A exhibits the lowest mean number of conjunctions per number of words, whereas the highest value is observed in AJ, with a statistically significant difference between the two ($p=0.0155$). From a diachronic perspective, the pattern is erratic, with a slight overall increase throughout the period studied. The minimum value is recorded in 2018 (0.0337) and the maximum in 2019 (0.0484), with the difference between them being statistically significant ($p=0.0009$).

Figure 6

Mean Number of Conjunctions per Number of Main Verbs

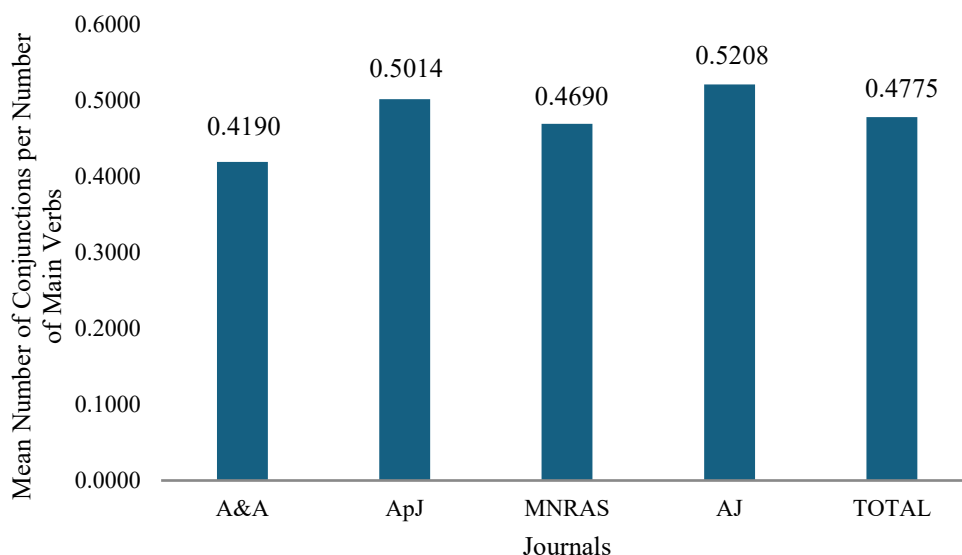


Figure 6 displays the mean number of conjunctions per number of main verbs. A&A once again shows the lowest mean number of conjunctions per number of main verbs, while AJ exhibits the highest, with the difference between both being statistically significant ($p=0.0028$). ApJ displays the second-highest mean, and the difference between ApJ and A&A is also statistically significant ($p=0.0178$). Notably, the

values for ApJ and MNRAS in Figure 6 are reversed relative to those in Figure 5. This result can be attributed to ApJ having the lowest mean number of main verbs per number of words, and MNRAS having the second highest (Méndez & Alcaraz, 2025a).

From a diachronic standpoint, the mean number of conjunctions per main verb shows a slightly irregular downward trend, in contrast to the overall increase observed in the mean number of conjunctions per number of words. The minimum value is recorded in 2018 (0.3803) and the maximum in 2019 (0.5953), with the difference between the two being statistically significant ($p=0.0015$).

Taken together, Figures 5 and 6 suggest that authors publishing in A&A (predominantly L2/L3, see previous sections) tend to use fewer conjunctions both per number of words and per number of main verbs. This pattern aligns with findings from a substantial body of research (Granger, 1998; Hinkel, 2003; Chen, 2006; Neary-Sundquist, 2013, to name a few), which reported lower conjunction usage among L2/L3 writers. A second explanation may lie in the syntactic structure of A&A RPAs, most of which consist of simple sentences without any conjunctions. Finally, the persistently higher values observed in AJ may be attributed to its more experimentally oriented scope, which typically requires more complex sentence construction to describe methodologies and present results.

