THE JOURNAL OF TEACHING ENGLISH FOR SPECIFIC AND ACADEMIC PURPOSES
Vol. 11, Nº 1, 2023, pp. 1–9

UDC: (811.111+005.336.6):374.823.4 https://doi.org/10.22190/JTESAP230120002S

Review research paper

LANGUAGE PROFICIENCY TEST DESIGN FOR A GROUP OF MILITARY PILOTS: CRITERIA, CONSTRUCT AND PURPOSE

Ana Lígia Barbosa de Carvalho e Silva

São Paulo State University (UNESP), Brazil

Abstract. *Despite the growing number of studies devoted to Aeronautical English, the role of language assessment to military pilots seems to be largely overlooked. To narrow this gap, this study aims at putting forward some findings of a broader study (Silva 2022)) which comprises a test design for a group of pilots from the Air Demonstration Squadron (EDA) of the Brazilian Air Force (FAB). Test design refers to the definition of crucial aspects of a test, prior to its development, such as: i) the criterion, or the communicative behavior expected in the target situation; ii) the construct, i.e., what the test is intended to assess; and iii) the purpose, which is mainly related to the reason why the test is designed and the interpretation and uses of its results. Language Assessment for Professional Purposes (LAPP) (Knoch and Macqueen 2020) is the main theoretical framework that guided this "Mixed Methods Case Study Design" work (Creswell and Clark 2018). Results indicate that the design of an English test for EDA pilots, based on a needs analysis, should be wide enough to encompass the use of Aeronautical English (phraseology and plain English for radiotelephony communications) and a more general language. As for the test purpose, data analyses indicate that test results could be useful in the selection of pilots who would be more suited to fly together as crew members. The expected contribution of the study is to provoke reflections on the main characteristics of an English proficiency test for military pilots, without losing sight of the commonalities it may have with language needs for other pilots around the world.*

Key words: *Aeronautical English, test design, pilots, language assessment for professional purposes (LAPP, language proficiency, ESP.*

1. SETTING THE RESEARCH SCENARIO

In recent years, more precisely after 2004, when ICAO published the first version of the Language Proficiency Requirements (LPRs), Doc 9835 (ICAO 2010), numerous researchers have shown interest in the specific language needed by pilots and air traffic controllers (ATCOs) for radiotelephony communications, known in the literature as Aeronautical English (Borowska 2017; Tosqui-Lucks and Silva 2020). Attachment A to Annex 1 of the same document even introduced a six-band analytical rating scale, which

Submitted January 20th, 2023, accepted for publication March 13th, 2023
Corresponding author: Ana Lígia Barbosa de Carvalho e Silva, São Paulo State University (UNESP), Rua Quirino de Andrade, 215 - Centro - São Paulo, Brazil
E-mail: ana-ligia.silva@unesp.br

has been used to establish levels of language proficiency to pilots and ATCOs' language abilities for certification purposes.

It is important to notice that, despite ICAO's efforts to reinforce good practices in language assessment for aviation, this international organization did not prescribe, at any point, a particular test to be used, leaving such responsibility to testing services groups and national agencies. Thus, many Aeronautical English tests started being developed in different parts of the world, not all of them with the same quality, as demonstrated in a report on a survey of Aviation English tests produced by C. Alderson (2008).

Although civil aviation pilots and ATCOs were the main target of ICAO language recommendations, the LPRs soon started being used to guide language teaching and assessment for aeronautical communications in the military scenario as well, according to previous studies (Santos, Pacheco, Reyes, et al. 2018; Katsarska 2017; Er, Kirkgöz 2018; Bratslavska 2020; Park 2020). Different ways to teach and assess this very specific language started being debated, due to the high-stakes nature of the decisions made, based on the interpretations of the results of exams guided by the LPRs. The construct of Aviation English, and even the appropriateness of ICAO rating scales, started being questioned worldwide (Knoch 2014; Kim and Elder 2015).

In that regard, a language needs analysis research for pilots from the Air Demonstration Squadron from the Brazilian Air Force, or "Esquadrilha da Fumaça" (EDA[1]) (Silva 2016) concluded that further studies were necessary to better understand what, exactly, needed to be assessed in an English language test for this specific group of pilots, who represent Brazil and its Air Force, not only in the air, with acrobatic air demonstrations, but also in other international events, as diplomatic representatives. To fulfil this gap in research, the aim of this study is to put forward the results of a broader investigation (Silva 2022) that comprises a test design for EDA's pilots, stating the criteria, construct and test purpose. With that in mind, the research questions that guided this study were:

1) what is the target language use domain, for EDA pilots? (criteria)
2) what are the knowledge, skills and abilities (KSAs), that have to be tested in a language proficiency test that is representative of the TLU situation for EDA pilots? (construct)
3) what are the potential uses of an English language proficiency test results for EDA pilots?

In short, stating the theoretical principles of an Aeronautical English test design for a group of military pilots, according to a language needs analysis, is the great contribution this study brings to the literature of English for specific Purposes (ESP), Language Assessment for Professional Purposes (LAPP) and Applied Linguistics, at large.

2. The Pathway for Data Collection and Analysis

The methodological approach chosen for this investigation was a Mixed Methods Case Study Design (MMCD) (Creswell and Clark 2018) described as a complex application of a convergent mixed methods, in which qualitative and quantitative data were used, within the framework of a case study. The final product achieved is a thorough in-depth description

[1] A few acronyms used in this paper will remain in the same way they are used, originally, in Portuguese. That is the way they are commonly referred to, which means that changing them would sound artificial and unnatural.

of a case, i.e., a test design for EDA pilots, with more details than a study in which only qualitative or quantitative data are used in isolation.

The twenty six participants in the study were categorized as follows: i) the Brazilian Air Force Academy commander; ii) one expert in international air traffic control, who applied the operational and Aeronautical English phraseology test (TAI) at EDA; iii) two Brazilian Air Force Academy English professors (one who teaches General English and the other who teaches ESP); iv) fourteen pilots, members of EDA during data collection, or *Actual Pilots* (AP); ii) eight EDA *Veteran Pilots* (VP), who had been members of EDA, but later on, in their careers, had also worked as pilots on the international civil aviation. This parameter was used for choosing these participants because they were familiar with both ways of assessing language proficiency for pilots in Brazil: the exams developed by the Brazilian Air Force, and the exam developed by the National Civil Aviation Agency (ANAC).

A workshop session was provided to APs, organized by this researcher in conjunction with an exam rater from ANAC, in order to introduce the *Santos Dumont English Assessment* (SDEA), the only Aeronautical English exam officially applied to civil pilots in Brazil. This procedure was considered necessary because the APs had never sat the SDEA, and only knew the assessment system adopted by the Brazilian Air Force. Interviews, questionnaires, observations and documental analyses were the methodological instruments used for data collection and triangulation. Qualitative data was analyzed through content analysis (Bardin 2016), whereas quantitative data, generated from closed questions in questionnaires, were analyzed through descriptive analysis.

3.THEORETICAL CONSIDERATIONS

In the past four decades, a lot has been debated about the main characteristics of ESP language tests. Douglas (2013: 368) clearly explains "first, that language use varies with context, second, that specific purpose language is precise, and third that there is an interaction between specific purpose language and specific purpose background knowledge". There seems to be no better context than aviation to illustrate this understanding of language for professional purposes and its specificities.

A conceptual difference, with great impact in test design, has been established between the umbrella term *Aviation English*, which encompasses all the language used by professionals in aviation at large, such as mechanics, engineers, and *Aeronautical English*, and the specific language used by pilots and air traffic controllers (ATCOs) to communicate over the radio (Borowska 2017; Tosqui-Lucks and Silva 2020). According to these authors, Aeronautical English has basically two elements: i-) *standard aeronautical phraseology*, a very specific register to be used use in normal situations; ii) *plain aeronautical English*, for abnormal or emergency situations.

The intriguing question that seems to remain unanswered in language assessment for specific/professional purposes, and consequently for Aeronautical English tests, is not only *how much*, but also *how*, language and operational-content knowledge could be assessed, in order to make the test as representative as possible of the communicative situations the pilots will encounter in real life. In this regard, Moder (2013:239) states that "it is essential to include both routine and unexpected radiotelephony tasks in Aviation English tests, making use of representative authentic combinations of phraseology and plain language".

Language test design is the moment *per se* to make crucial decisions regarding the necessary alignment between the test purpose, the language to be assessed in test tasks, with its corresponding assessing criteria, and, finally, the language use situation. A detailed definition of *what*, *how* and *why* we are testing language for professional purposes, based on a comprehensive needs analysis, operationalized in clear test specifications, can steadily guide test development and application.

A fundamental concept that illuminated this study was a socially oriented theory, followed by a model with samples of the linguistic repertoire in the workplace to help determine the test construct, in terms of language specificity. This model, in the field of Language for Professional Purposes, called *Codes of Relevance* (Knoch and Macqueen 2020) is presented in Figure 1.

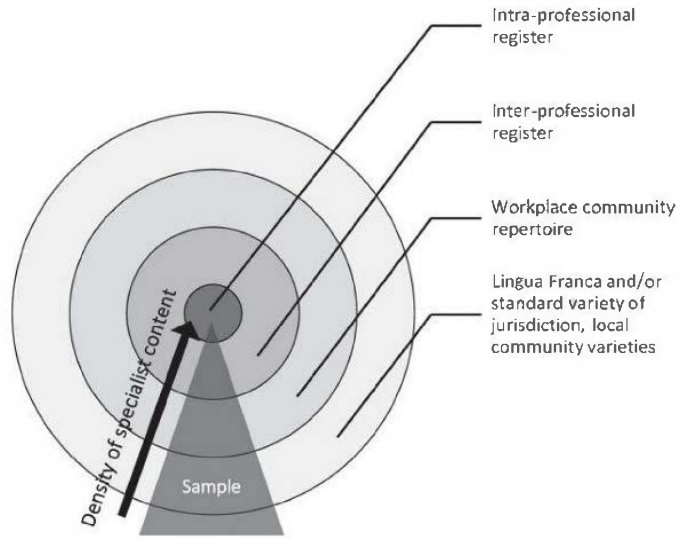

Fig. 1 Codes of Relevance (Knoch and Macqueen 2020).

This model introduces a current and innovative vision to explain the different contexts of language use for specific purposes, demonstrating the complexity of the interaction between language and specific content. The illustration, in Figure 1, shows how much task content in the assessment instruments can vary in relation to the density of their specificity, depending on the different interlocutors and contexts of language use in the target language use (TLU) domain. It can be observed that professional communications can vary from a high level of specificity, passing through a slightly less specialized register, but still between professionals of the same area of work, to reaching communications with the general public, with the use of local lingua franca varieties.

The most inner circle, called intra-professional register (Knoch and Macqueen 2020) represents the use of language by a reduced number of professionals who share a very specialized and precise linguistic repertoire at the workplace, due to a certain amount of shared background knowledge, with greater levels of inseparability between language and specific content.

As the circles move outwards, the degree of specificity decreases progressively, making communications more accessible to a larger number of individuals, until reaching the language used in everyday conversations among ordinary people. Thus, the sphere marked by the inter-professional record represents communications between professionals with some specific and shared background knowledge. The last two circles represent less density levels of specialized registers, getting closer to what is popularly known as "general language".

4. Answer to Research Question 1: Test Criteria/Domain

Data analyses, including triangulation of participants, instruments and methods, provide convincing evidence that there are basically two categories of language use situations for EDA pilots: i) while flying, in routine and in non-routine situations; ii) on the ground, in many different situations. Thus, the necessary language for such a wide domain should fall in different points in a continuum of specificity (Douglas 2000).

In order to answer research question 1, the language needs analysis previously conducted for EDA pilots (Silva 2016) had to be reviewed and updated.

Firstly, results from the Present Situation Analysis (PSA) indicated that both assessment systems used for pilots in Brazil (from ANAC and from the Air Force) should undergo a validation process. Underrepresentation of the construct, which is one of the threats to vality (Messick 1989), was observed in both contexts. This phenomenon occurs when "the test is too narrow and fails to include important dimensions or facets of focal constructs" (S. Messick 1989: 4). For example: the SDEA, from ANAC, does not directly assess the use of *standard aeronautical phraseology* in English, because ICAO considers it "operational knowledge", and not language knowledge. On the other hand, the assessment system at the Air Force does not directly assess the use of *plain aeronautical language* in abnormal or emergency situations. Therefore, none of the existing exams that have been used in Brazil, so far, seem to be adequate to assess EDA pilots' communication needs in English.

Secondly, results from the Target Situation Analysis (TSA) corroborates previous findings (Silva 2016), as the situations of language use for these pilots have remained unchanged, and spotted in a continuum of specificity, ranging from a very specific register (*aeronautical phraseology*), to be used in normal situations during flights, moving on to less specific register (*plain aeronautical English*), for abnormal and emergency flight situations, and a more general register, for situations on the ground, where EDA pilots have to perform social and professional tasks (in aeronautical events or not), acting as real representatives of their country and their Air Force.

5. Answer to Research Question 2: Test Construct

Taken altogether, quantitative and qualitative data analysis provide enough evidence that the theoretical and stated construct of an Aeronautical English proficiency test for EDA pilots is rather complex.

Evidence also indicates that knowledge, skills and abilities to be assessed cannot be totally separated from content or operational knowledge. The main linguistic skills to be assessed are listening and speaking, with use of precise specific vocabulary, integrated with note taking, in real time, while multitasking (communicating, piloting and navigating). It is also mandatory to assess certain strategic competences, such as asking for repetition and/or

reformulations, making inferences and clarifications, negotiating meaning and adapting to a less proficient interlocutor. Also important is the openness to regional accents and local air traffic rules, not always clearly explicated on flight plans. Other strategies to be assessed include the use of English language on the ground to recognize implicit meanings, express opinions, compare, demonstrate hospitality, express humor, answer questions, give interviews and speak in public.

As for test format, findings in the study show that language proficiency for EDA pilots should be assessed by means of a *performance* test, which simulates the real situation of language use, including phraseology in English and plain language, with the appropriate transition between these two registers. The use of technology in test design and administration is strongly recommended.

Of similar importance is the assessment of code-switching, as pilots can choose to communicate with ATCOs either in Portuguese, their mother tongue, or English, the international language of aviation, while flying in the Brazilian airspace. Other aspects that should never be disregarded are situational awareness, intercultural diversity and distress, in emergency situations. In order to comprise such an extended construct, more than one test instrument could be used (alternatively, one test with different parts).

To better illustrate the test construct here proposed, Figure 2 shows the different levels of specificity that needs to be operationalized, according to *Codes of Relevance* (Knoch and Macqueen 2018), as mentioned in Section 3.

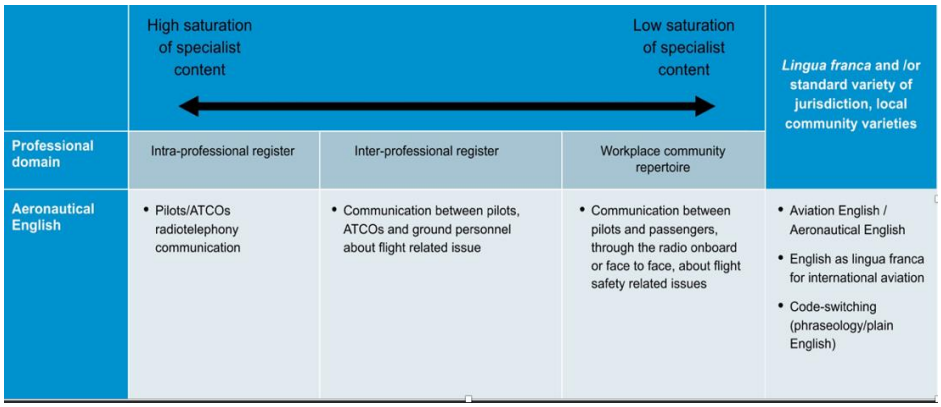

Fig. 2 Codes of Relevance for EDA pilots

5.1. Codes of relevance: Intra-professional register

Aeronautical phraseology, for normal situations during an international flight, is the most specific register that should be assessed in a language proficiency exam for EDA pilots. Alternatively, the use of plain English, for abnormal and emergency situations should also be assessed. The use of content knowledge, phraseology and plain language should not be considered separately, but in an integrated way, as close as possible to real life situations of language use.

5.2. Codes of Relevance: Inter-professional register

It is necessary to assess the language needed buy EDA pilots to interact with a variety of interlocutors, in different contexts of language use, i. e., on the ground, during stopovers and at final destination, to discuss many topics, such as those related to air demonstrations, the aircraft flown by EDA team, and its operational equipment. Therefore, the language register should include a combination of general and specific vocabulary, aviation related, or not.

5.3. Codes of Relevance: Workplace community repertoire

It comprises the necessary language that EDA pilots need to communicate with air show organizers, authorities, the public in general, and the press, before or after air demonstrations abroad. Other actions include taking part in global interchange programs, participating in aviation fairs and events, and hosting foreign visitors to EDA's headquarters. All these situations of language use involve a wide range of vocabulary, aviation related, or not.

6. Answer to Research Question 3: Test Purpose

Several findings of this study support the assumption that, in practical terms, the results of an Aeronautical English exam for EDA pilots could be useful to assist EDA's Command and Operational Sectors in making relevant decisions, such as choosing the most appropriate pairs of pilots to compose a military crew in a double-seat military aircraft for EDA's travels overseas, with a view to greater safety in aeronautical communications during these flights. Test results can also be used to help these officers choose those pilots who would most effectively and appropriately represent Brazil and its Air Force in certain events on the ground.

At this point, it is relevant to highlight that Brazilian Air Force pilots, who become EDA members, spend only a few years of their military careers in that position. After a period of no more than five years, they have to leave the group, in order to give way to other pilots, who also wish to join the most famous squadron in the Brazilian Air Force.

As it is easily presumed, the admission process to select a new member of the group is very competitive, including operational, administrative and personal criteria. To these days, English language proficiency is not one of these criteria. Maybe in the near future, in case language policy at the Brazilian Air Force changes, results of a language proficiency test, built upon a serious language needs analysis, could be one more useful tool to help them make assertive and appropriate decisions in choosing the best of the best pilots, to represent Brazil and its Air Force internationally, including language-wise, in the air and on the ground.

7. Conclusion

In the process of designing or validating a language test for professional purposes, we may often times ask ourselves: Are the test tasks, test instruments and rating criteria appropriate and adequate to the test purpose? Is the language being assessed really representative of the language needed in real life situations? What are the possible interpretations and uses of test results? What is the potential impact of the decisions made, based on test results, in terms of fairness and justice? It remains clear, by the end of this study, that answers to these questions

can only be supported by a solid tripod, made of a clear definition of the *test criteria, test construct and test purpose*, based on a comprehensive needs analysis. English language test design for military or civil pilots, all the same, should follow suit. Still, future studies will have to look at ways to put together research and discussions that have been taking place in both aviation sectors, civil and military, so that we can all benefit from the contributions that could be given to one another, for safer and better skies to all of us.

REFERENCES

Agência Nacional de Aviação Civil (ANAC). *Language Proficiency.* Brasília, 2022. <https://www.gov.br/anac/pt-br/assuntos/regulados/profissionais-da-aviacao-civil/processo-de-licencas-e-habilitacoes/proficiencia-linguistica-2> (16.01.2023) .

Alderson, J. C., *Final Report on a Survey of Aviation English Tests.* 2008. <http://www.ealta.eu.org/documents/archive/alderson_2008> (16.01.2023).

Borowska, A. P. *Avianlinguistics: The Study of Language for Aviation Purpose*. Frankfurt: Peter Lang, 2017.

Brasil. Teste *Diagnóstico em Idiomas Estrangeiros* (TDIE). ICA 37-324, 2016. <https://www2.fab.mil.br/unifa/images/proeei/ICA_37_324.pdf.>(16.01.202).

Brasil. *Fraseologia de Tráfego Aérea. MCA 100-16. 2020* <https://www.sislaer.fab.mil.br/terminalcendoc/acervo/detalhe/10896?guid=1605830408521&retu%20rnUrl=%2Fterminalcendoc%2Fresultado%2Flistar%3Fguid%3D1605830408521%26quantidade%20Paginas%3D1%26codigoRegistro%3D10896%2310896&i=20> (16.01.2023).

Bratslavska, A., *Special features of forming the readiness of future military pilots for professional interaction when using the common air space*. (in:) "Adaptive Management: Theory and Practice, Series Pedagogics", 8(15), 2020. < https://doi.org/10.33296/2707-0255-8(15)-09)> (16.01.2023).

Bruno, M. L. *Uma crítica à avaliação de proficiência linguística, em inglês aeronáutico, aplicada aos pilotos da força aérea brasileira*. Monograph. Escola de Aperfeiçoamento de Oficiais da Aeronáutica, Rio de Janeiro, 2018.

Douglas, D. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press, 2000.

Douglas, D. ESP and assessment. (in:) Paltridge, B., Starfield, S. (eds)., The handbook of English for Specific Purposee (p. 367-383). Chichester, UK: John Wiley & Sons, 2013.

Er, M. and Kirkgöz, Y., *Introducing Innovation into an ESP Program: Aviation English for Cadets.* (in:) Key Issues in English For Specific Purposes in Higher Education.179-199. 2018. <http://dx.doi.org/10.1007/978-3-319-70214-8_11> (16.01.2023).

Garcia, A. C., *What do ICAO language proficiency test developers and raters have to say about the ICAO Language Proficiency Requirements 12 years after their publication? A qualitative study exploring experienced professionals' opinions*. 2015. <https://www.icaea.aero/wp-content/uploads/2015/10/Masters-Dissertation_Angela-Garcia_28052015.pdf> (16.01.2023).

Katsarska, V. *Harmonization of University Aviation Courses*. (in:)The Journal of teaching English for specific and academic purposes, 5(3), 449-458. 2017. <http://espeap.junis.ni.ac.rs/index.php/espeap/article/view/538> (16.01.2023).

Kim, H. and C. Elder. *Interrogating the construct of aviation English: Feedback from test takers in Korea*. Language Testing, 32(2), 129-149. 2015. <https://journals.sagepub.com/doi/10.1177/0265532214544394> (16.01.2023)

Knoch, U. Using subject specialists to validate an ESP rating scale: The case of the International civil Aviation Organization (ICAO). English for specific Purposes, 33, 777-86. 2014. <https://www.sciencedirect.com/science/article/abs/pii/S088949061300063X> (18.01.2023).

Knoch, U. and Susy Macqueen. *Assessing English for Professional Purposes*. New York: Routledge, 2020.

Messick, S. *Validity*. (in:) R. L. Linn (ed.). Educational measurement. 3rd. ed., 13-103, 1989.

Moder, C. L. Aviation English in B. Paltridge and S. Starfield (eds.), The handbook of English for Specific Purposes (pp-227-242). Chichester: John Wiley & Sons, 2013.

Monteiro, A. L. and N. V. Bullock, *A broader view of communicative competence for aeronautical communications: Implication for teaching and high-stakes testing*. (in:) "The ESPecialist", 41(3, 1-29, 2020. ≤https://revistas.pucsp.br/index.php/esp/article/view/47361 ≥ (16.01.2023).

Pacheco, A. *English for Aviation: Guidelines for teaching and Introductory research*. 1. EDIPUCRS, 2019.

Park, M. Investigating Target Tasks, Task Phases, and Indigenous Criteria for Military Aviation English Assessment. *Language Assessment Quarterly*, v. 17, Issue 4, 2020. <https://www.tandfonline.com/doi/full/10.1080/15434303.2020.179922> (18.01.2023).

Petrashchuk, O. Aviation Radiotelephony discourse: an issue of safety, (in:) M. Grygiel (ed.), cognitive Approaches to specialist Languages. Cambridge, 114-134, 2017.

Santos, M. G., V.A Pacheco, R. R. Reyes, G. V. Vargas, *Analysis of the methodology applied in Aviation English course at a military school in Salinas.* (in:) "Journal of Education and Human Development", 39-44. 2018.<http://jehdnet.com/journals/jehd/Vol_7_No_3_September_2018/5.pdf>(16. 01.2023).

Silva, A. L. B. C. *Uma análise de necessidades de uso de inglês para os Oficiais Aviadores do Esquadrão de Demonstração Aérea (EDA) da Força Aérea Brasileira (FAB).* Master's Thesis in Applied Linguistics. State University of Campinas. Brazil. 2016. (URL <https://repositorio.unicamp.br/acervo/detalhe/973844>) [Accessed on 19.01.2023].

Silva, A. L. B. C. *Avaliação de proficiência em inglês para pilotos da Esquadrilha da Fumaça: Da análise de necessidades ao desenho de um exame*. (URL https://repositorio.unicamp.br/Acervo/Detalhe/1244211?returnUrl=%2FIndicador%2FIndex%3Fc%3D1244211) [Accessed on 19.01.2023].

Tosqui-Lucks, P. and M. A. Prado. *New routes in the study of Aviation and Aeronautical English*. (in:) "TheESPecialist" 41(4), 1-9. 2020. (URL https://revistas.pucsp.br/index.php/esp/article/view/51059) [Accessed on 16.01.2023]

Tosqui-Lucks, P. and A. L. B. C. Silva, *Aeronautical English: Investigating the nature of this specific language in search of new heights*. (in:)The ESPecialist, São Paulo, 41(3). 2020. <https://revistas.pucsp.br/index.php/esp/article/view/47826>. (16.01.2023].