



UNIVERSIDAD DE MURCIA
ESCUELA INTERNACIONAL DE DOCTORADO
TESIS DOCTORAL

Register and English for Specific Purposes pedagogy: English
for Military Navy submariners.

Registro y pedagogía del inglés para Fines Específicos:
Inglés para submarinistas de la Marina Militar.

D^a. Yolanda Noguera Díaz

2023



UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

Register and English for Specific Purposes pedagogy: English
for Military Navy submariners.

Registro y pedagogía del inglés para Fines Específicos:
Inglés para submarinistas de la Marina Militar.

Autor: D^a. Yolanda Noguera Díaz

Director: D. Pascual Pérez-Paredes

2023

*To my parents,
María y Alfonso.*



**DECLARACIÓN DE AUTORIA Y ORIGINALIDAD
DE LA TESIS PRESENTADA EN MODALIDAD DE COMPENDIO O ARTÍCULOS PARA
OBTENER EL TÍTULO DE DOCTOR**
Aprobado por la Comisión General de Doctorado el 19-10-2022

D./Dña. YOLANDA NOGUERA DIAZ

doctorando del Programa de Doctorado en

869 PROGRAMA DE DOCTORADO EN ARTES Y HUMANIDADES: BELLAS ARTES,
LITERATURA, LINGÜÍSTICA GENERAL E INGLESA.

de la Escuela Internacional de Doctorado de la Universidad Murcia, como autor/a de la tesis
presentada para la obtención del título de Doctor y titulada:

Register and English for Specific Purposes pedagogy: English for Military Navy submariners.
Registro y pedagogía del Inglés para Fines Específicos: Inglés para submarinistas de la Marina
Militar.

y dirigida por,

D./Dña. PASCUAL PÉREZ-PAREDES

D./Dña.

D./Dña.

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita, cuando se han utilizado sus resultados o publicaciones.

Además, al haber sido autorizada como compendio de publicaciones o, tal y como prevé el artículo 29.8 del reglamento, cuenta con:

- *La aceptación por escrito de los coautores de las publicaciones de que el doctorando las presente como parte de la tesis.*
- *En su caso, la renuncia por escrito de los coautores no doctores de dichos trabajos a presentarlos como parte de otras tesis doctorales en la Universidad de Murcia o en cualquier otra universidad.*

Del mismo modo, asumo ante la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada, en caso de plagio, de conformidad con el ordenamiento jurídico vigente.

En Murcia, a 28 de MARZO de 2023

Fdo.: Firmado por YOLANDA NOGUERA DIAZ -
NIF: ***9160** el día 28/03/2023 con un
certificado emitido por ACCVCA-120

Información básica sobre protección de sus datos personales aportados

Responsable:	Universidad de Murcia. Avenida teniente Flomesta, 5. Edificio de la Convalecencia. 30003; Murcia. Delegado de Protección de Datos: dpd@um.es
Legitimación:	La Universidad de Murcia se encuentra legitimada para el tratamiento de sus datos por ser necesario para el cumplimiento de una obligación legal aplicable al responsable del tratamiento. art. 6.1.c) del Reglamento General de Protección de Datos
Finalidad:	Gestionar su declaración de autoría y originalidad
Destinatarios:	No se prevén comunicaciones de datos
Derechos:	Los interesados pueden ejercer sus derechos de acceso, rectificación, cancelación, oposición, limitación del tratamiento, olvido y portabilidad a través del procedimiento establecido a tal efecto en el Registro Electrónico o mediante la presentación de la correspondiente solicitud en las Oficinas de Asistencia en Materia de Registro de la Universidad de Murcia

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	2
SUMMARY	4
INTRODUCTION	12
I. CONCEPTUAL FRAMEWORK	15
I.I. SPECIALIZED LANGUAGES AND ENGLISH FOR SPECIFIC PURPOSES (ESP)	15
I.II. ESP: MARITIME ENGLISH	19
I.III. CORPUS LINGUISTICS AND CORPORA: THE CMSC AND THE SAR	22
I.IV. CORPORA FOR TEACHING AND LEARNING: DATA-DRIVEN LEARNING (DDL)	25
I.V. MULTI-DIMENSIONAL ANALYSIS IN THE SUBMARINE ENGLISH REGISTER	28
II. STUDY 1	31
III. STUDY 2	33
IV. STUDY 3	34
V. OUTLINE OF RESEARCH QUESTIONS AND ANSWERS.	35
VI. CONTRIBUTION OF THIS PHD AND FUTURE RESEARCH.....	38
REFERENCES.....	44

ACKNOWLEDGEMENTS

Muchas gracias a mis guías de luz en el cielo y a mis padres en la tierra por todo y, por tanto.

Me siento agradecida a todos los que de alguna forma me han impulsado o apoyado en estos largos años que ha conllevado mi tesis, y también a los que me han puesto trabas, ya que todo forma parte de mi aprendizaje.

Es complicado resumir mis agradecimientos, por lo que me ceñiré al ámbito académico. La primera persona por cronología, es mi primera maestra, Doña María José Rosque Conesa, que supo plantar y abonar la semilla de la lectura y la curiosidad en una niña de 6 años.

La realización de una investigación de esta dimensión no puede ser el fruto de una persona sola que nunca antes había intentado escribir ciencia. Aunque la tarea de recopilación de datos, así como la presentación de este trabajo sea tarea de la autora, nunca hubiera llegado a buen puerto sin la tutorización de la profesora doña Purificación Sánchez-Hernández bajo la imprescindible batuta de dirección del profesor don Pascual Pérez-Paredes, al que le estoy agradecida por haberse aventurado a subirse en este mundo submarino para dirigir, corregir y matizar esta tesis.

Su capacidad de trabajo, honestidad, generosidad, perfeccionismo y talento están al alcance de muy pocos.

Las gráficas que aparecen en este trabajo tan didácticas, nacieron gracias al compañero informático de la UPCT don Israel Baños Rivas, que convierte en arte todo lo que pasa por sus manos y al que le agradezco su enorme paciencia y tan buen hacer.

En el ámbito militar, me gustaría agradecer in memoriam al Almirante García-Huix por confiar en mí y en mi trabajo, por abrirme puertas cuando otros me las intentaban cerrar. Un hombre grande por todas sus aristas.

No puedo olvidarme de quien al llegar a la Escuela de Submarinos de Cartagena, llamada 'Almirante García de los Reyes', me guio por los cometidos profesionales en mi periodo de enseñanza militar: el Capitán de Navío Suárez-Bárcena . Su entrega y pasión por la vida naval y militar fueron contagiosamente positivas, y que, junto con sus extensos conocimientos teóricos y vivencias en los Submarinos, supusieron un importante pilar en mi formación. Junto a él, el TeCol médico Cervantes cuyos relatos sobre su práctica médica a bordo de los estrechos submarinos convertidos en enfermerías de urgencia en un mar bastante oscuro además de sus invitaciones para la observación directa de los entrenamientos del Tanque de escape (clases en directo de supervivencia) me hacían reflexionar sobre esta forma especial de trabajar y estas gentes hechas, sin duda, con otra materia prima. Por último, y no menos relevante, agradecer también al Capitán de Navío Bernal González-Villegas ya que siempre conté con su importante apoyo logístico y moral.

AD UTUMQUE PARATUS (Lema del submarinista : siempre preparados)

SUMMARY

Este estudio profundiza en el análisis del Inglés para Submarinistas de la Marina Militar (Submarine English, SE), un área poco investigada del Inglés para Fines Específicos y Lenguas Especializadas. Utilizando la lingüística de corpus y la pedagogía basada en corpus, esta investigación aborda el análisis lingüístico de este registro, así como algunas aplicaciones pedagógicas relacionadas. Utilizando la gramática de constituyentes y el análisis multidimensional desarrollados por Douglas Biber, nuestra investigación identifica patrones de modificación y dimensiones subyacentes de uso en los dos corpus analizados. Además, un diseño semi-experimental de aprendizaje basado en datos (DDL) utiliza algunos de los datos del corpus con submarinistas profesionales de la Armada. El SE se caracteriza por una fuerte premodificación nominal, una baja premodificación adjetiva, una baja postmodificación de la frase preposicional y por el uso predominante de sustantivos apositivos en los huecos de postmodificación. El análisis multidimensional del registro analizado descubrió diferencias significativas entre los dos corpus. Estos resultados pueden ser útiles para la enseñanza profesional de idiomas, ya que proporcionan información sobre los patrones de variación en las distintas dimensiones de uso. Los patrones de variación convergentes y divergentes hallados en los dos corpus analizados constituyen un primer paso en la integración de la variación de registro en los currículos y materiales basados en corpus.

