LANGUAGE, IDENTITY, AND VOICE:

EUROPEAN PERSPECTIVES ON ENGLISH AS THE

INTERNATIONAL LANGUAGE OF SCIENCE

By

FRANCES M. JUNNIER

Master of Arts in Politics and International Relations
University of Aberdeen
Aberdeen, United Kingdom
2001

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2022

LANGUAGE, IDENTITY, AND VOICE:

EUROPEAN PERSPECTIVES ON ENGLISH AS THE

INTERNATIONAL LANGUAGE OF SCIENCE

Dissertation Approved:

Dr. Carol L. Moder
_____
Dissertation Adviser

Dr. An Cheng
_____

Dr. Stephanie M. Link
_____

Dr. Wouter Hoff
_____

ACKNOWLEDGEMENTS

I would like to thank my PhD advisor, Dr. Carol Moder, for her valuable mentorship. I

would also like to thank my advisory committee, Dr. Wouter Hoff, Dr. Stephanie Link,

and Dr. An Cheng, for their continued guidance and support. Finally, I would like to

acknowledge the contributions of the European research scientists who acted as

informants for this project.

Name: FRANCES M. JUNNIER

Date of Degree: JULY, 2022

Title of Study: LANGUAGE, IDENTITY, AND VOICE: EUROPEAN PERSPECTIVES ON ENGLISH AS THE INTERNATIONAL LANGUAGE OF SCIENCE

Major Field: ENGLISH

Abstract: Drawing on theories from Applied Linguistics research, this doctoral dissertation brings together European perspectives on *language*, *identity*, and *voice* in an examination of English as the international language of science (EILS). Using the theoretical frame of Dynamic Systems Theory (DST), each of the chapters adopts a different perspective to examine EILS. At the macro-level, I conduct a comparative Critical Discourse Analysis (CDA) of English as a lingua franca discourse in the published declarations and communications of the European Higher Education Area (EHEA) and the European Research Area (ERA). The analysis adds nuance to claims that the EHEA has a disregard for language related matters and uncovers intra-institutional conflicts in how language is conceptualised between actors and across portfolios in the European Union (EU). At the meso-level, I use the lens of Conversational Analysis to enrich understandings of how European research scientists construct linguistic and disciplinary identities through turns-at-talk in a series of semi-structured interviews. The analysis emphasizes the importance of the research interview's interactional features, which are often ignored in more dominant approaches to interview in English for Academic Purposes (EAP) research. Finally, at the micro level, a Usage-Based Linguistics (UBL) study reveals new insights into the development of a European research biologist's authorial voice across timescales and contexts. In particular, the study highlights the importance of discourse context in the analysis of voice features. In sum, this doctoral dissertation contributes to an important and productive site of scholarly debate by adopting epistemological approaches that foreground the importance of interaction and context in understanding language as a social practice. It provides original perspectives on the conceptualisation of language, the construction of identity, and the development of voice in the dynamic and complex system that is EILS.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

English has become the global language of scientific communication (Ammon, 2001). Its increasing dominance is seen in number of scientific domains such as higher education, research and development, and scientific publishing. In the higher education sector, internationalisation efforts have led to a rapid expansion in English-medium instruction (EMI) in non-Anglophone countries (Dimova et al., 2015; Macaro et al., 2019); in research and development, national research policy and assessment structures have promoted transnational research activities resulting in the prevalence of English-medium scientific collaborations and exchanges (Plo Alastrué, 2015); and in scientific publishing, power and prestige continue to be concentrated in the hands of English-medium journals and their Anglo-American gatekeepers resulting in the dominance of English-medium publication (Canagarajah, 2002; Swales, 2000). While the degree to which English is used across these domains varies by language communities and countries (Ammon, 2001), the general trend over the past several decades has been towards more English in non-English environments (Mauranen et al., 2010).

One such 'non-English' environment experiencing the dominance of English as the international language of science (EILS) is mainland Europe. To use Kachru's (1990)

taxonomy, Europe (excluding the UK, Ireland, and Malta) forms part of the Expanding

Circle of English, meaning that across the continent, English has no official status, but is

introduced in schools and universities as a foreign language. However, over the past

twenty years, the instrumental use of English as a communicative 'tool' in higher

education and research contexts across Europe has rapidly expanded (Plo Alastrué,

2015). This rapid expansion has led to a productive site for scholarly debate in

conferences, research articles, and edited volumes, many of which were generated

through a research partnership between European universities called the *English in*

Europe: Opportunity or Threat? project which ran from 2012 to 2014.

1.1 Overview

Drawing on theories from Applied Linguistics research, this doctoral dissertation

brings together European perspectives on *language*, *identity*, and *voice* in an examination

of EILS. In order to approach the topic in an integrated way that recognizes both the

relational and dynamic interplay of elements and agents within complex systems (Larsen-

Freeman, 2012; Douglas Fir Group, 2016), each of the three studies in this dissertation

uses a different perspective/level of analysis to interrogate EILS. At the macro (societal)

level, I begin with a comparative critical discourse analysis of English as a lingua franca

(ELF) in European higher education and research contexts. At the meso (community)

level, I use the lens of Conversation Analysis (CA) (Sacks et al., 1974) to analyse how

European research scientists construct linguistic and disciplinary identities in semi-

structured research interviews. Finally, at the micro (individual) level, I take a Usage-

Based Linguistics (UBL) approach to track the textual development of authorial voice of

a European research biologist over a period of nine years as he moves across physical, linguistic, and disciplinary boundaries.

1.2 Theoretical Connections

In the previous section, I alluded to EILS as being part of a complex system or dynamic interplay of elements and agents. This conceptualisation of EILS is informed by Dynamic Systems Theory (DST) (de Bot, Lowie & Verspoor, 2005 & 2007) and Complex Dynamic Systems Theory (CDST) (Larsen-Freeman, 2012), which posit that language itself is a dynamic system. Dynamic systems can be characterised as sets of variables that develop over time in complex and non-linear ways through interaction with their environment and through internal self-reorganization (de Bot, Lowie & Verspoor, 2007). The relevance of dynamic systems to Applied Linguistics, was highlighted by Larsen-Freeman (1997) in her seminal work on chaos/complexity theory and her analogy between language acquisition and complex, non-linear systems existing in nature. A key concept of DST/CDST that informs this dissertation is *complete interconnectedness*. Complete interconnectedness means that changes in one variable will impact the other variables in the same system and that systems are always part of other systems, nested one within the other (de Bot, Lowie & Verspoor, 2007). Therefore, although each chapter in the dissertation is a discrete study, the macro, meso, and micro units of analysis are convenient abstractions that are better viewed as systems nested one within the other as part of a complex system representing English as the global language of scientific communication in Europe.