4.2.4 Pronouns

Figure 7

Mean Number of Pronouns per Number of Words

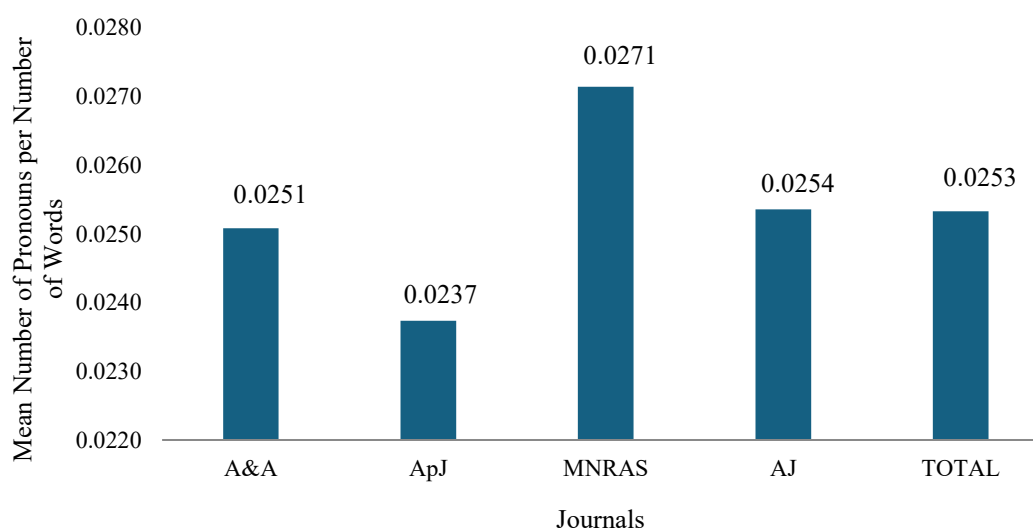


Figure 7 shows the mean number of pronouns per number of words. MNRAS presents the highest mean number of pronouns per number of words, and ApJ the lowest, with the difference between both being nearly statistically significant (0.0531). Considering that the pronoun “we” accounts for more than half of all pronouns used (see Table 2), it is not surprising that MNRAS, the journal with the second-highest percentage of active voice constructions (Méndez & Alcaraz, 2025a), tops the scale. The second-highest value in this ranking is observed in AJ, which has the highest percentage of active voice usage (Méndez & Alcaraz, 2025a).

From a diachronic perspective, the pattern is erratic, with a slight overall increase over the study period (from 0.0222 in 2011 to 0.0297 in 2021, $p=0.0180$). The maximum value occurs in 2021 (0.0297), and the minimum is recorded in 2013 (0.0212), with a statistically significant difference between both ($p=0.0073$). This overall increase in pronoun usage parallels the documented decrease in the use of passive voice constructions during the same period (Méndez & Alcaraz, 2025a).

This exploratory study provides explanations that illuminate the observed patterns, acknowledging that journal-specific conventions, editorial practices, or other contextual factors may also play a role while offering a coherent interpretation of the results.

5.0 Conclusion

In this study, we conduct a detailed analysis of various types of function words, excluding helping verbs, in a sample of RPAs collected from the four leading astrophysical journals written in English during the period 2011–2021.

Our main findings indicate that prepositions are the most frequently used function words in astrophysical RPAs, followed by determiners and conjunctions. Conjunctions occupy third place and pronouns fourth, with the latter two appearing at much lower frequencies. Among individual items, “of” is the most used preposition, “the” is the most frequent determiner, “and” ranks highest among conjunctions, and “we” is the most common pronoun.

The US-based journals exhibit a higher frequency of indefinite articles and demonstratives, whereas the British journal MNRAS shows a higher proportion of possessive determiners. Additionally, A&A, the only journal in the sample with structured RPAs, shows the highest percentage of coordinating conjunctions, primarily used to join phrases or clauses, and the lowest percentage of subordinating conjunctions introducing dependent clauses.

Focusing on intrinsic variables, A&A exhibits the highest mean values of prepositions per number of words, as well as per number of nouns and pronouns, in contrast to MNRAS, which consistently shows the lowest values for these metrics. A&A also ranks highest in the mean number of determiners per number of words and per number of nouns and pronouns, whereas both US-based journals display the lowest figures for these parameters. A reverse pattern is observed in the use of conjunctions: the mean number of conjunctions per number of words and per number of main verbs is highest in AJ and lowest in A&A. As for pronouns, MNRAS displays the highest value, closely followed by AJ, which also reports the highest proportion of active voice structures. ApJ, the other US-based journal, ranks lowest in this category.

From a cross-journal perspective, most of the differences identified may be interpreted through the lens of an L1 versus L2/L3 authorial dichotomy. In this regard, it is important to note that the use of specific function words often correlates with the use of verbal constructions and vice versa (Méndez & Alcaraz, 2025a). Additional variation may stem from the stylistic preferences typical of British versus US academic writing. Finally, some differences could also be attributed to the journals’ scopes, ranging from generalist to more experimentally oriented publications.

