ISRG Journal of Multidisciplinary Studies (ISRGJMS)

ISRG PUBLISHERS
Abbreviated Key Title: isrg j. multidiscip. Stud.
**ISSN: 2584-0452 (Online)**
Journal homepage: https://isrgpublishers.com/isrgjms/
Volume – IV, Issue - IV (April) 2026
Frequency: Monthly


# Application of Small Corpora in Aeronautical English Teaching

**Zeqi LIU**

University of Malaya, Kuala Lumpur, Malaysia

| **Received:** 11.04.2026 | **Accepted:** 16.04.2026 | **Published:** 19.04.2026

***Corresponding author:** Zeqi LIU

## Abstract

*Since the 1990s, corpus linguistics has been increasingly integrated into language pedagogy. However, research regarding corpus applications within English for Specific Purposes (ESP) remains relatively sparse, primarily due to the inherent complexity of ESP disciplines and the challenges of acquiring specialized corpora. This paper utilizes Aviation English as a case study to explore practical pathways for implementing small corpora in ESP instruction.Initially, the paper establishes that small ESP corpora must adhere to principles of representativeness, authenticity, and operability. By aligning with the professional requirements of Aviation English learners, a specialized corpus was constructed using authentic civil aviation data across various genres. The application of this corpus is examined through three dimensions: identifying linguistic features, providing direct classroom materials, and supporting autonomous learning. This methodology aims to correct common ESP teaching pitfalls, such as knowledge-oriented biases and over-reliance on instructor experience.A case study on verb nominalization (comparing malfunction and its nominalized form malfunctioning) illustrates the linguistic distinctions between Aviation English and General English, confirming the utility of small corpora in feature analysis. The findings suggest that self-built small corpora effectively heighten students' sensitivity to ESP linguistic nuances and bolster professional communication skills. Nevertheless, the study acknowledges limitations in representativeness and suggests that small corpora be used in tandem with large-scale corpora for enhanced precision.*

**Keywords:** *Small Corpus, English for Specific Purposes (ESP), Aviation English, Language Teaching, Verb Nominalization*

## INTRODUCTION

Corpus linguistics has remained a cornerstone of linguistic research since the 1990s, with language teaching serving as a primary application field. Meanwhile, aviation English, as a highly standardized "aviation linguistic language," involves not only linguistic rules but also information interpretation in complex professional environments (Borowska, 2018). To improve teaching quality, linguists should collaborate with aviation experts (SMEs) to develop methods that compensate for the lack of specialized background knowledge in purely language-based instruction (Borowska & Simon, 2023). Furthermore, research topics for


**DOI: 10.5281/zenodo.19648910**

applying corpora to language teaching are concentrated in: corpus informed methods (developing appropriate teaching materials based on word frequency, collocation, and structure, as well as information on usage context and register), corpus-based language description (using corpora to verify researchers' hypotheses), and corpus-driven data analysis (forming hypotheses or even conclusions about language use through observation).

The role of corpora in higher education language teaching is also

significant. Corpora are widely used in areas such as textbook compilation and language testing. The establishment of learner corpora also allows teachers to understand students' English usage more clearly. However, research on the use of corpora in English for Specific Purposes (ESP) teaching is uncommon. This is because ESP involves numerous disciplines. Teachers often cannot find corresponding corpora for specific disciplines, making application to teaching out of the question. In reality, however, if teachers understand the principles and methods of corpus construction, they can build small corpora themselves to be applied to various aspects of ESP teaching.

## CONSTRUCTION OF CORPUS

**Principles of Construction**

The establishment of an ESP corpus must adhere to both the basic principles of corpus construction and the special nature of ESP, possessing the following characteristics:

1) Representativeness: The ESP corpus should be established based on a needs analysis of the students. Data should be selected according to the needs of students to communicate effectively in their future professional contexts, and the selected data should cover all aspects of teaching content to the greatest extent possible.
2) Authenticity: One of the basic principles of ESP teaching is to use authentic materials as textbooks. Exercise design and both in-class and extracurricular activities should reflect the socio-cultural context of ESP. Therefore, an ESP corpus should consist of authentic data from specific thematic fields (such as Aviation English). Furthermore, in the aviation context, the authenticity of the corpus is directly related to flight safety; therefore, it must cover standard communication as well as language use in abnormal situations (Borowska, 2025).
3) Operability: The intended users of an ESP corpus are ordinary teachers and students. Therefore, the corpus should possess strong operability to be truly applied to ESP teaching.

**Construction of Aviation English Corpus**

Taking Aviation English as an example, its teaching targets will primarily be engaged in civil aviation flight, ground control, operational support, and aviation engineering technology in the future. Combined with actual work scenarios and job requirements, their professional communication needs are mainly reflected in: being able to understand professional civil aviation technical manuals, operation specifications, and equipment instructions; being able to consult civil aviation technical literature and operational bulletins; being able to use English to fill out technical records, work reports, and various business documents; being able to understand English air-ground communications, post instructions, and business communication content; and being able to use English to carry out civil aviation business exchanges, work coordination, and on-site communication. Based on the above needs analysis, civil aviation English materials of different genres were selected to construct a small specialized corpus. The number of texts and tokens (words) included in each genre is as follows:

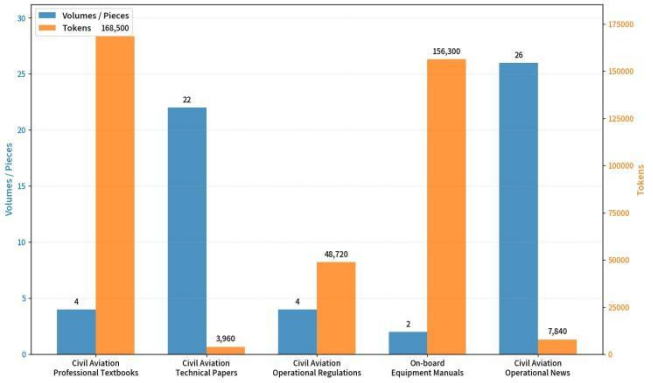


Figure 1 Genres, Volumes/Pieces, and Tokens of the Self-built Aviation English Corpus

## APPLICATION OF CORPUS IN AVIATION ENGLISH TEACHING

Although there is currently much research on ESP teaching and most higher education institutions offer corresponding ESP courses, the teaching results are often unsatisfactory. The author found that ESP teachers easily fall into two traps. First, the courses are knowledge-oriented. English is merely a medium for conveying professional knowledge; students only hope to clarify involved concepts and principles, while their attention to linguistic phenomena is seriously insufficient, leading to a lack of sensitivity to linguistic features. Second, ESP teachers conduct teaching activities based on personal experience. For example, deciding which word is important and needs a detailed explanation, or which word is unimportant and can be ignored. In reality, such "experience-based" talk can easily mislead students. If an ESP corpus can be effectively applied to ESP teaching, these two problems can be well resolved.

**Using Corpus to Help Students Recognize Characteristics of Aviation English**

TBy performing part-of-speech tagging on the self-built Aviation English corpus and then using its statistical characteristics, teachers can help students recognize the features of Aviation English. Furthermore, negation structures in aviation discourse are a key factor influencing comprehension; analyzing the decoding process of these structures through corpus analysis can help reduce communication risks (Borowska, 2017). For example, Aviation English uses fewer personal pronouns to reflect the objectivity of technical discourse and reduce personal emotional color. The frequency of nouns and noun phrases is higher than in General English, showing stronger nominal characteristics. This is because Aviation English contains more descriptions, definitions of objects, phenomena, and states, while General English describes more of daily life and personal emotions, showing more verbal characteristics. Passive structures appear more frequently than in General English, also for the sake of precise and objective expression. Based on these features, teachers can consciously conduct corresponding training for students, such as requiring them to avoid personal pronouns in translation exercises, or rewriting verb structures into noun structures and active voice into passive voice. Similar exercises can ground students' focus on ESP in


**DOI: 10.5281/zenodo.19648910**

"language," which is greatly beneficial for improving their sensitivity to linguistic features.

**Direct Use of Corpus as Classroom Material**

HThere are two main ways to use a corpus directly as a textbook: one is for the teacher to select a batch of sentences or an article from the corpus targeting a certain vocabulary, grammatical phenomenon, or topic for use in class. In this way, students can encounter a large amount of authentic data centered on the corresponding topic, vocabulary, and grammar. Another way is for students to directly participate in the material selection process, viewing the exploration of corpus resources as one of the activities in the language acquisition process. For example, the teacher guides students to search for sentences containing "wing" in the self-built corpus. Students can first set "wing" as the search term, which appears 82 times in the self-built corpus. Students then conduct group discussions, capture keywords, filter out useless information, and quickly browse these 82 sentences. For parts they are unsure of, they can enter the corresponding text via links for deep reading. Ten sentences are randomly sampled below:

1. The wing structure bears most of the aerodynamic forces during flight and provides lift for the aircraft...
2. Wing flaps are used to increase lift at low speeds during takeoff and landing...
3. The leading edge of the wing is designed to reduce resistance and improve aerodynamic efficiency...
4. Wing fuel tanks store a large amount of aviation fuel and affect the center of gravity..
5. Each wing is connected to the fuselage through strong attachment points...
6. Wing spoilers can reduce lift and increase drag to help the aircraft decelerate...
7. Ice on the wing surface seriously affects flight safety and must be removed before takeoff...
8. The wing includes ribs, spars, skin, and other structural components...
9. Navigation lights are installed at the wing tips to identify the aircraft position...
10. Wing design directly determines the performance, stability, and economy of the aircraft...

From the sample data above, students can find a wealth of information about the wing within the corpus. They can extract relevant content to summarize its definition, structure, and functions, after which each group reports, shares, and discusses the summarized information. Similar teaching methods can also be used for summarizing vocabulary collocations and inducing discourse frameworks and rhetorical features. This corpus-based classroom teaching avoids being confined to a single textbook and changes the traditional teaching model of teacher transmission and passive student acceptance. It is greatly beneficial for mobilizing student initiative and increasing the amount of information in the classroom. Furthermore, this teaching method, based on real-world communication scenarios, aligns with the evolving perspectives and needs of modern aviation English training (Borowska & Enright, 2016).

**Using Corpus to Assist Students' Autonomous Learning**

Although corpus-based classroom teaching can increase information volume and improve efficiency, classroom time is limited. There are certain limitations to learning Aviation English within this timeframe. A more effective method is to encourage students to use the corpus for autonomous learning in their spare time. As a type of technical English, Aviation English has very distinct linguistic features. If the frequency of students' language input can be significantly increased, it can activate their language awareness and improve their sensitivity to Aviation English characteristics.

Flowerdew (1993) once designed several learning activities for students using corpora for autonomous learning, aimed at stimulating curiosity, critical spirit, and reflective spirit, which is conducive to a series of activities cultivating innovative thinking:

1) Interactive Learning: Students complete a task in groups, such as observing and summarizing the collocation and usage of a certain word (e.g., observing whether "error" is used more as a verb or a noun in Aviation English and what words it usually collocates with). Students will consult and discuss during the problem-solving process.
2) Open Learning: Designing problems where answers have multiple possibilities or can be solved through various methods.
3) Analytical/Inductive Learning: Observing corpus-driven data, classifying and inducing at multiple levels based on the similarities and differences of linguistic features, and summarizing regularities. For example, using the corpus to observe whether the use of prepositions and articles in Aviation English differs from General English, or whether the top ten most frequent words in the Aviation English corpus are all functional words as in General English, and the reasons why.

## VERB NOMINALIZATION BASED ON AVIATION ENGLISH CORPUS

**Research Purpose**

The study aims to compare the usage of verb nominalization in General English and Aviation English, analyze the differences, and explore the expressive effects brought by these differences.

**Research Subject**

A small Aviation English corpus was built using randomly selected various types of Aviation English discourse. The free corpus analysis software AntConc was used for data extraction and analysis. The representative British National Corpus (BNC) was selected as the object for comparative research.

**Research Method**

As a type of technical English, Aviation English mainly describes objects, phenomena, and states, and rarely involves personal emotions and daily life. Therefore, the data will use more nouns and noun phrases. Compared with General English, it also shows stronger nominal characteristics. Verb nominalization is selected here as the entry point to study this nominal characteristic.

Halliday (1998) pointed out: "Nominalization may have formed earlier in scientific or technical registers." Some scholars even attempt to determine text genres by identifying grammatical metaphors of verb nominalization. Based on this, the author hypothesizes that Aviation English is more inclined to use verb nominalization structures compared with General English. To give students a more concrete understanding, the word malfunction (common in Aviation English) and its nominalized form *malfunctioning* were selected. The usage of these two words in the self-built Aviation English corpus and the BNC was counted. To


**DOI: 10.5281/zenodo.19648910**

avoid register overlap, non-technical registers were selected from the BNC for comparison.

**Result Analysis and Discussion**

The search results for the *malfunction* class (including *malfunction, malfunctions, malfunctioned*) and the *malfunctioning* class (including *malfunctioning, malfunctionings*) in the self-built corpus and the BNC are shown in Figure 2.

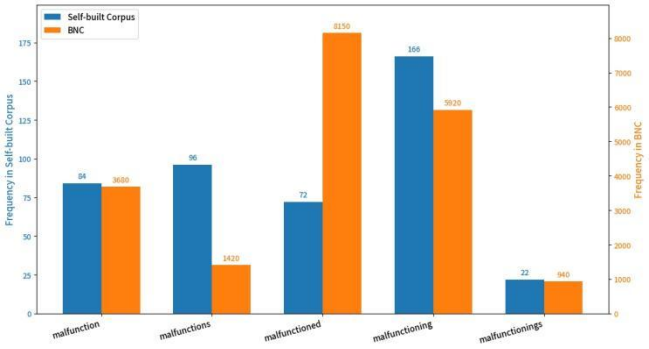


Figure 2 Frequency Comparison of malfunction and malfunctioning in the Self-built Corpus and BNC

Because the *malfunction* class includes three forms while the *malfunctioning* class includes only two, simply comparing percentages cannot lead to a conclusion about linguistic inclination. The total counts (440 and 20,110) were treated as the sizes of two new corpora, and a Chi-square test was used to analyze whether Aviation English tends to use verb nominalization represented by *malfunctioning*.

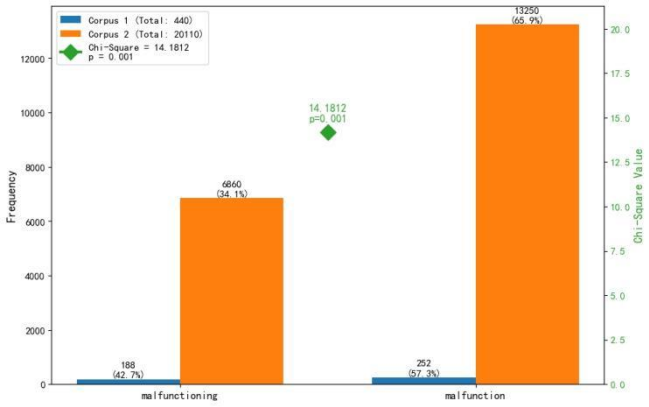


Figure 3 Malfunction and malfunctioning chi-square values

The Chi-square value calculation result is shown in Figure 3. It can be seen that the *malfunctioning* class has a higher frequency in the self-built Aviation English corpus, and its Chi-square value is statistically significant at the 0.001 level. Therefore, the hypothesis that "Aviation English is more inclined to use verb nominalization structures compared with General English" is established.

Based on these results, the expressive effects of nominalization are analyzed. Nominalization refers to "the process of forming a noun from some other word class or the process of deriving a noun phrase from an underlying clause." This phenomenon belongs to the category of conceptual metaphor mentioned by Halliday (1998). Non-linguistic major students do not need to master the origins of nominalization, but should understand its effects and try to use similar structures in papers and reports.

Halliday (1998) believes the use of nominalization is related to register. Written language often presents a nominalization trend because it treats experiences and phenomena as "objects," thus using more nouns. Conversely, informal registers, especially spoken language, describe events as "actions," thus using more verbs. In the process of nominalization, modal verbs and particles are often omitted, greatly reducing ambiguity and making sentence meanings more certain and objective. This is one reason why nominalization is more frequent in technical registers like Aviation English.

Noun phrases can contain one or more modifiers, allowing them to carry a large amount of information that verb phrases cannot match. For example, *A fault in the control system may cause continuous malfunctioning of the navigation and communication equipment.* If rewritten with a *malfunction*, it would require a clause: *A fault in the control system may lead to the result that the navigation and communication equipment would malfunction continuously.* In contrast, the sentence using *malfunctioning* is concise. Verb nominalization can highly concentrate information and make the expression more succinct, meeting the communicative needs of Aviation English.

Furthermore, verbs often require an agent (the doer), while noun phrases can "blur" the agent by not using personal pronouns, making the style more objective and concise. For example, *Since records of repeated malfunctioning show abnormal system status, perform the following inspection before the next flight.* The use of malfunctioning omits the agent while accurately and objectively expressing the meaning.

## CONCLUSION

Based on a self-built Aviation English corpus, authentic data can be retrieved around specific topics, vocabulary, and grammar for classroom teaching; students can also be inspired to conduct autonomous learning. Quantitative analysis of malfunctioning and malfunctioning classes demonstrated that Aviation English tends to use verb nominalization. It can be concluded that a self-built small corpus helps students focus on linguistic phenomena, improve sensitivity to language features, and enhance future professional communication skills. However, teachers should note that small corpora have gaps in representativeness and persuasiveness compared to large corpora. They should be used in conjunction with large corpora to obtain more precise data and conclusions. Furthermore, future teaching practices should further promote a joint training model for language teachers and industry experts to ensure consistency between assessment and teaching (Borowska & Petrashchuk, 2023).

## REFERENCES


1. Borowska, A. P. (2025) English for Aeronautical Communication, London: Routledge UK Limited Taylor and Francis Group.
2. 2.Borowska, A. & Simon, T. (2023) Collaborative Development: How Linguists and Aviation SMEs can Best Serve the Aviation Language Community, Aviation English Special Volume: Enhancing Efficiency in Aeronautical Communications/9th GEIA Proceedings, Applied Linguistics Papers, 27/2: 2023, 55-62.
3. 3.Borowska, A. & Petrashchuk, O. (2023) Best Practices for Test Construction based on Test of English for Aviation Personnel (TEAP) (in:) Aviation English Special Volume: Enhancing Efficiency in Aeronautical Communications / 9th GEIA Proceedings, Applied Linguistics Papers Papers 27/2: 2023, 63-78.
4. 4.Borowska, A. (2017) Decoding negation in aeronautical discourse, Kwartalnik Neofilologiczny, Nr



**DOI: 10.5281/zenodo.19648910**

3/2017, Warszawa, ISSN 0023-5911, ERIH List (NAT), pp. 366–379.
5. 5.Borowska A. & Enright, A. (eds) (2016) Changing Perspectives on Aviation English Training, Studia Naukowe – Vol. 29, Warszawa: Wydawnictwo Naukowe Instytutu Komunikacji Specjalistycznej i Interkulturowej, Uniwersytet Warszawski, pp. 111, ISBN 978-83-64020-33-9.
6. 6.Flowerdew, J. (1993). Concordancing as a tool in course design. System, 21(2), 231-244.
7. 7.Halliday, M. A. K. (1994). An introduction to functional grammar (2nd ed.). Edward Arnold.
8. Halliday, M. A. K. (1998). Things and relations: Regrammaticising experience as technical knowledge. In J. R. Martin & R. Veel (Eds.), Reading science: Critical and functional perspectives on discourses of science (pp. 185-235). Routledge.


**DOI: 10.5281/zenodo.19648910**