Another key concept, which DST/CDST has helped to inform, is that of *agency*.

CDST maintains the structure-agency complementarity while also highlighting that agency is relational and spatially-temporally situated (Larsen-Freeman, 2019). Put another way, agency is not something that is *possessed* by the individual but is *achieved* by the individual "orienting to the different affordances of the social and material worlds" (Larsen-Freeman, 2019, p.73). Understanding the interplay of structure and agency is important as each of the studies in the dissertation is fundamentally concerned with how elements and agents shape, and are shaped by, EILS as a dynamic complex system. To conclude, in recognising language as a dynamic system, it becomes impossible to delimit language as a cognitive tool that individuals use to communicate one thing or another. The theoretical approach to language undergirding this dissertation recognizes that language "inextricably involves cognition, emotions, consciousness, experience, embodiment, brain, self, human interaction, society, culture, mediation, instruction, and history in rich, complex, and dynamic ways." (Douglas Fir Group, 2016, p. 39). This definition may indeed seem broad, and it is unsurprising that the 15 scholars who form the Douglas Fir Group come from traditions in Second Language Acquisition (SLA) and Applied Linguistics research that foreground context and interaction (i.e. socio-cultural theory, usage-based linguistics, conversation analysis, and social identity theory). In this next section, I discuss how the frameworks I have chosen to investigate EILS draw on these traditions to form an epistemological as well as theoretical coherence to the dissertation.

1.3 Epistemological Connections

In addition to their topical and theoretical connections, the three studies adopt

complementary epistemological approaches to the study of EILS. CDA, CA and UBL

share fundamental beliefs about how we might know language — all foregrounding the

importance of interaction and context in understanding language as a social practice.

Context is critical in CDA because social and political power relations are constructed

and reflected in the discourse (Fairclough & Wodak, 1997). Thus, although CDA

researchers might adopt various methodological approaches to discourse, they are

essentially interested in examining both the overt and covert manifestations of power,

dominance and discrimination as they are produced in language (Wodak, 2011). While

CA's intellectual traditions and empirical foci differ substantively from CDA (Wooffitt,

2005), both approaches foreground the primacy of context in the study of naturally

occurring text and talk. Indeed, CA has become a well-developed analytical approach in

Applied Linguistics and other social sciences as it attempts to "describe, analyze and

understand talk as a basic and constitutive feature of human social life." (Sidnell, 2010).

Moreover, like CDA, CA advocates a detailed study of "the structures of text and talk

and their interactional and social functions" (Van Dijk, 1999, p. 460). Finally, Usage-

based Theory also emphasises the importance of context and interaction as it posits that

cognitive representations for language are built as language users experience, encode and

categorize utterances based on form, meaning and context (Bybee, 2013). Our language

is, in effect, the mental organization of our experiences with language in *use*. In sum, the

dissertation brings together perspectives on EILS using different levels of analysis and

distinct but complementary analytical approaches to the study of text and talk. The

purpose of this introductory chapter is to provide an overview of the context of EILS in

Europe (section 1.4); to outline the disciplinary frames through which I investigate EILS

(section 1.5); and to provide theoretical perspectives on the constructs of language, identity, and voice (section 1.6, 1.7, 1.8).

1.4 EILS in Europe

Non-Anglophone universities across Europe have integrated English in aspects of instruction, research and administration (Villares, 2019). Most notable has been the exponential growth in EMI across the continent, particularly at the Masters and PhD levels (Dimova et al., 2015), as well as a gradual shift to English-only publication contexts for European academics (Giannonni, 2008; Curry & Lillis, 2010). Hultgren (2014) argues that stances taken to the "Englishisation" of scientific life in Europe often group around two binaries. From one perspective, scholars suggest that while processes of marketization and internationalisation influence English use, they argue that "Englishisation" is also driven by agentive choices at local levels (Coleman, 2013; Ferguson, 2015), by happenchance (Brumfit, 2004), by no "central hand" (Montgomery, 2004, p. 1334). On the opposite side, English expansionism is characterised as being driven by nefarious hegemonic forces (i.e. Phillipson, 2006). Critical theorists have characterized the spread of English in Europe as a form of linguistic imperialism (Phillipson, 1992) or epistemicide (Bennet, 2015). From these critical perspectives, the expansion of English and Anglo-Saxon epistemological stances in European higher education and research environments threaten local languages and knowledge paradigms (Bennet, 2015; Phillipson, 2017).

In addition to these more theoretically focused concerns, research has also considered how European multilingual scholars and researchers actually negotiate

6

differential power relations in order to gain access and participate in academic knowledge production (Curry & Lillis, 2004). A significant body of this scholarship focuses on how non-Anglophone European scholars negotiate the discursive and non-discursive demands of publishing in English-medium journals (i.e. Curry & Lillis, 2004; Giannoni, 2008; Kuteeva, 2015; Martin, et al., 2014; Mur-Dueñas, 2013). Indeed, over the past several decades, various branches of Applied Linguistics' research have sought to describe, explicate, and theorize the phenomenon of English as the international language of scientific publishing in Europe and beyond.

1.5 EILS in Europe and Applied Linguistics Research

Over the past thirty years, English for academic purposes (EAP) has developed as a distinct branch of Applied Linguistics to "help learners gain access to ways of communicating that have accrued cultural capital in particular communities, demystifying academic discourses" (Hyland & Shaw, 2016, p. 6). More recently, English for Research Publication Purposes (ERPP) has emerged as a sub-disciplinary field of EAP with a specific focus on issues faced by scholars publishing in English as an Additional Language (EAL). EAP's practice-oriented approach and its early prescriptive tendencies led to criticisms that it was un-critical and accommodationist (Benesch, 2001; Pennycook, 1997). Critical EAP (CEAP) sought to address these concerns by exploring EAP in relation to "students' and teachers' complex and overlapping social identities: class, race, gender, ethnicity, age and so on." (Benesch, 2009, p. 81).