From a diachronic perspective, the variables analysed display highly erratic patterns, with no substantial statistically significant trends over time. Overall, only slight variations are observed: (1) a decrease in the mean number of prepositions per number of words and per number of nouns and pronouns, accompanied by an increase in the mean number of determiners in both metrics; (2) a rise in the mean number of conjunctions per number of words, contrasted by a decline in the mean number of conjunctions per number of main verbs; and (3) a general increase in the mean number of pronouns per number of words.

Collectively, these findings indicate how subtle function word choices shape cohesion, clarity, and rhetorical style in astrophysical writing, revealing both discipline- and journal-specific conventions. By highlighting these patterns, this study offers practical guidance for non-native authors seeking to produce more coherent and rhetorically effective scientific texts and provides a foundation for future corpus-based research investigating how linguistic features influence scientific communication and writing pedagogy across disciplines.

This corpus-based empirical study has several limitations. The observed pairwise differences require cautious interpretation because the analysis is exploratory and not intended to test a single omnibus hypothesis. Furthermore, the relatively small sample size and restricted temporal coverage of 11 years limit the breadth of the inferences that may be drawn. Future research would benefit from larger datasets

spanning longer periods, as well as from the use of variance-adjusted omnibus tests (e.g., Welch ANOVA) followed by planned post hoc comparisons, which may enhance the reliability and generalisability of the findings. In addition, the methodology employed can be extended to other disciplinary domains to facilitate cross-field comparisons.

Acknowledgements

We would like to thank the anonymous reviewers for their insightful and constructive feedback, which greatly contributed to the improvement of this manuscript. Their thoughtful suggestions and recommendations were invaluable in refining our analysis and interpretation.

Conflict of Interest

The authors declare that no conflicts of interest exist.

Authors' Contribution Statement

DIMA: Conceptualisation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing. MAAA: Conceptualisation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing.

Funding

This research received no external funding.

Ethics Statement

This research did not require IRB approval because it used publicly available secondary data comprising published research paper abstracts, with no involvement of human participants and no collection of identifying personal information.

Data Access Statement:

Research data supporting this publication are available upon request from the corresponding author.

Authors Biography

David I. Méndez holds a PhD in astrophysics from La Laguna University (Canary Islands, Spain) and is a lecturer in the Department of Physics, Engineering Systems and Signal Theory at the Polytechnic University College of the University of Alicante, Spain. He teaches undergraduate and master's courses in fundamental engineering physics, optics, and acoustics in both English and Spanish. He has published numerous research articles on astrophysics, optics, non-linear oscillations, information science, and English linguistic matters in leading international journals.

María Ángeles Alcaraz Ariza is retired and was formerly a lecturer in the Department of English Studies at the University of Alicante, Spain. She is the author of numerous research articles, conference papers, and other publications on the linguistic and pragmatic-rhetorical analysis of English-, French-, and Spanish-written specialised discourses, particularly in medicine and astrophysics.

References

- Aarts, B. (2001). *English syntax and argumentation*. New York: Palgrave
- Alawerdy, A. S., & Alalwi, F. S. (2022). Enhancing English as a foreign language university students' writing through explicit instruction of conjunctions as cohesive devices: An experimental study. *Frontiers in Psychology*, 13, 1053310. <https://doi.org/10.3389/fpsyg.2022.1053310>
- Alonso, R. A. (2017). L1 influence on second language acquisition and teaching. *New Trends and Issues Proceedings on Humanities and Social Sciences*, 2(9), 136–149. <https://doi.org/10.18844/gjhss.v2i9.1094>
- Antes, T. A. (2019). Countability versus flexibility: Rethinking vocabulary and determiner instruction. *TESL-EJ*, 24(2), 1–18. <https://www.tesl-ej.org/pdf/ej90/a3.pdf>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Amsterdam: John Benjamins Publishing Company.
- British Council. (2025, March 1). *Definite article*. British Council LearnEnglish. <https://learnenglish.britishcouncil.org/grammar/english-grammar-reference/definite-article>
- Budiharto, R. A., Azizah, L. T. M., & Yazunka, Z. W. (2022). Investigating mother tongue effect in the acquisition of English prepositions by Indonesian-speaking learners. *ELTR Journal*, 7(1), 34–47. <https://doi.org/10.37147/eltr.v7i1.163>
- Chen, C. W. (2006). The use of connective adverbials in the academic papers of advanced Chinese ESL learners. *International Journal of Corpus Linguistics*, 11(1), 113–130. <https://doi.org/10.1075/ijcl.11.1.05che>
- Chung, C. K., & Pennebaker, J. W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343–359). Psychology Press
- Corver, N., & van Riemsdijk, H. (2013). *Semi-lexical categories: The function of content words and the content of function words*. De Gruyter Mouton. <https://doi.org/10.1515/9783110874006>
- Crystal, D., & Robins, R. H. (2025, April 11). *Language*. Encyclopaedia Britannica. <https://www.britannica.com/topic/language>
- Derkach, K., & Alexopoulou, T. (2024). Definite and indefinite article accuracy in learner English: A multifactorial analysis. *Studies in Second Language Acquisition*, 46(3), 710–740. <https://doi.org/10.1017/S0272263123000463>
- Divjak, D., Romain, L., & Milin, P. (2023). From their point of view: The article category as a hierarchically structured referent tracking system. *Linguistics*, 61(4), 1027–1068. <https://doi.org/10.1515/ling-2022-0186>
- Du, H., & Du, S. (2024). The acquisition and application of prepositions by Chinese middle school students: Taking *in*, *on*, and *at* as examples. *SHS Web of Conferences*, 187, Article 01026. <https://doi.org/10.1051/shsconf/202418701026>
- Fries, C. C. (1952). *The structure of English: An introduction to the construction of English sentences*. Harcourt, Brace
- Granger, S. (1998). *Learner English on computer*. Routledge. <https://doi.org/10.4324/9781315841342>
- Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attraction. *Journal of Pragmatics*, 36(2), 293–317. [https://doi.org/10.1016/S0378-2166\(03\)00090-0](https://doi.org/10.1016/S0378-2166(03)00090-0)
- Harnad, S. (2023, September 16). *Content words, function words, and LLM analysis of dictionaries*. Skywritings. <https://generic.wordpress.soton.ac.uk/skywritings/2023/09/16/content-words-function-words-and-llm-analysis-of-dictionaries/>

- Hartley, J., & Benjamin, M. (1998). Use of abstracts when reading journal articles: The first stages of information retrieval. *Journal of the American Society for Information Science*, 49(7), 640–645
- Harwood, N. (2005). “Nowhere has anyone attempted ... In this article I aim to do just that”: A corpus-based study of self-promotional *I* and *we* in academic writing across four disciplines. *Journal of Pragmatics*, 37(8), 1207–1231. <https://doi.org/10.1016/j.pragma.2005.01.012>
- Hertwig, R., Benz, B., & Krauss, S. (2008). The conjunction fallacy and the many meanings of “and.” *Cognition*, 108(3), 740–753. <https://doi.org/10.1016/j.cognition.2008.06.008>
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37(2), 275–301. <https://doi.org/10.2307/3588505>
- Huckin, T. (2001). Abstracting from abstracts. In M. Hewings (Ed.), *Academic writing in context: Implications and applications* (pp. 93–103). University of Birmingham Press
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. University of Michigan Press ELT. <https://doi.org/10.3998/mpub.6719>
- Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of the American Society for Information Science and Technology*, 54(3), 251–259. <https://doi.org/10.1002/asi.10204>
- Hyland, K., & Jiang, F. K. (2018). In this paper we suggest: Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, 51, 18–30. <https://doi.org/10.1016/j.esp.2018.02.001>
- Kang, S., Shin, Y. K., & Yoo, I. W. (2024). Using lexical bundles to teach prepositions to Korean EFL students: Corpus-based instructed SLA. *Journal of Second Language Studies*, 7(1), 75–98. <https://doi.org/10.1075/jsls.00022.kan>
- Kirmizi, O., & Karci, B. (2017). An investigation of Turkish higher education EFL learners’ linguistic and lexical errors. *Educational Process: International Journal*, 6(4), 35–54. <https://doi.org/10.22521/edupij.2017.64.3>
- Kojima, T., & Popiel, H. A. (2023). Correct use of articles and prepositions in academic writing: Advice for non-native English-speaking researchers. *Journal of Korean Medical Science*, 38(48), Article e417. <https://doi.org/10.3346/jkms.2023.38.e417>
- Kragh, H. (2023). A new literary style of science: The rise of acronyms in physics and astronomy. *Physics in Perspective*, 25, 175–198. <https://doi.org/10.1007/s00016-024-00306-9>
- Leech, G., & Svartvik, J. (2013). *A communicative grammar of English*. London: Routledge. <https://doi.org/10.4324/9781315836041>
- Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. Routledge. <https://doi.org/10.4324/9781315840161>
- Leung, C. (2005). *A comparison of the use of major English conjunctions by American and Hong Kong university students* [bachelor’s thesis, Lund University]. Lund University Publications. <https://lup.lub.lu.se/luur/download?func=downloadFile&recordId=1331671&fileId=1331672>
- Li, Q. (2023). A corpus-based study on the use of high-frequency prepositions between Chinese English learners and native speakers. *Curriculum and Teaching Methodology*, 6, 104–113. <http://dx.doi.org/10.23977/curtm.2023.060616>
- Li, S., & Akram, H. (2024). Navigating pronoun-antecedent challenges: A study of ESL academic writing errors. *SAGE Open*, 14(4). <https://doi.org/10.1177/21582440241296607>

- Liu, D., & Gleason, J. L. (2002). Acquisition of the article “the” by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition*, 24, 1–26. <https://doi.org/10.1017/S0272263102001018>
- Loke, D. L., Ali, J., & Anthony, N. N. Z. (2013). A corpus-based study on the use of preposition of time “on” and “at” in argumentative essays of Form 4 and Form 5 Malaysian students. *English Language Teaching*, 6(9), 128–135. <https://doi.org/10.5539/elt.v6n9p128>
- Lorincz, K., & Gordon, R. (2012). Difficulties in learning prepositions and possible solutions. *Linguistic Portfolios*, 1, Article 14. https://repository.stcloudstate.edu/stcloud_ling/vol1/iss1/14
- Mahardika, A. A. A., & Bram, B. (2022). Preposition mistakes in English for Specific Purposes: The case of Indonesian accounting study program students. *Pedagogy: Journal of English Language Teaching*, 10(2), 97–110. <https://doi.org/10.32332/joelt.v10i2.4560>
- Mahmoodzadeh, M. (2012). A cross-linguistic study of prepositions in Persian and English: The effect of transfer. *Theory and Practice in Language Studies*, 2(4), 734–740. <https://doi.org/10.4304/tpls.2.4.734-740>
- Master, P. (2013). A contrastive study of determiner usage in EST research articles. *International Journal of Language Studies*, 7(1), 33–58. <https://ijls.net/sample/71-2.pdf>
- Méndez, D. I., & Alcaraz, M. A. (2020). Research paper abstracts in Monthly Notices of the Royal Astronomical Society (1943–2018). *English Text Construction* 13(1), 62–93. <https://doi.org/10.1075/etc.00035.alc>
- Méndez, D. I., & Alcaraz, M. A. (2025a). Content Words and Compound Structures in English-Written Astrophysical Research Paper Abstracts (2011–2021). *Atlantis. Journal of the Spanish Association for Anglo-American Studies* 47(2), 25–46. <https://doi.org/10.28914/Atlantis-2025-47.2.2>
- Méndez, D. I., & Alcaraz, M. A. (2025b). The Use of Verbs in Astrophysical Research Paper Abstracts: Sociolinguistic Implications. *Journal of Linguistics, Culture and Communication* 3(1), 88–119. <https://doi.org/10.61320/jolcc.v3i1.88-119>
- Méndez, D. I., & Alcaraz, M. A. (2025c). The Role of Adverbs in Astrophysical Research Paper Abstracts. *Indian Journal of Language and Linguistics* 6(2), 17–31. <https://doi.org/10.54392/ijll2523>
- Mindt, D., & Weber, C. (1989). Prepositions in American and British English. *World Englishes*, 8(2), 229–238. <https://doi.org/10.1111/j.1467-971X.1989.tb00658.x>
- Neary-Sundquist, C. A. (2013). The development of cohesion in a learner corpus. *Studies in Second Language Learning and Teaching*, 3(1), 109–130. <https://doi.org/10.14746/ssllt.2013.3.1.6>
- Nghi, T. T., & Phuc, T. H. (2022). A corpus-based study on the most frequently used English prepositions: A conceptual transfer perspective. *Theory and Practice in Language Studies*, 12(2), 213–220. <https://doi.org/10.17507/tpls.1202.01>
- Nordquist, R. (2024, July 20). *Definition and examples of function words in English*. ThoughtCo. <https://www.thoughtco.com/function-word-grammar-1690876>
- Pangaribuan, T., Haddina, E., & Manik, S. (2018). The students’ error in using conjunctions (*because*, *since*, *as*, *in case*) in sentences. *English Language Teaching*, 11(4), 91–100. <https://doi.org/10.5539/elt.v11n4p91>
- Parisse, C., & Maillard, C. (2019). Determiners. In *The SAGE encyclopedia of human communication sciences and disorders* (pp. 578–579). SAGE Publications. <https://doi.org/10.4135/9781483380810.n188>

- Pennebaker, J. W. (2011). Your use of pronouns reveals your personality. *Harvard Business Review*, 89(12), 32–33. <https://pubmed.ncbi.nlm.nih.gov/22250353/>
- Purba, D. R., Safnil, & Pulungan, R. (2017). A comparative study of types of conjunctions used in the results and discussion section of students' RAs and TEFLIN JAs published in 2015. *Journal of English Education and Teaching*, 1(1), 65–73. <https://doi.org/10.33369/jeet.1.1.65-73>
- Quirk, R., & Greenbaum, S. (1973). *A university grammar of English*. London: Longman.
- Rahayati, N., & Herlina, R. (2021). Discourse markers in abstracts of international journals. *Journal of English Education Program (JEEP)*, 8(2). [http://dx.doi.org/10.25157/\(jeep\).v8i2.6426](http://dx.doi.org/10.25157/(jeep).v8i2.6426)
- Riadi, A., Nathania, N., & Barona Valentino, Y. (2023). Types and appropriateness: Students' usage of conjunctions in academic writing. *Linguists: Journal of Linguistics and Language Teaching*, 9(2), 165. <https://doi.org/10.29300/ling.v9i2.2590>
- Rocha, J. V. P. (2022). Meanings of the preposition “of” in the oral interlanguage of A2 and B2 Brazilian learners. *Revista X*, 17(3), 1031–1053. <https://doi.org/10.5380/rvx.v17i3.85093>
- Ruan, Z. (2018). Structural compression in academic writing: An English Chinese comparison study of complex noun phrases in research article abstracts. *Journal of English for Academic Purposes*, 36, 37–47. <https://doi.org/10.1016/j.jeap.2018.09.001>
- Sanjaya, A. A., & Bram, B. (2021). Investigating preposition usage problems of English language education study program students. *Saga*, 2(1), 19–33. <https://doi.org/10.21460/saga.2020.21.65>
- Sanosi, A. B. (2024). Revisiting cohesion in academic writing: A corpus-based analysis of EFL learners' use of conjunctions. *Indonesian Journal of Applied Linguistics*, 14(1), 64–78. <https://doi.org/10.17509/ijal.v14i1.70418>
- Snape, N. (2019). The acquisition of articles: The story so far. *Second Language*, 18, 7–24. https://doi.org/10.11431/secondlanguage.18.0_7
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press
- Swales, J. M. (2004). *Research genres: Exploration and application*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524827>
- Tang, W. (2020). The development of a functional function word list from form and meaning. *Bulletin of University Education Center (Fukuyama University)*, 6, 51–61
- Tarone, E., Dwyer, S., Gillette, S., & Icke, V. (1981). On the use of the passive in two astrophysics journal papers. *The ESP Journal*, 1(2), 123–140. [https://doi.org/10.1016/0272-2380\(81\)90004-4](https://doi.org/10.1016/0272-2380(81)90004-4)
- Tarone, E., Dwyer, S., Gillette, S., & Icke, V. (1998). On the use of the passive and active voice in astrophysics journal papers: With extension to other languages and other fields. *English for Specific Purposes*, 17(1), 113–132. [https://doi.org/10.1016/S0889-4906\(97\)00032-X](https://doi.org/10.1016/S0889-4906(97)00032-X)
- Weinberger, C. J., Evans, J. A., & Allesina, S. (2015). Ten simple (empirical) rules for writing science. *PLOS Computational Biology*, 11(4), Article e1004205. <https://doi.org/10.1371/journal.pcbi.1004205>
- Xue, Q., & Ge, T. (2021). A corpus-based study on phrasal complexity in computer science abstracts of novice and advanced writers. *Open Journal of Modern Linguistics*, 11(5), 808–822. <https://doi.org/10.4236/ojml.2021.115062>
- Yakonothova, T. (2020). English writing of non-Anglophone researchers. *Journal of Korean Medical Science*, 35(26), Article e216. <https://doi.org/10.3346/jkms.2020.35.e216>

- Yuan, H. C. (2014). A corpus-based study on the influence of L1 on EFL learners' use of prepositions. *Theory and Practice in Language Studies*, 4(12), 2513–2521. <https://doi.org/10.4304/tpls.4.12.2513-2521>