Este trabajo llena un vacío en el análisis de los lenguajes de especialidad en entornos militares y en la generación de conocimiento lingüístico que informe el diseño de materiales didácticos para submarinistas de la Armada basados en corpus. En muchos sentidos, esta tesis doctoral es única, ya que examina una aplicación del inglés militar que, hasta donde sabemos, no ha recibido atención académica hasta el momento. Esta tesis persigue promover la generación de conocimiento sobre las características formales del registro SE y fomentar la investigación en el área más amplia del inglés militar.

En el primer estudio de esta tesis por compendio se aborda el análisis del primer corpus recopilado en esta tesis doctoral: Cartagena Military Submarine Corpus (CMSC). La investigación previa en inglés marítimo (Maritime English, ME) no ha prestado atención hasta la fecha a la variedad de textos y lenguaje a los que están expuestos los submarinistas de la Armada durante su formación y carrera profesional. Este primer trabajo de investigación examinó los patrones de modificación de grupos nominales en un corpus de textos profesionales en el mencionado registro SE en el Cartagena Military Submarine Corpus (CMSC). Utilizando una combinación de análisis cuantitativo y cualitativo, descubrimos que el SE se caracteriza por un patrón de modificación único. Estos rasgos distintivos del registro SE exigen una pedagogía sensible al registro que desentrañe estas características y las presente en su contexto. Sostenemos que la contribución de la lingüística de corpus es esencial para explorar registros que, por diferentes razones, no han sido abordados o descritos lingüísticamente en el pasado. Del mismo modo, sostenemos que el análisis y la

enseñanza de los grupos nominales es esencial para comprender las tendencias actuales en la escritura y la comunicación profesionales.

Este tipo de postmodificadores es infrecuente en otros registros merece la atención de profesores y desarrolladores de materiales pedagógicos. La investigación en el área de la complejidad sintáctica ha prestado atención al grupo nominal principalmente en inglés académico (Biber & Gray, 2016) e inglés para aprendices (i.e. Díez-Bedmar & Pérez-Paredes, 2020). Sin embargo, la exploración de los grupos nominales en registros especializados y profesionales como el SE ha recibido menos atención. Es posible que esta falta de interés se deba al grado de especialización y especificidad del género analizado y, por tanto, a la falta de masa crítica en este ámbito. Un análisis de una gama más amplia de grupos nominales en los dos corpus resultará útil para profundizar en este registro de especialidad y fundamentar un currículo especializado basado en el análisis de corpus lingüísticos. Si bien nuestros hallazgos son de interés, los submarinistas militares deben familiarizarse con un número más amplio de rasgos léxico-gramaticales relevantes para comprender cómo se utiliza su registro profesional. Los trabajos futuros deberían abordar una gama más amplia de rasgos lexicogramaticales.

En el segundo estudio, se investigó las reacciones y actitudes de dieciséis submarinistas profesionales en formación ante tareas de Aprendizaje Basado en Datos (DDL) que exploraban el uso de acrónimos en un corpus especializado. Una gran cantidad de estudios en la última década ha apoyado el uso de la lingüística de corpus en la enseñanza de idiomas (Boulton y Cobb 2017; Pérez-Paredes 2019), pero solo algunos de ellos han combinado los métodos de la lingüística de corpus, DDL y ESP.

Hace más de quince años que Gavioli (2005) aplicó el uso de DDL para enseñar el lenguaje disciplinar y mejorar las experiencias autónomas de aprendizaje de idiomas de los estudiantes de medicina. La investigación en este campo, sin embargo, no parece haber avanzado significativamente (Pérez-Paredes 2019). La mayoría de los investigadores parecen estar de acuerdo en que, como señalan Crosthwaite y Cheung (2019: 20), el uso de DDL expone a los estudiantes de idiomas a evidencia lingüística que les permita comprender los rasgos lingüísticos característicos implicados en la producción de géneros disciplinares, mejorando así su comprensión de las complejidades de la alfabetización dentro de su campo disciplinar. La literatura especializada ha señalado que, a través de la interacción de los aprendices con materiales basados en corpus, los estudiantes identifican características lingüísticas particulares, las prácticas discursivas y las habilidades comunicativas de la comunidad profesional a la que pertenecen.

En este segundo estudio se recopiló y analizó un segundo corpus de Salvamento y Rescate de Submarinos en el que el uso de acrónimos, área muy descuidada en el inglés con fines específicos (ESP), desempeña un papel sustancial. Utilizando una metodología mixta, esta investigación examinó, por primera vez en una Escuela Militar de la Armada Española, la interacción de los estudiantes con las tareas, así como las entrevistas posteriores con los alumnos.

La postmodificación aposicional motivó la selección de acrónimos en el presente segundo estudio. En este trabajo utilizamos el segundo corpus que integra esta tesis doctoral, el corpus de Salvamento y Rescate de Submarinos (SAR), usado para generar una selección de actividades DDL en la enseñanza de acrónimos de

especialidad. Mediante el diseño de investigación, perseguimos saber cómo los profesionales militares abordaban el aprendizaje de una lengua a través de un enfoque DDL. Nuestro objetivo lingüístico no era sólo contribuir con una enseñanza informada por la investigación sobre los acrónimos de especialidad, sino introducir a los estudiantes por primera vez (se trata de un curso de posgrado) en un enfoque informado por corpus para el aprendizaje del inglés submarino a través de líneas de concordancia en contextos profesionales y militares. Esta es en sí misma una contribución importante de esta tesis doctoral ya que, hasta donde sabemos, DDL nunca se había utilizado con las Fuerzas Armadas y muy raramente fuera de la educación terciaria o secundaria. Se utilizaron tareas DDL deductivas e inductivas con líneas de concordancia de siglas con dos grupos de estudiantes de diferente rango militar. Ambos grupos encontraron las tareas desafiantes y mostraron opiniones encontradas hacia la explotación de las líneas de concordancia en el aula. Este trabajo supone una novedad al trasladar DDL a aulas no universitarias e interactuar con estudiantes más maduros en estudio profesionales no académicos. En este trabajo se argumenta que las conversaciones con los estudiantes adultos sobre la naturaleza del aprendizaje de lenguas, en general, y el uso de líneas de concordancia y DDL, en particular, son absolutamente esenciales para promover un papel de estudiante más activo.

En el tercer artículo de esta tesis, se exploran las diferencias y las similitudes encontradas en dos corpus representativos de dos registros de relevancia para los submarinistas de la Armada Española. En una serie de análisis estadísticos basados en el análisis multidimensional, se discute cómo este enfoque puede informar la

enseñanza profesional de idiomas en tales contextos. Argumentamos que la variación lingüística entre los textos dentro de los registros permite identificar patrones de variación tanto convergentes como divergentes a través de dimensiones de uso identificadas mediante análisis multidimensional.

El análisis de la variación lingüística mediante el análisis multidimensional (AD) no ha sido ampliamente adoptado en la enseñanza de lenguas profesionales y especializadas fuera de contextos universitarios. Dada la falta de investigación en el registro SE, y con la intención de familiarizar a investigadores y profesores de idiomas con la naturaleza lingüística de los textos de nuestros dos corpus, estudiamos la variación en ambos corpus siguiendo el enfoque de análisis multidimensional de Biber (1988, 2019). Aunque los dos corpus usados en esta investigación representan dos tipos de texto de interés para los submarinistas de la Armada (revistas profesionales y manuales de salvamento y rescate), ambos comparten un dominio temático y una comunidad. Los resultados cuantitativos y cualitativos del análisis del registro SE nos ofrecen una imagen compleja de las características lingüísticas y funcionales de estos registros, además de ofrecer a los instructores la posibilidad de utilizar estos datos. Las diferencias lingüísticas entre las Dimensiones tienen un gran impacto en la enseñanza de idiomas. Ambos corpus muestran una puntuación alta en sustantivos y pasivas sin agente (D1 y D4), pero divergen en D2, 3 y 4, lo que significa que muestran diferencias de registro que tienden a ignorarse en las aplicaciones de investigación. También hay diferencias en el análisis textual inter-corpus e intra-corpus que son útiles y pedagógicas. Una contribución importante del Estudio 3 es la discusión sobre las orientaciones convergentes y divergentes en el análisis multidimensional y su

aplicación en la enseñanza de lenguas especializadas. La comprensión de la variación lingüística en la SE implica una inspección minuciosa en dos niveles diferentes, tanto de los textos individuales como en el nivel superior de los corpus. Los patrones convergentes de variación se centran en textos que se comportan de manera similar en una determinada dimensión de uso. Los patrones de variación divergentes muestran cómo los dos corpus presentan frecuencias y distribuciones de rasgos lingüísticos que permiten interpretaciones funcionales divergentes en una dimensión determinada. en un diseño semi-experimental para la enseñanza del inglés militar. En términos sencillos, si quisiéramos subrayar la enseñanza de rasgos lingüísticos asociados al tiempo verbal pasado, los pronombres de tercera persona o el uso del aspecto perfecto, los textos más adecuados se pueden encontrar en el CMSC, ya que este corpus muestra una orientación más narrativa. Si, por el contrario, buscamos trabajar con oraciones de relativo, subordinación condicional y, por ejemplo, modales, el corpus de rescate y salvamento puede proporcionarnos textos adecuados a este perfil, ya que este corpus es más informativo y persuasivo. Como hemos señalado en este trabajo, la investigación especializada ha tendido a centrarse en las diferencias. Sin embargo, en el análisis lingüístico profesional y especializado, el estudio de los patrones convergentes de variación puede repercutir en la evaluación de los textos que pueden informar a un enfoque pedagógico basado en corpus. Este extremo debe seguir explorándose en futuros estudios. Otros trabajos podrían retomar nuestro análisis de MD y aplicar el marco convergente/divergente sugerido en esta tesis

Las lenguas militares profesionales no han recibido suficiente atención en la literatura de la ESP y, en general, en la lingüística aplicada. Conrad (2019) afirma que

el análisis de registro identifica eficazmente las cualidades que están presentes en todos los textos y que son ubicuas en un grupo de textos o corpus. Para ella, "el análisis de registro es útil porque los hallazgos son directamente aplicables a la enseñanza, que es un objetivo en muchos estudios de registro ESP [y] los hallazgos se utilizan en la creación de nuevos materiales didácticos que integran el desarrollo de habilidades de escritura con el contenido y la práctica [ESP]" (p. 190). Esta tesis ha seguido la sugerencia de Conrad de que el campo de ESP puede recurrir al análisis de corpus para analizar y comprender el uso especializado de la lengua. Hemos abordado un escenario específico de uso en el que se han explorado el contexto de producción y recepción, una amplia gama de rasgos lingüísticos y las relaciones funcionales entre las cualidades situacionales y los rasgos lingüísticos. Combinado con el análisis de los rasgos frasales y las actividades de clase, esta tesis ofrece, hasta donde sabemos, la primera aplicación del DDL inglés-militar y el primer análisis del SE utilizando una perspectiva de registro del MD.

CXXIX

ADRIFT! A little boat adrift!
And night is coming down!
Will no one guide a little boat
Unto the nearest town? (...)

Emily Dickinson (1830–86).

INTRODUCTION

This research work is the result of a personal and professional admiration for the world of the sea and its people.

Working in a military base gives you certain unspoken and subtle lessons that provide you with a different vision. My period of work at the Cartagena Submarine School, called Almirante García de los Reyes, for more than five years, was a before and after as a lecturer of English as a foreign language.

I started teaching academic English based on the current European profiles (A2, B1, B2, and C1) although listening to the doubts that emerged in class, which had nothing to do with that programme, made me think that something had to be modified. Therefore, I started to combine academic English lessons with standard maritime vocabulary regulated by the International Maritime Organization (IMO). Then, I began the phase of teaching English for Specific Purposes (ESP) to Officers, Petty Officers

and sailors along with general academic English. We can assume it was the beginning of Maritime English (ME) teaching in my lessons.

As I progressed with these classes, I realized the surrounding vessels did not match the descriptions we made in the lessons. It was necessary to change the target domain. We had to concentrate on a different type of vessel with unique and distinctive characteristics: submarines.

In the search of didactic material related to Submarine English (SE) I realised the scarcity or, indeed, the non-existence of materials. This was an empty niche.

Therefore, there was a first stage in which I started to prepare the lessons with the manuals, submarine naval engineering English books from the Base library and I collected articles from the specialised magazines found there.

In a second stage, I expanded the search in the Submarine School's curriculum project for subjects that were taught in an unrestricted way in order to be able to consider a Content and Language Integrated Learning (CLIL) and include some subjects that I could teach transversally, such as Submarine Salvage and Rescue. To do this, I consulted Spanish naval sources and some expert commanders of the Submarine School. I wanted to provide my military students with appropriate and authentic materials and to design a good ESP course on Submarine English (SE). The didactic planning and materials were developed, but still lacked consistency.

This search for a scientific basis led to the birth of a professional objective and an academic target. This was the third stage: the search of an academic approach,

linguistic and pedagogical scientific entity to the register of Submarine English as a branch of Maritime English.

To carry out this third stage, we framed our research within the so-called specialised languages, Corpus Linguistics (CL), Data-Driven approach (DDL) and MDA (Multi-Dimensional Analysis). It is, therefore, an interdisciplinary work .

This is a PhD by publication. It is divided into six chapters.

- **Chapter 1** offers the theoretical framework of the study. We pay attention to the theories and approaches that have informed our work.
- **Chapter 2** is Study 1: Noguera-Díaz, Y. & Pérez-Paredes, P. (2019). Register analysis and English for Specific Purposes (ESP) pedagogy: noun-phrase modification in a corpus of English for Military Navy submariners. *English for Specific Purposes*, 53,118-130.
<https://doi.org/10.1016/j.esp.2018.06.002>
- **Chapter 3** is Study 2: Noguera-Díaz, Y., & Pérez-Paredes, P. (2020). Teaching acronyms to the military: A paper-based DDL approach. *Research in Corpus Linguistics*, 8(2), 1-27. <https://doi.org/10.32714/ricl.08.02.01>
- **Chapter 4** is Study 3: Noguera- Díaz, Y. (2022). A multidimensional analysis of two registers of English for Navy submariners. *Language Value*, 15(2), 81-111. <https://doi.org/10.6035/languagev.7019>
- **Chapter 5** offers a brief account of the main findings in the three studies.
- **Chapter 6** discusses the global contribution of this PhD to applied linguistics and future research, which may stem from this work

I. CONCEPTUAL FRAMEWORK

This study examines English for Military Navy Submariners (SE), an under researched area in English for Specific Purposes and Specialized Languages. Using corpus linguistics and corpus-based pedagogy, this PhD addresses a linguistic analysis of the register and related pedagogic applications.

This is a PhD by publication. Its main chapters are integrated by the three studies published in *English for Specific Purposes*, *Research in Corpus Linguistics* and *Language Value*.

This chapter offers the reader a conceptual framework in which to situate this PhD thesis. To this end, we will review some of the areas of knowledge in which to frame some of the research in specialised languages by means of corpus linguistics that has inspired our approach. Far from an extensive review of the literature used in other types of doctoral thesis (i.e. these others than a thesis by publication), we want to situate the present research work conceptually in three areas, which we will discuss below: specialised languages, corpus linguistics and corpus-based language pedagogy.

I.I. Specialized Languages and English for Specific Purposes (ESP)

The study of Specialized Languages (SL) as a branch of applied linguistics dates back to the 1960s. As Swales (1985) points out, it is the year 1962, which marks “the beginning of the story” with Barber's (1962) article “Some measurable characteristics of modern scientific prose”. However, it was Halliday, Strevens and McIntosh (1964)

with the publication of their book *The Linguistic Sciences and Language Teaching* who established the theoretical foundations of languages for specific purposes and SL as disciplines in the field of Applied Linguistics.

The analysis of specialised languages can potentially encompass any language used in the world, although, given the strength and importance of the English language in international contexts of specialised and scientific communication (Curry, 2021), most of the research in this area has been done on the English language (Bathia, Sánchez-Hernández & Pérez-Paredes, 2012). This fact may be interpreted as a drawback when it comes to articulating a comparative framework with other languages at the same level.

Today, the study of Specialized Languages is considered as an independent and multidisciplinary discipline (Bathia, Sánchez-Hernández & Pérez-Paredes, 2012). SL has its own working and research methodology, although it draws inspiration from other disciplines ranging from corpus-based research and genre analysis to ethnography of communication and communication studies.

One of the challenges facing the novice researcher in this field is terminology. The terms SL, Languages for Specific Purposes (LSP) and ESP are even interchangeable (Belcher, 2009). Some authors overlap their uses, and when they refer to LSP they are referring to ESP. As Belcher (2009:2) pointed out, 'there are as many types of ESP as there are specific learner needs and target communities in which learners wish to develop'.

ESP has traditionally been divided into two main branches: English for Academic Purposes and English for Occupational Purposes (EOP) (Hutchinson and Waters, 1987; Dudley-Evans and St Johns, 1998; Belcher, 2009). EAP focuses on the academic skills in formal settings at primary, secondary and tertiary levels. On the contrary, EOP is a field of language learning and teaching in which teaching focuses on the learner occupational needs instead of on the learner's academia (Anthony, 2018).

Upton (2012) reflects on LSP studies encompassing five decades (1960-2010). Its theoretical framework focuses mainly on needs analysis and language analysis. His theoretical framework focuses mainly on needs analysis and language analysis. In the area of Needs Analysis highlights the study by Dudley-Evans and St John (1998:4). They gave a referential definition of ESP, including a specific one on learner-centered focus: "ESP is designed to meet specific needs of the learner". In this view, "ESP is centered on the language appropriate to these activities in terms of grammar, lexis, register, study skills, discourse and genre" (p.4).

Belcher (2009:3) repeats the same perspective on prioritising the students' needs: "first and foremost (before, during and even after instruction), finding out what learners 'needs are."

In our work, the learners' needs, language analysis, and the use of specific materials go hand in hand to form a unit of action that guides our interest in the use of linguistic corpora and authentic specialised texts.

In the area of LSP language analysis, Upton (2012) emphasises the contributions of the vocabulary and structures used in specialised contexts, known as "register analysis", led by Halliday, McIntosh, and Strevens (1964), which are the precursors of the register analysis later inherited by Biber (1988).

There has been a huge number of studies on LSP/ ESP from the 1960s to the present day, each with different linguistic perspectives. Hyland and Jiang (2021) highlight 2013 as the beginning of ESP's modern age, a term coined by Johns (2013) to indicate the updating as well as the inclusion of concepts and methodology.

Hyland and Jiang (2021) have studied the most prominent ESP issues through bibliometric techniques by analysing 3500 papers on the social science citation index in the period 1990-2020. This work is extremely important because, by looking at the most studied areas and the most cited authors, we can recognise the evolution of the ESP discipline, and the influence it may have had on our own study. Hyland and Jiang (2021) argue that in the period 2005-2020, two researchers stand out as having made an exceptional and continuous contribution to the research and development of the field in question.

The first one is Bathia (2012) who advocates for a multidimensional and multi-perspective methodological framework to understand and analyse professional and academic genres, professional practices, and disciplinary and professional cultures as comprehensively as possible. The second one is Biber (1988) and Biber and Conrad (2009), who offer us the definition of register that we have used in our study applied to

the study of specialised registers (Biber et al. 1999) using corpus linguistics and a quantitative approach:

Register variation focuses on the pervasive patterns of linguistic variation across such situations in association with the functions served by linguistic features. Genre variation focuses on the conventional ways in which complete texts of different types are structured. (Biber and Conrad 2009:35).

I.II. ESP: Maritime English

The first register we review in this study under the framework of ESP is Maritime English (ME). We consider ME as a holistic register because of the number of branches into which it can be divided for its study according to the situational context. Cole, Pritchard, and Trenker (2007) framed it as a mixture of different domains depending on the field or topic (nautical, logistics, technical, legal, etc.), which makes it a complex register to study. The most influential definition of the term was provided by Trenker (2000:7) as it encompasses the entire maritime community that must communicate in English on board or in port:

ME is the entirety of all those means of the English language, which, being advisable for communication within the international maritime community, contribute to the safety of navigation and the facilitation of seaborne business.

In Spain, Bocanegra-Valle (2013:1) offered an interesting definition of ME, although due to the extent of unidentifiable branches, she did not mention sections such as sea tourism, naval engineering or marine communication, which are also considered important. For her, ME includes “five different sub varieties according to the specific

purpose they serve in the maritime context: English for navigation and maritime communications, English for maritime commerce, English for maritime law, English for marine engineering, and English for shipbuilding”. As the reader can appreciate, military applications of ME are not included.

The governing role of the International Maritime Organization was key as it officially established English as the official language of communication on ships and in marine environments (STCW78/1995 Convention and Code) in order to avoid accidents and failures by multicultural crews. That is, there was no debate in the selection of the lingua franca to be used in these international communication contexts. English became the fundamental language of international communication in the development of international maritime transit and trade. With the publication and approval of the Standard Maritime Communication Phrases (SMCPs 2001), the use of these standard phrases is regulated, and their use became mandatory for officers in charge of a navigational watch and radio officers. These International Maritime Organization (IMO) Standard Phrases are used ship-to-ship, ship-to-shore, and in shipboard communications in the civilian world. Despite the wide range they cover, they do not deal with lexical fields, which are also important in this register of ME, such as shipbuilding, marine engineering, meteorology, and maritime legislation on board and in port, and, in our case, ME used by the military in submarines. We highlight the IMO Model Courses, which were developed at the suggestion of several governments who requested such guidance courses, including Model Course 3.17: Maritime English. Although we are aware that a single ME syllabus does not cover all the needs of different trainees, at least this course offers a minimum content established and

regulated in the civilian world for the different maritime learning contexts (Vocational Schools, Universities, and Private Courses). As we can observe, the ME field has mainly dealt with civilian uses and communication contexts.

In the context of the Military Navy, there is a variety of textbooks and materials that are focused on meeting the needs of military students learning English as a foreign language as well as general Maritime English. NATO (North Atlantic Treaty Organization) publishes and regularly revises a Glossary of Military Terms and Definitions (AAP-6) in addition to the descriptive profiles Standard Nato Agreement (STANAG 6001). They are similar to the European framework established by the Common European Framework for Reference (CEFR) under the need to communicate among the different NATO countries whose L1 is not English.

To align with the civilian world, level descriptors were created in the Standard NATO Agreement 6001 (STANAG 2019). STANAG includes five levels ranging from level 1 (survival), 2 (functional), 3 (professional), 4 (expert) to 5 (native). These levels include various productive and receptive skills in listening, speaking, reading, writing and use of English. Orna-Montesinos (2016), for example, created a series of teaching materials based on an analysis of the needs of her students at the Spanish Military Academy of Zaragoza (Spanish Army) using these guidelines. The design of the course is based on intercultural communication as a transversal topic to overcome the challenges and barriers that cadets may encounter in the development of their profession. In general, current military English textbooks do not cater to particular armed forces curricula (i.e. Navy submariners), but combine various international military contexts under shared themes, such as humanitarian aid, peacekeeping

missions or manoeuvres (Campaign ed English for the Military series). These publications can be used for land, sea or air forces.

Nowadays, the Spanish Navy Submarine Warfare School trains future S-80 class submariners through a year-long programme according to their military category, which includes the subject of English as a foreign language, as these students carry out international manoeuvres, meetings and conferences under the NATO umbrella. It is in this Naval Military School where we concentrate our attention.

I.III. Corpus Linguistics and corpora: the CMSC and the SAR

Corpus linguistics (CL) has been used as a source of linguistic reference and a tool in the teaching-learning process of specialised languages. For some authors, the use of corpus linguistics is intrinsically associated with the study of ESP (Anthony, 2018). For Conrad (2019), corpus-based analyses have become widespread and valued in recent years for all areas in ESP. The corpus-informed approach in specialised registers encompasses the collection and analysis of texts (oral and/or written) to describe the nature of a register. The qualitative dimension of linguistic description gives it an evidence and credibility that is difficult to overcome with any other methodology. In the following lines, we offer a very brief theoretical overview of CL and an introduction to the corpora we put together for the analysis and the classroom exploitation of Submarine English (SE).

Sinclair (2004:31) defines a corpus as a collection of texts in electronic form, “selected according to external criteria to represent a language or language variety as a source of data for linguistic research”. A corpus is thus a collection of (1) machine

readable (2) authentic texts, which is (3) sampled to be (4) representative of a particular language or language variety (McEnery, Xiao & Tono, 2006:5). Tognini-Bonelli (2001:55) described corpora as computerized collections of authentic texts, amenable to automatic or semiautomatic processing or analysis. The texts are selected in order to capture the regularities of a language, a language variety or a sub-language. Tognini-Bonelli also highlights three important aspects in her approach to corpus linguistics that we embrace in our analysis: authenticity, representativeness and sampling. Authenticity concerns all linguistic material included in the corpus, whether written or oral, as in our study. This aspect must be clearly reflected when the data collection design is approached, otherwise one of the most important features of corpus creation would be nullified. The representativeness of the register under study is linked to the range of texts that represent the observed linguistic phenomena, ranging from top-level representations of very broad registers as in Biber et al. (1999) to very specific registers as in our own study (see the corpora used in the three research papers).

As for the sampling of the representative texts, the controversy about the importance of collecting a large corpus in order to address a wide range of target audience (i.e. readers or users of the language) is of limited importance in the field of specialised corpora such as ours (Gavioli, 2002). However, it is essential that we acknowledge the limitations in using corpora in linguistic research. For Egbert, Biber and Gray (2022:5), the corpus is just an artefact created by the linguist. The corpus is, essentially, designed to represent the targeted domain of language use, in our case Submarine English. The emphasis on the representativeness of the linguistic data included in the corpus becomes, therefore, essential. To do this, we must describe the

area of interest/ domain of our study. As suggested by Egbert et al. (2022), we should follow a three-step process from the general description of the domain (Egbert et al.2022:73), understood as the full universe of language use a researcher wants to learn about, through an operational domain (the texts actually available for inclusion in the corpus) and the corpus as such resulting from the final selection. In our case, the restricted or secret character of most of the texts belonging to the SE domain limited the area, making the operational domain and the corpus sample coincide without a biased selection or stratification. For Reppen (2010), the size of the corpus must be determined by two factors: representativeness and practicality. We must make sure that we have collected enough texts to be able to represent the researched language and to be able to answer the research questions.

In this study, we embrace the theoretical division established by Tognini-Bonelli (2001) between corpus-based and corpus-driven approaches. In corpus-based approaches, the corpus itself is used as a reference to exemplify, for example, the rules or linguistic phenomena to be collated, while in the corpus-driven approach, the corpus is the linguistic source that provides and unveils the linguistic elements to be investigated.

In this PhD, the Cartagena Military Submarine Corpus (CMSC) and the Salvage and Rescue of Submarines Corpus (SAR) are more than reference materials or sources. This is why our linguistic perspective follows, the parameters of a corpus-driven approach, as we build up our analysis of the SE register on the basis of the linguistics evidence from the two corpora (see Study 1 and 3).

The Cartagena Military Submarine Corpus (CMSC) is made up 822,755 words, 952 texts and comprises 12 years of curated professional specialised journals with specific articles about the field of submarines, which we will specify later. In the search for another sample or corpus to broaden the focus of the operational domain, we focused on a compulsory subject that our military students and future submariners must study as part of their academic training, and where, due to its transversal and international nature, the access to the texts was not restricted. The Salvage and Rescue of Submarines Corpus (SAR) contains 37,615 types and 717,446 tokens from 18 unclassified NATO books and manuals that are consulted and read by these students and trainers.

As for the size of our corpora, we can consider them small in size. According to O'Keeffe (2007), there are positive aspects to the small corpus approach since the compiler and researcher is usually the same person (as in our case), which helps the quantitative analysis of the corpus and the qualitative analysis of the results. Similarly, Flowerdew (2004) also highlights that there is no ideal corpus size, valuing in particular the small-specialised corpus that focuses on specific registers and is connected to teaching and its use in the classroom.

I.IV.Corpora for teaching and learning: Data-driven learning (DDL)

The main purpose behind the creation of our corpora has been the linguistic analysis of Submarine English (see Study 1 and 3). With the SAR corpus, however, we have used corpus tools to bring authentic language data into the classroom (see Study 2). It is widely accepted that corpora are effective when researching vocabulary in ESP

and SL (Coxhead et al., 2016; Anthony, 2018). Similarly, it has been shown that the use of data-driven learning is effective for the acquisition of a wide range of skills and linguistic knowledge such as vocabulary and register-specific features (Boulton & Cobb, 2017; Pérez-Paredes, 2022). In our case, based on our analysis of modification patterns in Study 1, acronyms turned to be essential to express reference in the Noun Phrase. One of our main findings in Study 1, and to some extent in Study 3, is that modifications patterns are unique to SE. Although some studies have developed lists of technical words, mainly in engineering, medicine, nursing, law, aviation, business, and chemistry (Coxhead, 2023) for their use in the teaching/learning process, to the best of our knowledge our DDL approach is unique in the field of SE, an area where the mastery of technical vocabulary in context is essential as a sign of community identity (Parkinson and Mackay, 2016).

To get into the classroom proper, we used Data-Driven Learning (DDL), a term coined by Tim Johns (1991:211) to describe the use in the classroom of “computer-generated concordances to get students to explore the regularities of patterning in the target language”. Years later, the definition is still being adapted for the same teaching purposes. DDL exploits the tools and techniques of corpus linguistics for pedagogical purposes (Gilquin and Granger, 2010:359) which presupposes an investigation and application of the use of corpus data for use in the classroom (O’Keeffe, 2021). DDL allows the learner and the teacher through a concordancer software, or through the examination of paper based concordance lines, to observe, read, predict, explore, and examine keywords in context (KWIC), frequency lists, n-grams and collocations in the corpus. DDL can be hands-on, whereby learners access the corpus or activities directly

via their computers or it can be hands-off, whereby the teacher prints out the worksheets (Boulton, 2012). This was our approach in Study 2.

It is a fact that most studies on DDL are in tertiary education (Boulton and Cobb, 2017), and that the teacher may find it useful to stagger activities, even if this involves learning to use corpora (Lenko-Szymanska, 2017). The problem lies in the belief of corpus integration in the researcher-teacher tandem or in teacher training programmes, as Farr (2021) states is currently happening. As an example of studies in ESP that confirm the good results of DDL, we have Friginal and Roberts (2018) in Aviation English, where the researchers developed materials for pilots and air traffic controllers, emphasizing the improvement of the linguistic level that they obtain through corpus studies in their area and invite research for individual job roles within a single industry (aviation, oil gas, shipping industry, etc...). This finding is in line with our own study of military navy submariners.

Despite the challenges in using corpus data in classroom contexts (Pérez-Paredes, 2022), Boulton and Vyatkina (2021:89) confirm that DDL works and the use of DDL is growing, and not only in Europe, but all over the world across levels and devices (Boulton and Pérez-Paredes, 2023). In Study 2, we used mixed-methods methodology with military submarine students through paper-based tasks with concordance lines of selected acronyms with two groups of students together with a series of subsequent interviews. Although Boulton (2012) recommends corpus tools in ESP over the traditional methods of textbooks, grammars, etc., our students showed curiosity and some general amazement with these totally new technologies, feeling more comfortable, however, with traditional methods. In our results, we can only agree with

Boulton and Pérez-Paredes' (2023) tips for teaching with DDL, in terms of reducing the time of use, familiarising students with Google tools, and capturing attention by making DDL interaction and engagement as simple and straightforward as possible.

I.V. Multi-Dimensional Analysis in the Submarine English register

Research in specialized languages offers classroom applications that include the design, compilation, and pedagogic exploitation of lexical features in specialized sub-genres (Albi et al., 2014), the study of move structure, lexical bundles (Cortes, 2006), and text patterning (Flowerdew, 2011) , as well as the analysis and the development of corpus materials for teaching of vocabulary in professional domains such as aviation (Bocorny, 2010) or dentistry (Crosthwaite and Cheung, 2019). However, there is a dearth of research that examines the potential of linguistic variation analysis as a contributing factor to the development of pedagogical approaches to teaching specialized languages in professional contexts (Biber, 1998; Ren & Lu, 2021).

In order to qualify and analyse the two corpora of the Submarine English (SE) register in great detail and reach particular conclusions, we used Multidimensional Analysis (MD) (Biber, 1988). In our study, five main dimensions of language use were identified in order to analyse the linguistic variation present in the two corpora. MD allows discourse/registers to be described quantitatively and qualitatively. Quantitatively, the dimensions quantify the extent to which the features associated with the dimension are used, and at the same time are based on frequencies of co-occurring features. Functionally, on the other hand, the dimensions are interpreted based on the analysis of shared functions of features, text fragments and register distribution.

Dimension 1 interprets involved vs informational production, Dimension 2, narrative vs non-narrative concerns, Dimension 3 explicit vs situation dependent reference, Dimension 4 overt expression of persuasion, and Dimension 5 abstract vs non-abstract information (Biber, 1988; Biber, 2019). These five dimensions of language use were calculated and each was given a factor score, which is a numerical value indicating the relative position of a text on an inherent factor in the factor analysis. For this purpose, the multi-Dimensional tagger (MAT), created by Nini, 2019, was used. According to Nini (2019:9), the dimension scores produced by MAT are reliable as they are based on Biber's (1988) work. Each dimension of use "comprises a distinct set of co-occurring linguistic features, and each has distinct functional underpinnings" (Biber, Reppen and Conrad 2002:459). The analysis of linguistic variation across registers (Biber, 2008;2019) can inform future decisions regarding the selection and use of corpus-based materials (Biber and Conrad, 2009) in professional environments.

MDA has been used in workplaces and academic English as a first step to understand how linguistic variation operates across domains and registers. In academic English, Gray (2013) looked at research articles in a corpus of high-quality journals in history, applied linguistics, philosophy, chemistry, political science, and physics. She found high levels of variation in the linguistic features across the disciplines analysed. Conrad (2017) examined the linguistic gap in civil engineering students' writing skills and professional writing in the workplace. Conrad found an overuse of some linguistics features such as the passive voice in the students' corpus. In other languages such as Spanish, there has been some interest in the use of MDA.

The combination of a specialized corpus, MD analysis and the development of data-driven learning (DDL) activities has been mainly examined in Cheung (2018) and Crosthwaite et al (2019). Both studies argue that, for the teaching of professional language, the distribution and frequencies of linguistic features in the corpora analysed needs to inform the elaboration of DDL tasks in the syllabus design of undergraduate Hong Kong university students of English for Dentistry. Other researchers have used MD for assessment. Lynch et al (2020) developed their own Aviation English Test, which screens and assesses incoming non-native English speaking flight students to their US aviation center. They claim to provide an effective language tool to focus on the candidates 'actual oral English skills separately from the other written skills. This emphasis on spoken language is also found in Pérez-Paredes & Sánchez-Tornell (2019), who performed a MD of a corpus of native speakers of English fulfilling RFL tasks like those found in oral performance interviews (OPIs). Their analysis of the underlying dimensions informs language assessment professionals and language teachers about the most prevalent linguistic features and communicative affordances of such tasks.

II. STUDY 1

Noguera-Díaz, Y., & Pérez-Paredes, P. (2019). Register analysis and ESP pedagogy: Noun-phrase modification in a corpus of English for military navy submariners. *English for Specific Purposes*, 53, 118-130.

This work was published in *English for Specific Purpose*, published by Elsevier. According to the Social Sciences Citation Index (SSCI), it is the 37th journal out of 195 in the area of linguistics, Q1 and percentile 81.28 in the aforementioned area in the year 2021. It has an impact factor of 2.417, with a 5-year impact factor of 3.191. The journal publishes world-class research on second language acquisition in specialised contexts, needs assessment, curriculum development and evaluation, discourse analysis, descriptions of specialised varieties of English, teaching and assessment techniques, effectiveness of various approaches to learning and teaching specialised languages, and teacher training for teaching English for specific purposes. To date, the work has been cited ten times in various journals in the field, most notably in Professor Gray's chapter entitled "What can a corpus tell us about registers and genres?" in the second edition of *The Routledge Handbook of Corpus Linguistics*, published by Routledge. B. Gray's work on academic genres and nominal group together with Douglas Biber is internationally recognised.

Research in Maritime English (ME) has paid no attention to the range of texts and language which Navy submariners are exposed during their training and professional careers. This research looked at Noun Phrase modification patterns in a longitudinal corpus of Submarine English (SE) professional texts in the Cartagena Military Submarine Corpus (CMSC). Using a combination of quantitative and qualitative analyses, we found that SE is characterised by heavy nominal premodification, low adjective premodification, low prepositional phrase postmodification and by the predominant use of appositive nouns in postmodifying slots. These distinctive features of SE call for a register-sensitive pedagogy that unpack these characteristics and present them in context. We argue that the contribution of corpus-linguistics is essential to explore registers which, for different reasons, have not been addressed or described linguistically in the past. Similarly, we maintain that the examination and teaching of NPs is essential to understand current trends in professional writing and communication.

III. STUDY 2

Noguera-Díaz, Y., & Pérez-Paredes, P. (2020). Teaching acronyms to the military: A paper-based DDL approach. *Research in Corpus Linguistics*, 8(2), 1-27.

Dialnet metrics: This journal is in 2021 in the first quartile with an impact index of 0.28. The position of this journal is third place out of 53. This article has been cited by five scholars .

This study investigated the reactions and attitudes of sixteen military submariner students towards Data-driven learning (DDL) tasks that explored the use of acronyms in a specialised corpus. The researchers collected and analysed a corpus of Salvage and Rescue of Submarines where the patterning of acronyms, neglected in English for Specific Purposes (ESP), play a substantial role. Using a mixed-methods methodology, this research looked, for the first time in a Military School, at the students' interaction with the tasks as well as the subsequent interviews with the students. Deductive and inductive paper-based DDL tasks with concordance lines of acronyms were used with two groups of students of different rank. Both groups found the tasks challenging and showed mixed emotions towards concordance lines. We suggest that conversations with adult students about the nature of instructed language learning and language patterning are absolutely essential to promote a more active learner role.

IV. STUDY 3

Noguera-Díaz, Y. (2022). A multidimensional analysis of two registers of English for Navy submariners. *Language Value*, 15(2), 81-111

Language Value Journal is in the first Quartile in the first position of Journals in Scopus and Dialnet with an impact of 0.37 in the year 2021. No citations have been included yet.

This research explores the differences and the similarities found in two corpora representative of two registers of relevance for Navy submariners in the Spanish Navy Submarine Warfare School. It shows cases in a range of analyses based on multi-dimensional analysis that can inform professional language teaching in such contexts. It is argued that linguistic variation among the texts affords the identification of both converging and diverging patterns of variation across dimensions of use.

V. OUTLINE OF RESEARCH QUESTIONS AND ANSWERS.

In this brief chapter, we outline the research questions in each of the three studies in Chapters 2-4.

STUDY	RESEARCH QUESTION	ANSWERS
STUDY 1	(1) Does Submarine English (SE) exhibit distinct Noun Phrase complexity features? If so, which are these?	<ul style="list-style-type: none"> - SE exhibits distinct Noun Phrase complexity in premodification as well as postmodification slots. -SE presents nouns as the most common type of NPs premodifiers and attributive adjectives as the second one. - SE shows a distinctive postmodification pattern with appositive nouns as the most common type of NPs post-modification.
STUDY 1	(2) Which areas of the NP behavior analysed, if any, are of interest to SE language education?	<ul style="list-style-type: none"> - Multiple pre-modifying nouns have been identified as a target feature in a corpus-based SE curriculum. - Semantic overlapping of Noun + Noun sequences as NPs premodifiers. - Syntactic properties of appositional postmodification. -A fourth area of significance to SE language education is the study and analysis of acronyms that appear in appositional position.

STUDY	RESEARCH QUESTION	ANSWERS
STUDY 1	(1) Does Submarine English exhibit distinct Noun Phrase complexity features? If so, which are these?	<ul style="list-style-type: none"> - SE exhibits distinct Noun Phrase complexity in premodification as well as postmodification slots. -SE presents nouns as the most common type of NPs premodifiers and attributive adjectives as the second. - SE shows a distinctive postmodification pattern with appositive nouns as the most common type of NPs post-modification.
STUDY 1	(2) Which areas of the NP behavior analysed, if any, are of interest to SE language education?	<ul style="list-style-type: none"> - Multiple pre-modifying nouns have been identified as a target feature in a corpus-based SE curriculum. - Semantic overlapping of Noun + Noun sequences as NPs premodifiers. - Syntactic properties of appositional postmodification. -A fourth area of significance to SE language education is the study and analysis of acronyms that appear in appositional position.
STUDY 2	How professional Military submariners' students understand the use of DDL in the process of language learning ?	<ul style="list-style-type: none"> -Professional Military submariners students interpret the use of DDL as a modern, yet unusual, approach to learning English. They perceived it as a challenge compared to the traditional grammar-translation method. -They showed a positive evaluation towards DDL as long as DDL is complemented with other lesson activities. -Qualitative interview data suggest a correlation between higher education and a better understanding of DDL in the learning process. -Attitudes to DDL seem to be affected by underlying ideologies about language learning.

STUDY	RESEARCH QUESTION	ANSWERS
STUDY 3	How can text linguistic approaches, in general, and MDA, in particular, strengthen the pedagogical potential of linguistic variation analysis in the teaching of the subject English for Navy Submariners?	<p>-The paper shows the ways in which Multidimensional Analysis treasures the potential to inform researchers and teachers about SE- specific target linguistic features and their distribution and spread in the corpus data.</p> <p>-MD can facilitate a variationist approach to the selection of materials in SE pedagogy (intra-text or inter-text variation).</p> <p>-MD provides register-specific linguistic insight for the teaching of the subject English for Navy Submariners in terms of how the scores of the underlying dimensions of use characterize the orientation of every single text in the two corpora analyzed.</p> <p>-The paper illustrates how converging and diverging approaches to the analysis of linguistic variation can be used in the teaching of specialized languages.</p>

VI. CONTRIBUTION OF THIS PhD AND FUTURE RESEARCH

This PhD fills a gap in the analysis of SE and the corpus-informed design of teaching materials for Navy submariners. In many ways, this PhD is unique as it examines an application of Military English that, to the best of our knowledge, has not received scholarly attention so far. This thesis sets out to promote the generation of knowledge about the formal characteristics of the SE register and encourage further research in the broader area of Military English.

Using a corpus-driven analysis of the CMSC corpus, we have examined the modification patterns in the most frequent Noun Phrases (NPs) in two corpora of Submarine English. Specifically, we found an above-average frequency of Noun +Noun modification, low adjectival premodification and prepositional phrase modification as well as heavy appositional postmodification. This type of postmodifiers in NPS is infrequent in other registers (Biber et al, 1999) and deserve the attention of lecturers and material developers. Research in the area of complexity has paid attention to the NP mainly in academic English (Biber & Gray, 2016) and learner English (i.e. Díez-Bedmar & Pérez-Paredes, 2020). However, the exploration of the NPs in specialized, professional registers such as SE has received less attention. It is possible that this lack of interest is due to the degree of specialisation and specificity of the genre analysed and, therefore, the lack of critical mass in this area. An analysis of a wider

range of NPs across the two corpora will be useful to provide further insight into SE and inform a specialized curriculum, while our findings are of interest, Military submarine students need to be familiar with other relevant lexico grammatical features in order to understand how their professional register is used. Future work should address a wider range of lexicogrammatical features.

Appositional postmodification motivated the selection of acronyms in our second study. We used a corpus-driven approach and a DDL pedagogic intervention in an ESP context for the first time in a Military Naval School. In this study, we used the Salvage and Rescue of Submarines corpus (SAR) to generate a selection of DDL activities for the teaching of acronyms. We wanted to know how a group of military professionals considered learning language through a DDL approach. Our linguistic goal was not only to contribute with research-informed instruction about acronyms, but also to introduce students for the first time (this is a postgraduate course after obtaining their degree at the Naval School) to a corpus-informed approach to learning Submarine English through concordance lines in professional and military contexts. This is *per se* a major contribution of this PhD as, to the best of our knowledge, DDL had never been used with the Military and had very rarely used outside tertiary or secondary education. We also highlight the importance of acronyms in specialized discourse and their occurrence in different disciplines, which was to some extent neglected in ESP research (Valipuri and Nassaji, 2013; Coxhead, 2016). Future work should expand on our initial findings by including other lexicogrammatical target features. Further reflection on the use of DDL in non-HE contexts can impact our understanding of DDL practices and contribute

to enlarging the scope of the use of corpora in language education and language teaching.

The analysis of linguistic variation using multidimensional analysis (MD) has not widely adopted in the teaching of professional and specialized languages outside university contexts. Given the lack of research in Submarine English, and with the intention of familiarize researchers and language instructors with the linguistic nature of texts in our two corpora, we studied variation in both corpora following Biber (1988, 2019) MD approach. Although they represent two types of text of interest to Navy submariners (professional journals and salvage and rescue manuals), they both share a topic domain and community. The quantitative and qualitative results of the analysis of the SE register gives us a complex picture of both linguistic and functional features of these registers, as well offering instructors the possibility of using these data. The linguistic differences across the Dimensions have a big impact on language teaching. Both corpora show a high score in nouns and agentless passives (D1 and D4) but they diverge in D2, 3 and 4, which means they display register differences that tend to be ignored in research applications. There are also differences in inter-corpus and intra-corpus textual analysis, which are useful and pedagogic too. In simple terms, if we wanted to engage learners with linguistic features associated to past tenses, third person pronouns or the use of the perfect aspect, then our texts are in the CMSC as this corpus shows a more narrative orientation. If, on the contrary, we seek to teach relative clauses, conditional subordination and modals, the SAR corpus can provide us the texts. This corpus is more informational and persuasive.

One major contribution of Study 3 is the discussion about converging and diverging orientations in MD analysis and their application in specialized language teaching. Understanding linguistic variation in SE implies a close inspection of both individual texts and corpora. Converging patterns of variation focus on texts that behave similarly on a given dimension of use. Diverging patterns of variation show how the two corpora display frequencies and distributions of linguistic features that allow diverging functional interpretations on a given Dimension. As we outlined in Study 3, research has tended to focus on the differences. However, in professional and specialized language analysis, the study of converging patterns of variation can impact on the evaluation of the texts that can inform pedagogy a corpus-pedagogy approach. This needs to be further explored in future studies. Further work could take up our MD analysis and apply the converging/diverging framework suggested in our study in an extended semi-experimental design for the teaching of Military English.

Professional Military languages have not received enough attention in the ESP literature and more generally in applied linguistics. Conrad (2019) states that register analysis effectively identifies qualities that are present throughout texts and are ubiquitous in a group of texts or corpora. For her, “register analysis is useful because the findings are directly applicable to teaching, which is a goal in many ESP register studies [and] the findings are used in creating new teaching materials that integrate writing skill development with [ESP] content and practice” (p.190). This thesis has followed Conrad’s suggestion that ESP can draw on corpus analysis to analyse and understand specialised language use. We have approached a specific scenario of use

where the context of production and reception, a wide range of linguistic features, and the functional relationships between the situational qualities and the linguistic features have been explored. Combined with the analysis of phrasal features and classroom activities, this PhD has offered, to the best of our knowledge, the first implementation of military-English DDL and the first analysis of SE using a MD register perspective.

One of the possible drawbacks of the DDL activities and approach is that because they are not included in the initial lesson design, students may perceive them as reinforcement activities and pay less attention to them than if they are included as part of the assessment performance. Therefore, if they are included as compulsory in the syllabus, other didactic actions will be imposed.

In addition, some pedagogically oriented research about the lexico-grammatical features of SAR corpus through MDA (**Study 3**) displays a possible focus on the conditional subordination and some types of prediction, necessity and possibility modals. The converging and diverging patterns of variation of CMSC and SAR gave us a global linguistic picture of both corpora. Linguistic features such as necessity modals exhibit diverging frequencies within the same corpus, that is, intra-corpus texts which would require special attention by teachers or syllabus designers. We should obtain further evidence of intra-corpus variation to provide us with more opportunities for pedagogical purposes. The identification of certain key areas such as passive voice and conditionals together with acronyms and NPs with multi-nouns premodification including other features of grammar, lexis and discourse are a good breeding ground

for the creation of a class manual for military submarine students for at least one semester.

REFERENCES

- Albi, A. B., Vidal, N. J., López, P. O., & Conde, T. (2014). A genre-based approach to the teaching of legal and business English: The GENTT specialized corpus in the LSP classroom. *Languages for Specific Purposes in the Digital Era*, 177-196.
- Anthony, L. (2018). Visualisation in corpus-based discourse studies. In *Corpus Approaches to Discourse* (pp. 197-224). Routledge.
- Barber, C. L. (1962). Some measurable characteristics of modern scientific prose. *Contributions to English syntax and philology*, 21-43.
- Belcher, D. (2009). What ESP is and can be: An introduction. *English for specific purposes in theory and practice*, 1-20.
- Biber, D. (2019). Multi-dimensional analysis: A historical synopsis. *Multi-dimensional analysis: Research methods and current issues*, 11-26
- Biber, D. (2008). Corpus-based analyses of discourse: Dimensions of variation in conversation. In *Advances in discourse studies* (pp. 110-124). Routledge.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
- Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Long-man grammar of spoken and written English. *Tesol Quarterly*.
<https://doi.org/10.2307/3587792>

-
- Biber, D., Reppen, R., & Conrad, S. (2002). Developing linguistic literacy: Perspectives from corpus linguistics and multi-dimensional analysis. *Journal of Child Language*, 29(2), 449-488.
- Bhatia, V., Pérez-Paredes, P., & Sánchez, P. (2012). Researching specialized languages. *Researching Specialized Languages*, 1-246.
- Bocanegra-Valle, A. (2013). Maritime English. *The Encyclopedia of Applied Linguistics*.
- Bocorny, A. E. P. (2010). The teaching of aviation terminology to ab initio pilots: the specialized communication context and the communication context in the ESP classroom. *Aviation in Focus-Journal of Aeronautical Sciences*, 1(1), 12-20.
- Boulton, Alex & Pérez-Paredes, Pascual (2023). Data-driven language learning: In Regine Hampel & Ursula Stickler (eds), *Bloomsbury Handbook of Language Learning and Technology*. Bloomsbury.
- Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning & Technology*, 25(3), 66-89.
- Boulton, A. (2019). Data-driven learning for younger learners: Obstacles and optimism. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. Routledge.
- Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348-393.
<https://doi.org/doi:10.1111/lang.12224>
- Boulton, A. (2017). Corpora in language teaching and learning. *Language Teaching*, 50(4), 483-506. <https://doi.org/10.1017/S0261444817000167>
- Boulton, A., & Pérez-Paredes, P. (2014). ReCALL special issue: Researching uses of corpora for language teaching and learning Editorial Researching uses of

-
- corpora for language teaching and learning. *ReCALL*, 26(2), 121-127.
<https://doi.org/10.1017/S0958344014000068>
- Boulton, A. (2012). Corpus consultation for ESP: A review of empirical research. *Corpus-informed research and learning in ESP: Issues and applications*, 261-291.
- Cole, C., Pritchard, B., & Trenkner, P. (2007). Maritime English instruction – ensuring instructors' competence. *Ibérica*, 14, 123–148.
- Conrad, S. (2017). A comparison of practitioner and student writing in civil engineering. *Journal of Engineering Education*, 106(2), 191-217.
- Conrad, S. (2019). Register in English for academic purposes and English for specific purposes. *Register Studies*, 1(1), 168-198.
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and education*, 17(4), 391-406.
- Crosthwaite, P., & Cheung, L. (2019). Learning the language of dentistry: Disciplinary corpora in the teaching of English for Specific Academic Purposes (Vol. 93). John Benjamins Publishing Company.
- Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. *The Routledge handbook of English for academic purposes*, 177-190.
- Coxhead, A. (2023). Corpora for Teaching and Learning Vocabulary in ESP. In *The Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 193-205). Routledge.
- Curry, N., & Pérez-Paredes, P. (2021). Stance nouns in COVID-19 related blog posts: A contrastive analysis of blog posts published in The Conversation in Spain and the UK. *International Journal of Corpus Linguistics*, 26(4), 469-497.

-
- Cortés, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397–423.
- Crosthwaite, P., & Cheung, L. (2019). *Learning the language of dentistry: Disciplinary corpora in the teaching of English for Specific Academic Purposes* (Vol. 93). John Benjamins Publishing Company.
- Díez-Bedmar, M. B., & Pérez-Paredes, P. (2020). Noun phrase complexity in young Spanish EFL learners' writing: Complementing syntactic complexity indices with corpus-driven analyses. *International Journal of Corpus Linguistics*, 25(1), 4-35.
- Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific Purposes. A Multi-Disciplinary Approach*. Cambridge: CUP.
- Egbert, J., Biber, D., & Gray, B. (2022). *Designing and evaluating language corpora: A practical framework for corpus representativeness*. Cambridge University Press.
- Farr, F., & Karlsen, P. H. (2022). DDL pedagogy, participants, and perspectives. In R. Jablonkai & E. Csomay (Eds.), *the Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 329-343). Routledge.
- Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. *Discourse in the professions: Perspectives from corpus linguistics*, 11-33.
- Flowerdew, L. (2011). *Corpora and language education*. Springer.
- Friginal, E., & Roberts, J. (2022). Corpora for materials design. In *The Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 131-146). Routledge.
- Gavioli, L. (2002). *Exploring Corpora for ESP Learning*. John Benjamins Publishing.

-
- Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed., pp. 359-370). Routledge.
- Gray, B. (2013). More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles. *Corpora*, 8(2), 153-181.
- Gray, B. (2022). What can a corpus tell us about registers and genres?. *The Routledge Handbook of Corpus Linguistics*, 235-249.
- Halliday, M. A. K., McIntosh, A., & Stevens, P. (1964). *The linguistic sciences and language. Teaching.* (Longmans' Linguistic Library.) London: Longman.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes.* Cambridge university press.
- Hyland, K., & Jiang, F. K. (2021). A bibliometric study of EAP research: Who is doing what, where and when? *Journal of English for Academic Purposes*, 49, 100929.
- IMO. (2001). *Imo standard marine communication phrases (SMCP).*
- International Maritime Organization. (2001). *Final Act of the Conference of Parties to the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW), 1978.*
- Johns, T. (1991). *Should you be persuaded: Two samples of data-driven learning materials* (Vol. 4, pp. 1-16). na.
- Johns, A. M. (2013). The history of English for specific purposes research. *The handbook of English for specific purposes*, 5-30.
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217–241.
- Lynch, J. K., & Porcellato, A. M. (2020). The case for an aviation English screening tool for US flight schools. *The ESPecialist*, 41(4).

-
- McEney, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Taylor & Francis.
- Nini, A. (2019). The multi-dimensional analysis tagger. *Multi-dimensional analysis: Research methods and current issues*, 67-94.
- O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- O'Keeffe, A. (2021). Data-driven learning, theories of learning and second language acquisition. In P. Pérez-Paredes & G. Mark (Eds.), *Beyond Concordance Lines: Corpora in language education* (pp. 35-55). John Benjamins
- Orna-Montesinos, C. (2013). English as an international language in the military: A study of attitudes. *LSP Journal-Language for special purposes, professional communication, knowledge management and cognition*, 4(1).
- Parkinson, J., Mackay, J., & Demecheleer, M. (2016). Situated learning in acquisition of a workplace genre. *Vocations and Learning*, 11, 289-315.
- Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning*, 35(1-2), 36-61.
- Pérez-Paredes, P., & Sánchez-Tornel, M. (2019). The linguistic dimension of L2 interviews: A multidimensional analysis of native speaker language. *Focus on ELT Journal*, 1(1), 4-26.
- Ren, C., & Lu, X. (2021). A multi-dimensional analysis of the Management's Discussion and Analysis narratives in Chinese and American corporate annual reports. *English for Specific Purposes*, 62, 84-99.
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge.

- Swales, J. (1985). English as the international language of research. *RELC journal*, 16(1), 1-7.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.
- Trenkner, P. (2007). Maritime English. An attempt of an imperfect definition. In *Proceedings of the 2nd IMLA Workshop on Maritime English in Asia* (pp. 1–8).
- Upton, T. A. (2012). LSP at 50: Looking back, looking forward *Ibérica*, (23), 9-28.
- Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, 12(4), 248-263.