EAP research has also been criticized for being too textually focused (Swales, 2019). The privileging of text above practice has led to claims that the field has treated

language and particularly writing "as primarily a linguistic, and perhaps even an autonomous, object rather than something which is socially embedded in particular lives, disciplines and contexts." (Hyland, 2018, p. 395). However, Tardy (2006) argues that while EAP's early attention to written genres may have focused on text-type description, research agendas in the field evolved to include more nuanced understandings of genres as social practices that shape and are shaped by human activity. Indeed, socially situated approaches to teaching and learning through genre foregrounded the importance of learner characteristics, including prior writing experience and learning goals (i.e. Cheng, 2006) as well how a focus on genre could help in the acquisition of disciplinary identity (i.e. Dressen-Hammouda, 2008). Theories from Academic Literacies (Ac Lits) (Lea & Street, 1998) have been influential in informing EAP research agendas that view genres and genre production as complex socially embedded activities. Ac Lits approaches emphasize the importance of learners' literacy histories in negotiating and instantiating academic practices as well as the deeply social and political nature of those practices (Casanave, 2002). Indeed, EAP research has used Ac Lits perspectives to add valuable insight into how postgraduate academic writers in Europe use prior writing experience, disciplinary knowledge structures, and future identities to produce academic genres (Kaufhold, 2015; Lehman & Anderson, 2017). The Ac Lits model also stood in contrast to more mainstream SLA approaches of the time that considered "the foreign language speaker as a deficient communicator struggling to overcome an underdeveloped L2 competence, striving to reach the 'target' competence of an idealized native speaker" (Firth & Wagner, 1997, p. 285). Although the SLA field had evolved much since then, much SLA terminology remains encumbered with deficit-oriented ideologies (Douglas

Fir Group, 2016). In short, CEAP and Ac Lits have been influential in EAP's social turn by focusing on the social and political context, literacy history, and multilingual resources of whole-bodied writers.

A further but important influence of these two theoretical frameworks on EAP research has been their critical lens on what actually constitutes the *English* in EAP. Curry and Lillis (2019) argue that fields indexed with the singular term *English*, such as EAP and ERPP, have reifying tendencies that promote a singular native-speaker standard of English. Indeed, Lorés-Sanz (2016) notes that in its attention to academic genres, EAP has tended to focus on so called 'native' speaker standards without recognising that "all writing in international journals is, by definition, for an international rather than an ENL (English as a native language) audience." (p. 54). The next section of this chapter seeks to illuminate how the English language might be conceptualised in a European EILS context and how issues of identity and voice are central to issues in multilingual academic writing.

1.6 *Language*

Scholarship that challenges the privileging of a native English speaker 'standard' is by no means new. The World Englishes paradigm sought to reconceptualise English as a plural language that "embodied multiple norms and standards" (Canagarajah, 2006). Furthermore, over two decades ago, Widdowson (1994) questioned who actually owned 'standard English' arguing that it is not owned by the people "in an offshore European island" (p. 382) but rather that Standard English – particularly in written form – is now the preserve of business and scholarly communities that "transcend communal and

cultural boundaries"(p. 382). The existence of contact languages, used among

communities who do not share the same first language, have been documented for many

centuries in Europe (Jenkins et al., 2011). However, as a distinct field of Applied

Linguistics, ELF scholarship is relatively new with the foundations of the research

agenda attributed to the European linguists, Jenkins (2000) and Seidlhofer (2001). ELF is

defined as a contact or vehicular language between users of different first languages,

including native-speakers of English (Jenkins et al., 2017). Moreover, it is motivated by

communicative need rather than linguacultural factors (Seidlhofer et al., 2006).

Seidlhofer (2001) described a conceptual gap between Applied Linguistics' research on

the English language and how the English language was actually used in a wide variety

of geographies and domains. In Chapter II, I address this conceptual gap by defining and

viewing the use of English in European education and research contexts as ELF. In other

words, I recognize that in the geopolitical domain that is Europe, English is most often

used by speakers of different first languages as a common contact language to

accomplish specific communicative purposes. In addition, while ELF approaches may be

a considerable way from weakening the hold of native speaker 'standard' positions, it has

been suggested that professional disciplinary communities will become more influential

in setting their own language standards (Mauranen et al., 2010). This echoes what

Montgomery (2004) characterised as two opposing linguistic movements in which the

centripetal forces of global English are countered by the centrifugal forces of specialized

disciplinary speech practices. Key within these two opposing movements is how

multilingual scholars navigate their linguistic and disciplinary identities within and across

discourse communities (Swales, 1990).

1.7 *Identity*

EAP scholarship that seeks to understand academic writing practices within their socio-political contexts recognises that issues of identity are central to the ways both novice and expert multilingual scholars construct knowledge (Curry & Lillis, 2004). Poststructuralist perspectives in Applied Linguistics research view identity as multiple, emergent, dynamic, and contested (Baker, 2015) and this is particularly relevant "when individuals move across geographical and psychological borders, immersing themselves in new sociocultural environments" (Block, 2007, p. 864). In Chapter III, I use Bucholtz & Hall's (2005) broad definition of identity as "the social positioning of self and other" and Burgess and Ivanic's (2010) framework for the discoursal construction of writer identity to investigate, through semi-structured interviews, how European research scientists, who have become internationally mobile, construct ELF and disciplinary identities. Possibilities for self-hood (Burgess & Ivanič, 2010) are particularly relevant when examining disciplinary identity construction as it is within distinct discourse communities that European research scientists shape identity. Swales (1990) defines discourse communities as socio-rhetorical groupings of novice and expert members who, among other features, interact to pursue common communicative goals using community specific genres and specialized terminology. The decision to select informants for Chapter III solely from the biological sciences is purposeful. The hard sciences are often framed in respect of their reputation as author evacuated (Harwood, 2005). However, Halloran (1984) argues that in the post-war period, the more entrepreneurial tenor of the biological sciences opened new possibilities for selfhood, thus resulting in new

rhetorical/ethical styles in academic writing in the biological sciences. Moreover, "writer identity is not optional; all texts say something about the writer, although some are more marked than others" (Matsuda, 2015, p. 146). One way in which identity can be marked in texts is through authorial voice.

1.8 *Voice*

Once the domain of literary studies, authorial voice has become an area of increased research and central significance in academic writing (Hyland & Guinda, 2012). Matsuda's (2001) pivotal characterization of voice as the "amalgamative effect of the use of discursive and non-discursive features that language users choose, deliberately or otherwise, from socially available yet ever-changing repertoires" (p. 40) is seen to recognise voice as the interplay between individual self-expressions and social context. In Chapter IV, I address this interplay by taking a usage-based linguistics (UBL) approach to analyse how a molecular biologist who writes in English as an additional language for research publication purposes constructs authorial voice over a period of nine years. Context and community are central in determining the ways in which individual voice emerges (Hyland & Guinda, 2012). The contexts and communities in which and for which a researcher conducts, writes and publishes their research, shape and constrain their developing disciplinary voice to take "into account the social worlds for and out of which a text is produced" (Tardy, 2012, p. 39). Moreover, a diachronic approach to the study of voice is adopted in Chapter IV to recognise that in addition to context, time is equally important in understanding the discoursal construction of authorial voice (Burgess & Ivanic, 2010).

1.9 Summary

In summary, this dissertation draws on theories from EAP, Academic Literacies, and ELF research paradigms to investigate EILS in education and research contexts. It does so at different levels of analysis – from large complex units of analysis like the European Union to small complex units of analysis like the individual multilingual writer. Furthermore, studies in the dissertation use approaches to language that share epistemological and theoretical perspectives about what language is and how meaning is constructed by our lived experiences.

CHAPTER II

MACRO PERSPECTIVE: A COMPARATIVE CRITICAL DISCOURSE ANALYSIS OF
ENGLISH AS A *LINGUA FRANCA* IN THE EUROPEAN UNION

2.1 Introduction

As a communicative medium of choice among users of different first languages
(Seidlhofer, 2011), English has been defined as the *lingua franca* of knowledge
dissemination in European academic life (Lasagabaster, 2015; Mur-Dueñas, 2013). While
non-Anglophone scholars are certainly motivated to learn English, using an additional
language while also learning the social practices associated with academic, research, and
publishing contexts has been described as presenting a dual burden for doctoral students
and early career researchers (Curry & Lillis, 2019). As a result, increasing attention is
being paid to the ways in which non-Anglophone doctoral researchers are prepared and
supported to use English as a *lingua franca* (ELF) of the academic and research
communities in Europe (Peréz-Llantada, 2018). A number of stakeholders are involved at
different levels of influence in developing research and education policy that impacts
doctoral student experience with ELF. Individual higher education institutions; national
governments; and supra-national bodies, such as the European Union (EU), all contribute
to a multi-layered higher education (HE) language policy (Soler-Carbonell et al., 2017).

Previous research has examined the ideologies and practices of these stakeholders with regards to ELF. In their edited volume, Hultgren et al., (2014) explore the 'Englishization' of Nordic universities, identifying opposing ideological discourses, which promote either instrumental (internationalist) or protectionist (culturist) stances to language policy in HE contexts. Instrumental approaches to language policy in the EU can be exemplified by the widely contested decision of the Politecnico di Milano to adopt English in all MA and PhD programs from 2014 (Santulli, 2015). More protectionist stances were illustrated in Bjorkman (2014) who found that Swedish university policy documents did not reflect or provide guidance on everyday ELF practices, but instead were heavily informed by top-down language ideologies, which adopted protectionist stances toward local languages. Soler-Carbonell et al., (2017) investigated the local, national and international dimensions of HE language policy in the Baltic region. They found that local (HE institutions) and international (EU) actors take a more instrumental approach to the positioning of English as the de facto language of internationalisation while at the national level the promotion of English in HE contexts ran contrary to national language policies. Lasagabaster (2015) also conducted a multi-level analysis to investigate how different stakeholders influenced language policy in Spain. In the context of Spanish universities, Villares (2019) found that university language policy documents adopt positive attitudes toward languages and the promotion of multilingualism, in line with national policy documents. In short, depending on their geographical locations and/or sociolinguistic contexts, HE institutions in Europe may take more protectionist stances or more instrumental approaches to ELF. At the national level, there is greater focus on the protection of local languages and an emphasis on language as a vehicle of

social and cultural identity (Soler-Carbonell et al., 2017). Nation-state protectionist stances are exemplified by France's Loi Toubon or Sweden's 2009 Language Act, which legally impose their respective national languages in public administration, commerce and state media. (Ferguson, 2018). However, previous research has noted that, at the supra-national level — namely the EU, there has been a conspicuous absence of explicit language policy as it relates to HE and research contexts (Hultgren et al., 2015). Second, it is noted that there is a mismatch between language rhetoric and on-the-ground language practices (Ferguson, 2018). Furthermore, much of the previous research on ELF policy in Europe has focused on differences across geopolitical contexts and frictions and inconsistencies *between* HE stakeholders, with less attention being given to differences or frictions that may exist *within* levels of influence. Power dynamics in complex organizations are commonly characterized by *intra*-institutional conflict (Guillén, 2007) and the various actors within a complex supranational political organism (Krzyżanowski & Wodak, 2011) such as the EU, may hold divergent ideologies and promote different practices in the same policy areas. Randour et al., (2020) suggest that comparative studies on the positions different social and political actors take on the same topics provide promising grounds for future research in political discourse. Moreover, as Duguid and Partington (2018) note, even when the absence of a linguistic phenomena is identified, the significance of the absence can only be appraised by comparing it with the behaviour of the phenomena elsewhere. Consequently, applying a comparative approach to different policy actors *within* a complex organisation like the EU could be a promising way to interpret the absence of certain linguistic phenomena while also potentially uncovering *intra*-institutional differences toward these phenomena.

16

Therefore, the purpose of this study is to use the discourse of two EU policy actors as 'entry points' to uncover the ideological positions and potential *intra-institutional* differences in European ELF policy at the supranational level. More specifically, I conduct a comparative analysis of ELF discourse in the published declarations and communications of two policy actors that have remits for HE and research: the European Higher Education Area (EHEA) and the European Research Area (ERA).

2.1.1 Research Questions

Following the work of Baker et. al (2008), methods associated with Critical Discourse Analysis (CDA) and Corpus Linguistics (CL) are used to answer the following research questions:

i.   How is ELF represented in general EU discourse?

ii.   How is ELF represented in the discourse of the EHEA and the ERA?

iii.   What are the frequent topics associated with language in EHEA and ERA discourse?

iv.   What are the attitudes toward language in EHEA and ERA discourse?

2.2 Language policy, ELF and EU actors

In order to analyse the discursive representation of ELF by the EHEA and the ERA, it is necessary to (i) describe the broader socio-political context of language policy in the EU, (ii) define what ELF means in the context of European education and research,

and (iii) describe the institutional positions the EHEA and the ERA occupy within these contexts.

2.2.1 Language policy and the EU

Spolsky (2004) suggests that in order to locate language policy, we need not only examine the explicit interventions of policy actors but also make inferences from the implicit beliefs and practices of those actors. As such, Spolsky (2004; 2019) suggests that language policy can be characterised by three independent but interconnected components: (i) *language ideology* refers to the beliefs attitudes, and assumptions assigned to language choices; (ii) *language practices* refers to the actual choices stakeholders make from their language repertoires in specific communicative contexts, and (iii) *language management* refers to the way in which stakeholders' attempt, either implicitly or explicitly, to shape language ideology and practices. Shohamy (2006) adds to Spolsky's (2004) characterisation in her critical framework of language policy by including the concept of covert and overt mechanisms, which are used as the means to affect, create, and perpetuate language policies. It is evident in the reading of various policy documents that the EU overtly values linguistic equality and diversity — viewing language as a direct expression of EU cultural identity (Iskra, 2022). The EU is one of the only supranational organizations of its kind to adopt a "completely egalitarian language policy with respect to its member states" (Cogo & Jenkins, 2010 p. 272). Linguistic diversity is described as one of the fundamental values of the EU and is given legal basis in the Treaty of the European Union (Iskra, 2022). Moreover, the European Commission's multilingualism policy states that by 2025 European citizens should

master two other languages in addition to their mother tongue (Katsarova, 2019) and a number of policy initiatives that impact doctoral and early career scientist experience (i.e. European Language Label and Erasmus+) are in place to realise this objective. For example, Erasmus+ states that multilingualism and linguistic diversity are integral to the European project and, as such, its aim is to support EU citizens in acquiring foreign languages (European Commission, 2022). However, Phillipson (2017) argues that despite this socio-cultural rhetoric, EU language policy has never been serious about multilingualism and that policy is actually driven by hegemonic market forces. This economic orientation to EU language policy was also identified by Krzyzanowski and Wodak (2011) who found that the EU's Knowledge Based Economy (KBE) macro-strategy of the new millennium has dominated EU language policy and silenced previously championed social and democratic arguments. In short, while EU discourse may overtly foreground social dimensions, macro-strategic aims advance an economic orientation to language policy. This economic orientation has been keenly been felt in HE and research contexts where, through the processes of internationalisation and marketization, the neo-liberal values of competitiveness, commodification and consumerism (Evans et al., 2018) have conceptualized language — and in particular English — as an important economic tool.

2.2.2 ELF and European higher education and research

In this study, ELF is defined as a contact or vehicular language between users of different first languages, including native-speakers of English (Jenkins et al., 2017). Motivated by communicative need rather than linguacultural factors (Seidlhofer et al.,

2006), ELF is used in a number of European education and research contexts. Non-Anglophone universities have integrated English in aspects of instruction, research and administration (Villares, 2019) and in academic publishing, English serves as "a cross-linguistic or international medium of communication" (Seidlhofer et al., 2006, p. 7). Despite, this prolific use, there has been an explicit absence of overt EU policy as it pertains to ELF in higher education and research contexts (Hultgren et al., 2015) and Soler-Carbonell et al., (2017) note that in the context of the KBE agenda and the internationalisation of European HE, "language does not seem to play a very significant role for EU policy makers" (p. 309). Cogo and Jenkins (2010) suggest that the absence of ELF discourse is a result of EU policy-makers' lack of awareness of how English has developed in the world as a *lingua franca* over the past few decades. Rather than understanding that "the most typical use of English in Europe as a whole is as a lingua franca, i.e. ELF" (Cogo & Jenkins, 2010, p. 279), EU policy-makers continue to view English in a "nation-state sense" (Cogo & Jenkins, 2010, p. 279) as either a native or foreign language. Consequently, language policy management in the EU remains politically sensitive and the expansion of English within the EU is viewed as a form of linguistic imperialism (Phillipson, 2017) or as a vector for inequality and domination through its association with global capitalism (Ferguson, 2018).

2.3 Institutional actors in context

Within the EU, the EHEA and the ERA have been described as two pillars of the knowledge-based society (Koch Christensen, 2005). As such, they offer an appropriate comparative context to situate this study as two distinct EU actors with remits for

education and research. The EHEA is an initiative of the Council of the European Union (consisting of portfolio-specific government ministers from each member state), while the European Commission (the EU's politically independent executive arm) drives ERA policy. Despite their different governance structures, they have overlapping remits for doctoral education and research policy and have a number of shared goals with language policy implications (Foroni, 2015). For example, both actors are concerned with researcher mobility; knowledge flow; and cooperation among universities, business and other research actors (European Commission, 2021).

2.3.1 The EHEA

Since 1999, HE institutions across Europe have been involved in a major process of reform, restructuring and harmonization as part of the Bologna Process (Tudor, 2005). The Bologna Process was seen as a strategic response to the exigencies of globalization and the EHEA, which was officially launched in 2010, can be characterized as the institutionalization of that process (Neave & Veiga, 2013). Structural, organizational and terminological reforms were the primary foci of the Bologna Process (Neave & Veiga, 2013) and these have been the most successful policy areas of the EHEA (Curaj et al., 2009). However, despite the significant linguistic consequences of these reforms, it has been widely noted that the 1999 declaration, which began the Bologna Process "did not devote a single word to language-related issues" (Hultgren et al., 2015, p. 7). This is despite the fact that the internationalization agenda of the Bologna Process was broadly seen to strengthen the position of English relative to other languages in higher education contexts (Phillipson, 2015).

2.3.2 The ERA

The ERA was formed in 2000 by the European Commission with the aim of creating an internal market of research. It thus manifested long-held European fears over research competitiveness with other global research actors such as the USA and Japan (de Alera, 2006). Since its inception, however, it has struggled with the competing interests of EU member states and institutional inertia at the European level (Banchoff, 2002) and it's relaunch in 2020 can be seen as a renewed institutional effort to re-focus its agenda. In effect, the ERA is built as part of the KBE and stress is placed on the economic side of research "understanding it mainly as a tool for economic development and neglecting its social dimension" (De Alera, 2006, p. 573).

2.4 Methodology

In order to compare the discursive representations, associations, and attitudes toward ELF in the discourse of the EHEA and the ERA, I employ methods that draw on both Critical Discourse Analysis (CDA) and Corpus Linguistics (CL) as described by Baker et al. (2008). CDA is an approach well suited to this study because it is particularly interested in how social institutions mediate and reproduce dominant ideologies through linguistic and semiotic practices (Wodak, 2001). Combining CDA approaches with CL provides a useful "methodological synergy" (Baker et. al, 2008, p.274) in which CL can be used to add levels of objectivity to the study while at the same time benefitting from the explanatory power of CDA theoretical frameworks (Gabrielatos & Baker, 2008). In addition, when combined, these two methodological approaches are well placed to deal with an already established feature of EU language policy — namely, absence. Textual

absence is seen as particularly problematic for CL approaches, however comparative

analyses, such as the one employed here, are well suited to interpret such absences

(Duguid & Partington, 2018). Moreover, critical approaches recognise that the norms and

values, which underlie texts, are often out of sight rather than overtly stated (Baker et. al,

2008).

2.4.1 Focal Corpora

In order to compare EHEA and ERA discourse on ELF, two comparable corpora

were compiled. According to Randour et al. (2020) research on political discourse tends

to focus on the political conditions in which the discourse is produced rather than the

communicative context of the discourse. In order to attend to both, the two focal corpora

for this study are compiled of publically available written declarations, communications

and reports which share the communicative purpose of informing EU stakeholders of

EHEA and ERA policy positions and initiatives between 1998 and 2020. The EHEA

corpus consists of all Bologna Process or EHEA declarations and communiqués issued by

the European Council of Ministers from 1998 to 2020 (see Appendix 2.1). At 25,941

words, it places at Aston's (1997) lower end of the range defining small, specialized

corpora. All documents were downloaded in their English version through the publication

office of the European Union and converted to plain text files for further analysis. The

ERA corpus consists of all ERA communications and reports issued by the European

Commission from its inception in 2000 to 2020 (see Appendix 2.2). To be included in the

corpus, the communication was required to have *European Research Area* or *ERA* in the

title of the communication and more general Commission communications on research

i.e. "*A renewed European agenda for research and innovation - Europe's chance to*

shape its future" and "*Partnering in Research and Innovation*" were excluded, as their scope went beyond ERA policy-making to include other frameworks and stakeholders. At 120,159 words, the ERA corpus is considerably larger than the EHEA corpus. ERA documents are published more frequently and are generally lengthier. However, it is still placed at Aston's (1997) lower-end of the range defining small, specialized corpora. All documents were downloaded in their English version through the publication office of the European Union and converted to plain text files for analysis.

2.4.2 Reference Corpus

In addition to the focal corpora, the study also uses a 'reference' corpus. The Eur-Lex (English) Corpus (Baisa et al., 2016) is a general corpus of EU legal documents including treaties, legislation, preparatory documents, parliamentary questions, and the official journal of the European Union (516,897,745 words). It therefore acts as a useful baseline of 'general' EU discourse against which the specialized EHEA and ERA focal corpora can be compared. To ensure comparability, the Eur-Lex (English) corpus is narrowed to the time period covering the creation of the EHEA and ERA (1998 onwards).

2.4.3 Corpus Analysis

Data were analysed using the corpus analysis tool SketchEngine, which allows for the analysis of lexical and lexico-grammatical patterns (McEnery & Hardie, 2012). The corpora were searched for tokens of *English, lingua franca*, *language*, *linguistic, and lingua** and frequencies were normalized per 10,000. The identified tokens of these terms in the focal corpora were coded according to Spolsky's (2004; 2019) domains of

language policy as either instances of language ideology, language practice, or language management. Next, in order to identify topics frequently associated with the search terms, collocation analysis was performed using SketchEngine's word sketch function. Collocation refers to the "above-chance frequent co-occurrence of two words within a pre-determined span of the node" (Baker et. al, 2008, p. 278) and thus provides information about the most salient ideas associated with the node (word in focus) (Gabrielatos & Baker, 2008). The analysis is statistically determined and takes into account the frequency of the node, the collocate, and the collocation. In this study, the node span was set at five words to the left and right of the node. Finally, tokens were analysed at the discourse-semantic level and coded by attitudinal stance. Stance can be expressed to differing degrees by using value-laden word choice, grammatical devices, and paralinguistic features (Biber, 2006). In this study, I focus on semantic and grammatically encoded attitudinal stance markers such as stance nouns (i.e. cooperation/barrier), stance verbs (i.e. foster/block) attitudinal adjectives (i.e. unequalled/unwieldy) and attitudinal adverbs (i.e. importantly/worryingly), which were coded along a positive/negative dimension. To mitigate interpretive bias, the data were coded for language policy domain and attitudinal stance by two additional coders. The coders were given a coding sheet (see Appendix 2.3) that provided explanations and examples of the codes to be used. A short calibrating session was also held to discuss the coding sheet and the examples. In the final analysis, the codes applied by coder 2 were excluded from the study as their coding was widely discrepant from the other two coders. More thorough training and calibrating sessions might have avoided these inter-coder agreement issues.

2.5 Analysis

The purpose of this study was to conduct a comparative analysis of ELF discourse in the published declarations and communications of the EHEA and the ERA as well as in general EU discourse. Analysis shows that explicit ELF discourse is absent from EHEA and ERA communications. Implicit stances to ELF are however taken in more general discourse on language policy. These implicit stances reveal intra-institutional divergence to language-related matters along social and economic dimensions. Moreover, these two contrasting stances are replicated in explicit representations of ELF, which do appear in general EU discourse over the same time period. The next section will present an analysis of these direct references to ELF in general EU discourse before turning attention to the covert stances taken to ELF by the EHEA and the ERA.

2.5.1 ELF in general EU discourse

Although ELF discourse is absent from the policy documents of the EHEA and ERA, tokens of *English as a lingua franca* and *lingua franca* were found in the reference corpus. Concordance lines were examined in detail to identify how and by whom the terms are used in general EU discourse. This was done to identify how implicit stances to ELF in the focal corpora may align or diverge from general EU discourse on ELF over the same time period. Thirty-three tokens of *lingua franca* were found in the reference corpus across 27 documents. Fifteen tokens appear in European Commission texts, nine tokens in European Parliament texts, 4 tokens in European Economic and Social Council (EESC) texts, three tokens in Court of Justice texts, and two tokens in Committee of the Regions (CoR) texts. Given the central role of the European Commission in driving ERA

policy and its increasing influence in EHEA policy decision-making (Phillipson, 2015),

the following analysis focuses on the representation of *English as a lingua franca* and

lingua franca in European Commission texts and other tokens directly connected to

education and research contexts.

2.5.1.1 ELF in European Commission discourse

Lingua franca is used in European Commission documents on general language

policy (5 tokens); language policy in the IT sector (5 tokens); language policy in the

education sector (4 tokens), and a Commission Decision on a financial merger (1 token).

Notably, the only text that discusses *English* as the *lingua franca* of the academic

community is not an education policy document, but rather a Commission Decision,

which adjudicated the merger of European publishing companies (extract 1).

(1)　According to the parties, academic publishing encompasses all academic
disciplines, including for example technology, life sciences, earth sciences,
economics, humanities and other social sciences. **Academic publications in
these subjects are very often in <u>English</u>, as the <u>lingua franca </u>of the
<u>academic community</u>** whereas publications for example in the field of law
normally focus on a specific national jurisdiction and are therefore published
in the language used in this country."

*Commission Decision of 29/07/2003 declaring a concentration to be compatible with the
common market (Case No COMP/M.3197 - CANDOVER / CINVEN / BERTELSMANN-
SPRINGER) according to Council Regulation (EEC) No 4064/89*

Unlike the certainty given to English as the *lingua franca* of the academic community in

extract 1, when *lingua franca* is used in the Commission's Action Plan on Language

Learning and Linguistic Diversity, the identity of the *lingua franca* is ambiguous (excerpt

2) or there is explicit dispreference for the dominance of only one *lingua franca* (excerpt

3).

(2)   The Socrates programme's Lingua action 2 will fund a series of
       transnational projects to develop materials for teaching language awareness
       and foreign languages ***other than lingua francas***

*Commission Document Promoting Language Learning and Linguistic Diversity: an
Action Plan 2004 – 2006*

(3)   Language skills are unevenly spread across countries and social groups. The
       range of foreign languages spoken by Europeans is narrow, being limited
       mainly to English, French, German, and Spanish. ***Learning one lingua franca
       alone is not enough***. Every European citizen should have meaningful
       communicative competence in at least two other languages in addition to his or
       her mother tongue.

*Commission Document Promoting Language Learning and Linguistic Diversity: an
Action Plan 2004 – 2006*

Indeed, all five tokens of lingua franca in European Commission language policy

documents appear in the phrase, "one lingua franca alone is not enough" or discuss

language learning "other than lingua francas". Moreover, *lingua francas* are framed as

foreign languages, preserving the native vs. foreign dichotomy of language. This is in

contrast to Commission documents that discuss IT policy, where four of the five tokens

explicitly name English as a lingua franca of technology. This analysis demonstrates that

within the European Commission itself stances toward ELF vary across Commission

portfolios. Education and language policy documents express a dispreference or

ambiguity toward the dominance of one lingua franca whereas IT and business merger

policy documents make explicit reference to English as *the* lingua franca. These stances

28

can be characterized as manifestations of the two competing language ideology

discourses — internationalist vs. culturist — identified by Hultgren et al. (2014).

2.5.2 ELF in EHEA and ERA discourse

Over the twenty-year period of analysis, explicit ELF discourse is absent in the

policy documents of the EHEA and the ERA. However, the two actors do discuss

language-related matters, albeit with different frequency and focus. As is shown in Table

2.1, tokens of language or linguistic appear more frequently in the EHEA corpus (4.62

per 10,000 words) compared to the ERA corpus (1.33 per 10,000 words). Comparing

frequency in the focal corpora to the reference corpus, language and linguistic appear

more frequently in the EHEA corpus and less frequently in the ERA corpus compared to

general EU discourse (2.58 per 10,000 words) over the same time period.

Table 2.1: Frequency analysis of lemmas language and linguistic in the focus and
reference corpora

	ERA	EHEA	Eur-Lex
all tokens	16	12	133,307
per 10,000 words	1.33	4.62	2.58
language (raw)	12	8	126,480
per 10,000 words	1.0	3.08	2.45
linguistic (raw)	4	4	6,827
per 10,000 words	0.33	1.54	0.13

In addition to observable differences in frequency, the tokens of *language* and *linguistic*

are used in distinctive language policy domains across the two focal corpora. Using

Spolsky's (2004; 2019) classification, in the EHEA corpus, nine of the twelve tokens of

language or *linguistic* were coded as language ideology (excerpt 4).

(4) **We believe** that mobility of students, early stage researchers and staff enhances the quality of programmes and excellence in research… It encourages **linguistic** pluralism, thus underpinning the multilingual tradition of the European Higher Education Area.

Leuven Communiqué, The Bologna Process 2020 - The European Higher Education Area in the new decade, 2009

However, in the ERA corpus, tokens of *language* or *linguistic* are exclusively used in stretches of text related to either language practice (excerpt 5) or language management (excerpt 6).

(5) As regards the presentation of the professional experience, **current practice** shows that a researcher's CV normally requires the listing of the different experiences in chronological order with precise references to education and work experiences, additional skills, proficiency in foreign **languages** etc.

Researchers In The European Research Area: One Profession, Multiple Careers, 2003

(6) **Potential measures** to further facilitate the international mobility of researchers include equal access to national research funding programs for foreign researchers, and increasing the portability of research grants. Additional measures include the further development of human resources procedures in research performing institutions. Pension right transferability and **language** competency for teaching requirements are evolving topics.

ERA Progress Report, 2016

In sum, the representations of language in the focal corpora are distinct, with the EHEA focusing on language ideology and the ERA focusing on language practice and management. Finally, the EHEA discusses language matters to a greater extent and the ERA to a lesser extent than general EU discourse.

2.5.3 Collocation of language and linguistic in EHEA and ERA discourse

Given the absence of explicit ELF discourse, collocation analysis was instead carried out on tokens of language and linguistic. In the EHEA corpus, the strongest collocate of language is culture and over half of all tokens of linguistic collocate with diversity (extract 7).

(7)   to preserve Europe's **cultural richness** and <u>**linguistic diversity**</u>, based on its heritage of diversified traditions, and to foster its potential of innovation and social and economic development through enhanced co-operation among European Higher Education Institutions.

Berlin Communiqué, 2003

In the ERA corpus, however, the strongest collocates of language were competency, regulation or, requirement (excerpt 8) and tokens of linguistic collocated with obstacle or barrier (excerpt 9).

(8)   Europe does not offer researchers from third countries particularly advantageous (material and administrative) conditions. The formalities to be completed are generally unwieldy. The **regulations** and <u>**languages**</u> also vary from one country to another. And the "brain drain", which some have claimed is being held in check, has not stopped.

Towards a European Research Area, 2000

(9)   It is thus essential to establish a single and open European labour market for researchers, ensuring effective "brain circulation" within Europe… At the same time, public authorities and research institutions need to work to remove the legal, administrative and practical (e.g. <u>**linguistic**</u>) **barriers** to geographical and inter-sectoral mobility…

The European Research Area: New Perspectives 2007

The collocation analysis reveals additional differences in the discourse of the two actors toward language-related matters in European HE and research. While the EHEA collocates language with a social dimension of culture and values, the ERA positions language as an economic skill or tool. Table 2.2 shows the frequency of these collocations in the focal corpora compared with the reference corpus. We can see that language as an extension of culture appears more frequently in the EHEA corpus compared to general EU discourse and language as an economic tool appears more frequently in the ERA corpus than general EU discourse.

Table 2.2: Collocates of *language* in the focal and reference corpora

	ERA	EHEA	Eur-Lex
language + culture (raw)	0	4	499
per 10,000 words	-	1.54	0.01
language + competency (raw)	4	0	484
per 10,000 words	0.33	-	0.01

It would appear that the focal actors not only take opposing positions on the 'language as an extension of culture' vs. 'language as a tool' dichotomy, but that they do so more markedly than positions taken in general EU discourse across the same time period.

2.5.3.1 Mobility co-text

During the analysis, it also became evident that, while not a collocate, mobility frequently appeared in the co-text of tokens of language and linguistic in both of the focal corpora. As previously noted, researcher mobility is a key aim of both the EHEA and ERA. In the EHEA corpus, mobility appears 41.2 per 10,000 words and in the ERA

corpus 20.6 per 10,000. In the Euro-Lex corpus, mobility occurs 0.73 per 10,000 words

(see Table 2.3).

Table 2.3: Frequency analysis of mobility in the focus and reference corpora

	ERA	EHEA	Eur-Lex
mobility (raw)	248	107	37,991
per 10,000 words	20.6	41.2	0.73

Despite the prominence of mobility in the focal corpora and its frequent appearance in the

co-text of tokens of *language* and *linguistic*, there are key differences in the semantic

relations of language and mobility in the focal corpora. For example, in Excerpt 4 from

the EHEA, mobility "encourages linguistic pluralism" whereas in Excerpt 9 from the

ERA, language "is a barrier to effective brain circulation". In short, even in an area of

mutual interest, the EHEA and the ERA frame the relationship between language and

mobility quite differently. In the ERA corpus, language is a barrier to mobility whereas in

the EHEA corpus, mobility is a key to language. This distinct framing is also seen in the

attitudinal stances taken by the EHEA and the ERA to language and linguistics more

generally.

2.5.4 Attitudinal stance to language and linguistic in EHEA and ERA discourse

Negative attitudinal stances were taken in eleven of the sixteen tokens of

language and linguistic in the ERA corpus. This is done overtly (excerpt 10) or through

the use of metaphors indexing difficulty (excerpt 11).

(10)   **On the negative side**, discussions on the proposal for the creation of a
       Community Patent are still blocked in the Council, the main points of

disagreement being **language use** and translation arrangements.

The European Research Area: Providing New Momentum Strengthening - Reorienting - Opening Up New Perspectives, 2002

(11)   However, in practice such movement is very limited due to **obstacles of various types**: legal, administrative and regulatory, practical, cultural and **linguistic**, information related, etc. The activities undertaken on this theme aim to remove these **obstacles**.

The European Research Area: Providing New Momentum Strengthening - Reorienting - Opening Up New Perspectives, 2002

However, in the EHEA corpus, salience is given to the benefits and opportunities of

language and linguistic and overt positive stances are taken in ten of the twelve tokens

(i.e. excerpts 12 &13).

(12)   Ministers reaffirmed that efforts to promote mobility must be continued to enable students, teachers, researchers and administrative staff to **benefit from the richness of** the European Higher Education Area including its democratic values, diversity of cultures and **languages** and the diversity of the higher education systems.

Prague Communiqué: Towards the European Higher Education Area: Communiqué of the meeting of European ministers in charge of higher education, 2001

(13)   The EHEA is a unique cooperation, built on trust, where public authorities and higher education stakeholders work together to define and achieve shared goals. **Thanks to the diversity of** our cultures, **languages** and environments, and to our shared commitment to quality, transparency and mobility, our higher education systems offer **unequalled opportunities** for learning, teaching, research and innovation.

Rome Communiqué, 2020

In summary, although explicit ELF discourse is absent from the focal corpora, implicit

stances to ELF are taken in the actors' representation, association and attitudes towards

language-related matters. The EHEA articulates frequent positive attitudes towards its language beliefs and foregrounds linguistic pluralism and diversity of languages as beneficial to European HE and to the mobility of researchers. On the other hand, the ERA discusses language policy to a lesser extent and when doing so foregrounds the negative aspects of language practices and management issues in Europe — viewing language as a barrier and obstacle to EU research and the mobility of researchers.

2.6 Discussion

This study used the discourse of two EU policy actors as 'entry points' to uncover the ideological positions and potential intra-institutional differences in European ELF policy at the supranational level. The analysis reveals how different actors within symbolic elites, such as the EU, attempt (both overtly and covertly) to construct knowledge, reproduce their chosen ideologies, and ultimately control public discourse (van Dijk, 2001) on an issue such as ELF. In general EU discourse, particularly that of the European Commission, the role of lingua francas is explicitly discussed but direct discussion of English as the or even a lingua franca is confined to IT and business portfolios, with education policy documents preferring not to specify the lingua francas or stating an explicit dispreference for the dominance of only one lingua franca.

Importantly, over the institutional lifetime of the two supranational actors with remits for HE and research, English as a lingua franca, the role of English in HE contexts, or the role that lingua francas might play in HE and research environments do not appear in any official policy communications. Given the increasing dominance over this time period of English in non-Anglophone European academic and research environments

(Mauranen et al., 2010), these discoursal absences are conspicuous. Arguments made by Cogo and Jenkins (2010) over a decade ago for the integration of ELF into EU language policy frameworks continue to be disregarded. However, this study also provides some nuance to claims that language-related matters are absent from EHEA discourse (Hultgren et al., 2015). As previously noted, the perceived absence of a linguistic phenomenon can only be appraised by comparing that absence with the behaviour of the phenomenon elsewhere (Duguid & Partington, 2018). While language matters might not be as central to EHEA policy as applied linguists might expect, the frequency analysis shows that language and linguistic appear more frequently in EHEA discourse compared to general EU discourse and ERA discourse over the same time period. In other words, the EHEA actually foregrounds language related matters compared to an analogous institutional actor as well as the larger institutional organism in which it is nested. Nonetheless, I do recognise that EHEA discourse foregrounds more abstract ideological positions to language with less focus on language management and no discussion of on-the-ground practice. