CREATING AN AVIATION ENGLISH LESSON: A CORPUS-INFORMED MATERIAL DESIGN

Deni Sapta Nugraha\textsuperscript{1,2}, Ilza Mayuni\textsuperscript{1}, Miftahulkhairah Anwar\textsuperscript{1}
\textsuperscript{1}Universitas Negeri Jakarta, Jakarta, Indonesia
\textsuperscript{2}Politeknik Penerbangan Indonesia Curug Tangerang, Indonesia

Abstract. Since ESP materials call for authenticity and specific content related to the field, books available on the market are sometimes thought to be inadequate for students’ target situations. As a subset of ESP, aviation English in this instance demands teachers to create teaching materials that cater to the students’ needs, wants, and necessities. For many teachers, creating course materials in the aviation setting continues to be challenging. As a result, this article offers an alternative approach to designing aviation English lessons using the corpus. There are essentially two approaches to designing corpus-based material: direct and indirect. Nevertheless, combining the two approaches can be a new challenge in getting students interested in the learning activity. A corpus-data-driven learning activity is created and embedded in a lesson. This activity potentially involves students examining the linguistic patterns, lexical categories, word frequencies, synonyms, and other pertinent vocabulary tasks in the corpus. The activity exerts a considerable influence on enhancing their vocabulary knowledge and vocabulary size. Alternatively, aviation English teachers offer an indirect approach by selecting reading materials from the free online corpus, such as, among others, COCA. Choosing a text from the corpus is believed to ensure the text’s authenticity. The teachers can automatically generate word or phraseology lists related to the field being learned from the text. Here, the corpus assists aviation English teachers in choosing suitable vocabulary for aviation students. This article highlights the approach of corpus-informed material design in an aviation context.

Key words: English for Specific Purpose, Corpus-informed material design, data-driven learning

1. INTRODUCTION

Material for language teaching, such as textbooks, is an inextricable element of language teaching and learning. For some teachers, commercial textbooks provide avant-garde and prospective approaches to language teaching and learning. They are regarded as the most relevant material for pedagogical purpose in language teaching due to their portability, usability, and ease of use. In English language instruction, textbook serve two major purposes:
to encourage the development of L2 competency and to provide comprehensive support for the teaching and learning of foreign languages (Vitta 2021).

Concerning English for general purpose (EGP), some literature discussed how to use and develop learning materials (Tomlinson 2011; Tomlinson 2012; Tomlinson 2016). According to Tomlinson (2016), teaching materials are everything exploited by teachers or students to bridge language learning either in a classroom or outside the classroom. However, the definition suggested is somewhat broad as it embraces a wide range of learning resources. In this case, we need to distinguish between teaching materials and teaching resources. Mishan & Timmis (2015, 12) illustrates that teaching resources are infinite sources, and the potential to turn them into teaching materials is guided only by our pedagogical thoughts. Therefore, the designer of teaching materials must be able to build pedagogical goals in the activities that go along with the teaching materials based on the principles of the language acquisition approach (Kuci 2020; Spiro 2013).

In ESP (English for Specific Purposes), the role of teaching material possesses an additional meaning compared to EGP, because students often pay attention to both aspects at once; language and content (Woodrow 2018, 346. Hutchinson and Waters (1987) have historically had a significant impact on ESP. They conceptualize a centralized learning approach in ESP that leans on a learning principle that the student entirely determines the content. They claim that ESP does not use a specific language or method, but is an approach to language learning based on the learner’s needs. However, research investigating language patterns in large-scale corpora could be the basis to state that ESP has a particular genre used in its learning (Hanford 2010; Paltridge 2013; Paltridge and Starfield 2011). Spiro (2013) further emphasized that ESP learners need to have specific discourse competencies in their socio-cultural context.

ESP teachers can exploit the corpus in preparing teaching materials to understand how a person uses language naturally in their community or cultural context. The corpus is utilized to know how a person speaks the language, not how a person must speak the language (Biber and Reppen 2015; Bloch 2013; Crosthwaite and Cheung 2019; Reppen 2011; Spiro 2013). Using authentic material through the corpus is also in line with the tradition of Systemic Functional Linguistics (SFL), which has used texts and corpus in studying linguistic phenomena (Wu 2009). Corpus-based research is also increasingly influential in exploiting authentic languages that can develop teaching materials (Bryun and Puquot 2021; Flowerdew 2015; Friginal 2018; Granger 2015; Mishan and Timmis 2015; Römer 2011; Römer 2008). These studies suffice to answer Chomsky’s criticism, which states that speech or written texts collected in a corpus are unnecessary and counterproductive because they are considered impractical to record all language structures (Andor 2004).

Regardless of this controversy, there is a critical aspect in ESP teaching materials, namely that ESP teaching materials need to reflect the objectives of the communicative situation. Therefore, ESP teaching materials usually include authentic material (Anthony 2018; Basturkmen 2010; Widodo 2016). This aligns with Richard (2001, 252), who classifies teaching materials into two types of approaches, namely authentic teaching materials and created teaching materials. Authentic teaching materials are not designed for language learning purposes (Basturkmen 2010). The texts are intended for experts in their fields that enable these materials to be easily accessed and exploited from language learning. In developing ESP teaching materials, ESP instructors and teachers of vocational courses need to work together. This is important to assist ESP teachers in choosing texts
relevant to their vocational fields. However, preparing course materials in the ESP context remains challenging for several ESP teachers.

The process of creating language learning material is intricate and involves multiple stages. For instance, conducting needs analysis, identifying target audience (e.g., local, national, or global), deciding on curriculum and examination requirements, accounting for current teaching technique, being open to innovation, considering local infrastructure, and ensuring that the material accounts for the requirements of language acquisition (Tomlinson 2016). When creating new resources, like textbooks, the teachers and students who will utilize the material are involved in the earliest stage (Richards 2015). In contrast to material created for a worldwide audience, it is simple to define a thorough profile of the target users and take into account the cultural context for local audiences (McDonough, Shaw, and Masuhara 2013). The materials writer must consider how teachers and students will use the textbook in either case.

Another thorny problem in developing material is choosing and determining approaches to language use and language learning (Richards, 2006, 2015; Tomlinson, 2016). Teachers incorporated their theories and opinions about how people acquire languages into their practice. Thus, in order to choose concepts that are helpful in creating content, such as the role of internal learner factors (i.e., affective and cognitive domains), language input, the learning environment, and teaching methodologies, language learning material writers need to have some experience and expertise in second language acquisition theories (Mishan and Timmis 2015). These considerations for ethical materials design influence the writers’ selection of content categories, tasks, and their order (Macalister and Nation 2020).

Corpus-informed materials are relatively unique because they are based on authentic norm. The contexts in which words and grammar structures used are genuine. Materials that are informed by corpora are created using research on particular categories of reliable corpora sources. The use of corpora-informed materials can support the selection of linguistic items appropriate for a given proficiency level, such as structures, vocabulary, collocational patterns, and contexts of use. They can also include research results on a variety of corpora (e.g., learner/expert, L1/L2, different genres), present results of corpus research in the appropriate form for learners, such as charts, graphs, concordances, and frequency lists, and serve for authentic materials sourcing for listening and reading. The use of learner corpora research to identify the needs of a specific learner group is a beneficial use of corpus research into the creation of instructional materials (McCarten 2010).

Creating corpus-informed materials makes use of both the findings of corpora research and guiding principles in materials design. In order to decide what content will be included and how it will be presented, teachers or writers might employ corpora in the textbook writing process. The corpus research outputs are chosen based on “what is worthwhile, what one’s students want, need, feasible and practicable” (Meunier and Reppen 2015, 566). Correspondingly, corpora support a logical design process and instructional decision making rather than developing an alternate framework for materials creation. A corpus-based method of materials design may be used to contrast this. After decisions regarding the materials content have been made, the corpus serves as a source of authentic example (Meunier and Reppen 2015). Previous research showed the effectiveness of implementing corpus-informed materials, either indirect or direct application (DDL) in many areas. To take for example, language teaching and learning (Flowerdew 2017), language assessment (Huang et al. 2018; Taylor and Barker 2008; Xi 2017; Kong 2017; Thirakunkovit et al. 2019), needs analysis for material development, (Staples 2019; Karpenko-Seccombe 2018), feedback and error
correction (Crosthwaite 2017), register and lexis (Crosthwaite, Wong, and Cheung 2019), teaching grammar (Lin 2021) and vocabulary in ESP context (Đurović and Bauk 2022). Despite the growing interest in corpus-informed material, little is known about how corpus-informed material are designed. This article is addressed to help ESP teachers, especially aviation English teachers, to provide a lesson by incorporating two different approaches: direct and indirect approach.

2. APPROACHES OF CORPORA IN LANGUAGE TEACHING

According to Crosthwaite (2020), DDL is a pedagogical strategy that directly engages learners with corpus data through printed materials or hands-on corpus consultation utilizing corpus tools. As part of the DDL methodology, students use corpus data to act as “language detectives” as they query, modify, and analyze a variety of output data, including concordance, word frequency, collocation, and multimodal forms of data. Smart (2014) advises two characteristics of DDL, the use of student-centered, discovery-based learning activities and the utilization of real language data as a basis for the creation of language learning resources. This method involves students participating in computer-based activities and conducting their own research by directly accessing corpora and concordance software. The second method, on the other hand, gives students indirect access to corpus data through corpus-informed, paper-based activities and materials that their teachers have developed beforehand (Boulton and Cobb 2017). These two methods have been referred to by several names, including the hard versus soft version of DDL, and learner-corpus interaction versus teacher-corpus interaction, direct versus indirect consultation of corpora, and hands-on and hands-off corpus usage (Sun and Hu 2020; Römer 2008).

![Diagram of the use of corpora in language learning and teaching](image-url)

Fig. 1 Application of corpora in language teaching (Romer, 2008)
In this article, the term direct approach refers to the teachers and learners having direct access to corpus data to investigate the language. Direct means introducing students to corpora and getting them to work with concordance lines, frequency, etc., so as they can explore and discover the language feature by themselves. By contrast, indirect approach refers to the use of teaching and learning materials which is informed with corpus data. In this case, students are not involved inductively in the process of language investigation. Learners do not interact directly with corpus tools. Indeed, they are likely not even aware of them.

3. BUILDING A CORPUS

There are many choices for teachers in using corpus tools for preparing corpus-based activities. Teachers can use available corpus online, offline, or they even can build their own corpus. However, what corpus is suitable for students depend on the purpose, context and students’ background. Mcenery & Hardie (2011) recommend the following software to build your own corpus.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Subscription</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>AntConc</td>
<td>free</td>
<td><a href="http://www.laurenceanthony.net/software.html">http://www.laurenceanthony.net/software.html</a></td>
</tr>
<tr>
<td>wordsmith</td>
<td>Purchase</td>
<td><a href="https://www.lexically.net/wordsmith/downloads/">https://www.lexically.net/wordsmith/downloads/</a></td>
</tr>
<tr>
<td>LancsBox</td>
<td>free</td>
<td><a href="http://corpora.lancs.ac.uk/lancsbox/download.php">http://corpora.lancs.ac.uk/lancsbox/download.php</a></td>
</tr>
</tbody>
</table>

Building a corpus is the last option for ESP teachers because there are many credible and large corpora available in the internet. However, as ESP teachers, especially aviation English teachers, sometimes we cannot find suitable or appropriate corpus for our teaching goal. Therefore, creating mini corpus is considered relevant for local audience purpose. This present study tried to give the example in creating a mini corpus by using AntConc version 4.0.5 released on February 2022 (Anthony 2022). It is used because the software is open access (free), it can be operated without internet data, and the interface of the window is user friendly. LancsBox can also be an option since the software can be utilized either offline or online.

3.1. Case 1 Using AntConc

Software can be downloaded at the following link: https://www.laurenceanthony.net/software/antconc/. Before downloading the software, we need to familiarize with our own computer and choose a suitable software that is compatible with our computer. Download and install the software to the computer. When creating a corpus, we need to determine the goal of language teaching in order to limit or set the criteria of the text being collected. After setting the criteria, collect suitable texts with our teaching purposes. That should be authentic use of the text, for example, building a mini dictionary of one particular aviation topic, e.g “aviation disaster”. In this case, texts should be classified only related to aviation disaster, we cannot use text beyond the topic to limit and specify the profile of the text. After collecting the text, convert all the text into plain text .txt file in a notepad application. However, if we use the
latest version of AntConc, source files can be in form of docx, pdf, and txt. Figure 2 shows the example of source text that will be inputted into AntConc.

![Source Text Example](image)

**Fig. 2 Example of sources text**

Process all the texts into AntConc when the teachers considered that the texts are sufficient for facilitating their teaching purpose. As the files are successfully loaded, the mini corpus is created. Teachers can start using corpus for their teaching. There are some interesting features that can be utilized for teaching, such as word frequency analysis, concordance analysis, N-gram and collocation analysis. Each feature has unique function depending on the purpose.

### 3.1.1. Doing Analysis: word frequency

Based on Figure 3, we can see that the *aviation disaster* corpus has 3911 word tokens with 1196 word types. The total number of texts is 10. From the number of tokens, we can classify these words into word classes (part of speech/POS) so that the source of the vocabulary that will be inputted into the dictionary is in accordance with needs of teachers and students.
3.1.2. Doing Analysis: Concordance

We can analyze how the word aircraft is in a context. Students can observe the context that forms the meaning of the word. In this case, students can know authentically the lexical bundles of the word aircraft via corpus. In addition, students can also see the real context based on sentences in the text and understand the word through context (lexical bundles). We can click the word on the list, the context of the accompanying sentence will appear. It is commonly known as keyword in context (KWIC) as shown in Figure 4.
3.1.3. Doing Analysis: N-gram

Based on Figure 5, we can see common expressions that appear in the corpus in which the aircraft lexical bundle of size 2 is followed by the word e.g., descend, arriving, altimeter, exploded, etc.

![Fig. 5 N-Gram of the word aircraft](image)

Furthermore, we can do a collocation analysis to know how one particular word collocates with another word via corpus (see Figure 6). It is essential in vocabulary activities to enlarge students’ vocabulary knowledge. Figure 6 shows that the word descend, in the present corpus, collocates with the words alert, saying, descend, and sounded. However, to enlarge the word list, teachers can add more text in the corpus.

![Fig. 6 Collocations with the word “descend”](image)
4. CORPUS-INFORMED ACTIVITIES

Several pedagogical activities can be designed from one particular task. Teachers for example can make use of reading the text drawn from the corpus, then creating appropriate practice for students activities such as vocabulary, grammar, structure, writing tasks, to utilize the reading section. This article advises how to identify and analyze vocabulary from authentic material.

![Diagram showing the process of creating teaching materials from a corpus](image)

Fig. 6 Indirect corpus-informed material design

COCA is used to select and analyze appropriate material that relates to specific knowledge of aviation. Teachers can log in COCA (https://www.english-corpora.org/coca/) and write in a search bar one specific vocabulary related to the topic. To sort out the most appropriate topic, then go to the context, select one of the sources, COCA will directly take the user into source text. You can retrieve the reading materials from the available URL link in COCA. However, if teachers want to create vocabulary activities, COCA has Analyze text features. It helps material designer to list appropriate vocabulary, words frequency that has been classified into low, mid, and high. Here, teachers also can analyze every single word with their synonym, antonym, collocation, and also definition. To design language learning materials, we can also use the lexutor site, in the corpus we can design multiple choice worksheets, cloze tests, open ended questions, matching, and so on. The corpus is web-based and is free. Here is the address at which it can be visited https://www.lexutor.ca/cloze/vp/users/The_Problem_with_Lexie.html.

4.1. Identifying and analyzing vocabulary with COCA

4.1.1. Procedure for teachers

1) Access the online tool corpora from the sites
2) Look for the material from the internet, for example the article from Lacagnina, 2014 entitled “Mismanaged descent” https://www.smartcockpit.com/docs/identified-threats-when-transitioning-from-imc-to-vmc.pdf
3) Copy and paste the text into the analyze text box as following:
Fig. 7 Analyse text feature in COCA

4) After the lists are generated, decide which words to focus on in classroom activities. For example, focus on low-frequency items, such as airplane, airspeed, autopilot, auto throttle, descent, glidepath and go around, as useful for both receptive (e.g., comprehension or main idea question), and productive (e.g., group discussion or summary writing) activities related to articles.

5) Some lexical items such as descend are relevant as multiple parts of speech (POS) and COCA permit students to discover this distinction on their own. Alternatively, teachers can further categorize the subject-specific vocabulary based on their own content knowledge. For instance, this article contains many parts of an airplane; or teachers can directly pick the vocabulary related to the discipline such as the following:
Creating an Aviation English Lesson: a Corpus-Informed Material Design

This model can be implemented as indirect approach of a corpus-informed material design. Corpus assists teachers to identify vocabulary in many aspects such as collocation, synonym, antonym, definition, parts of speech, word frequency, context, concordance and many more. Teachers can create several activities for each aspect by providing paper-based activities or internet-based activities. In this model, students are not involved in investigating vocabulary through corpus. Teachers create and design the material with the help of corpus.

![Figure 8: Vocabulary list](image)

This model involves teachers and students to directly access corpus data. They utilize corpus to investigate the language phenomena as their teaching and learning activities. In this model, both teachers and students should be familiar with the corpus. Therefore, the ability of using corpus is the prerequisite before teaching and learning process. The following is the example of direct procedure of corpus-informed activities.

![Figure 9: Direct corpus-informed material](image)
4.1.2. Task for students

Analyze vocabulary identified by the teacher through discovery and exploration. After the instructor has determined the word list from the article, students can follow procedure for students. An example is given that permits the students to discover part-of-speech distinctions used in the article (i.e., a word as a noun and as a verb). Instructor could also indicate in the word list which part-of-speech to focus on.

Ask students to go to COCA. Click on “Frequency List” on the left side of the page. Teachers ask students to consider the following excerpt from the article “Mismanaged descent”:

“The indicated airspeed was on target, and the approach path lights showed that the airplane was just slightly high as it descended below 500 ft, the point at which the stability of a visual approach typically is judged.”

Students are asked to answer guided question for direct corpus activities. Some useful questions might be addressed such as: What part-of-speech is descended in this context? Which synonyms could you use for descended in the context of this excerpt? Which definition do you think fits the context of this excerpt? (click New: DEFIN+SPEC+GENL)

In the text, click “descended”, and choose the part-of-speech that matches the context described earlier.

![POS category](Fig. 10)

After answering the questions, ask students to examine the collocates for descended. Click on a verb collocate and take a look at the concordance entries. Ask them to identify whether they see any patterns on the list.
Further activities can be added by asking them to choose another collocate such as adjective. Take a look at the concordance entries, then ask students whether they see any particular pattern on preposition. Last activities can be in the form of a written activity. Teachers asked students to create original sentences that contain the collocate and descended, with at least one sentence using an aviation topic for instance. The detailed procedure is available in Appendix 1.

5. CONCLUSION

Direct and indirect approach of corpus-informed material design can be an alternative way for aviation English teachers in designing their own material for the purpose of local audience. Customizing teaching material is not an easy task. There are some principles in material development that should be considered, including needs, wants, and necessities. Additionally, teachers should include their theories and opinions about how people acquire languages into their practice. To choose concepts like the importance of internal learner variables, language input, the learning environment, and teaching approaches, they frequently need experience in language learning and SLA theories. Teachers should include their theories and opinions about how people acquire languages into their lessons. To choose concepts like the function of internal learner variables, language input, the learning environment, and teaching approaches, they frequently need to have familiarity with language learning and SLA theories.
REFERENCES


Creating an Aviation English Lesson: a Corpus-Informed Material Design


### APPENDIX 1

<table>
<thead>
<tr>
<th>Lesson Plan</th>
<th>Aviation – Identifying and Analyzing Vocabulary: Mismanaged descent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>50 minutes</td>
</tr>
<tr>
<td>Resources</td>
<td><a href="https://www.english-corpora.org/coca/">https://www.english-corpora.org/coca/</a></td>
</tr>
<tr>
<td>Pre-task</td>
<td>10 minutes</td>
</tr>
</tbody>
</table>

1. Elicit the knowledge, skills, abilities and personal qualities required to be a pilot in emergency situation, accident or non-routine situation.
2. List on board and elicit types of questions that can be asked to get this information.

### Dos for teacher
1. Teacher verifies vocabulary related to the topic with students’ prediction;
2. Teacher provides vocabulary list, synonyms; and definitions, alternatively ask students to search from the corpus by following the procedures;
3. Teacher delivers a piece of excerpt from the text to be discussed.

E.g. *The indicated airspeed was on target, and the approach path lights showed that the airplane was just slightly high as it descended below 500 ft, the point at which the stability of a visual approach typically is judged*.

### Task for students
Read the procedures

1. Go to COCA Click on “Frequency List” on the left side of the page.
2. Consider the following excerpt from the article “Mismanaged descent”:

   *The indicated airspeed was on target, and the approach path lights showed that the airplane was just slightly high as it descended below 500 ft, the point at which the stability of a visual approach typically is judged*.

SS complete the following questions

<table>
<thead>
<tr>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. What part-of-speech is descended in this context?</td>
<td>Verb</td>
</tr>
<tr>
<td>b. In the search bar, type “descended,” and choose the part-of-speech that matches the context described earlier.</td>
<td>Verb</td>
</tr>
<tr>
<td>c. Which definition do you think fits the context of this excerpt?</td>
<td>Free fall under the influence of gravity; move downward and lower.</td>
</tr>
<tr>
<td>d. Which synonyms could you use for descended in the context of this excerpt? Choose two of them.</td>
<td>Go down, fall</td>
</tr>
<tr>
<td>e. Examine the collocates for descended. Click on a verb collocate and take a look at the concordance entries.e.g. minimize. Do you see any patterns?</td>
<td>e...descended + a + N. ..descended + pref [from, for, in, into, on]. …descended + the f. The aircraft start to descend rapidly due to engine failures</td>
</tr>
<tr>
<td>f. Create two original sentences that contain the collocate and impact, with at least one sentence using an aviation topic.</td>
<td></td>
</tr>
</tbody>
</table>
Choose another collocate that is adjective. Take a look at the concordance entries e.g. steep. Do you see any patterns? Notice prepositions in particular.

Create two original sentences that contain the collocate and descended, with at least one sentence using an aviation topic.

1. Discuss the answer (10 Minutes)

<table>
<thead>
<tr>
<th>Post Task</th>
<th>10 Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflection</td>
<td>Ask a volunteer to express his/her experience using corpus in identifying vocabulary.</td>
</tr>
</tbody>
</table>

Worksheet

Analyze vocabulary identified by the teacher through discovery and exploration:

1. Go to COCA Click on “Frequency List” on the left side of the page.

2. Consider the following excerpt from the article “Mismanaged descent”:
   “The indicated airspeed was on target, and the approach path lights showed that the airplane was just slightly high as it descended below 500 ft, the point at which the stability of a visual approach typically is judged.”

3. What part-of-speech is descended in this context?

4. In the search bar, type “descended,” and choose the part-of-speech that matches the context described earlier.

5. In the search bar, type “descended”, find out the definition by clicking the POS either N or V. Which definition do you think fits the context of this excerpt?

6. In the search bar, type “descended,” Select POS N on word impact. Which synonyms could you use for descended in the context of this excerpt? Choose two of them.

7. Examine the collocates for descended. Click on a verb collocate and take a look at the concordance entries, e.g. minimize. Do you see any patterns?

8. Create two original sentences that contain collocate and impact, with at least one sentence using an aviation topic.

9. Choose another collocate that is an adjective. Take a look at the concordance entries, e.g. significant. Do you see any patterns? Notice prepositions in particular.

10. Create two original sentences that contain the collocate and impact, with at least one sentence on an aviation topic.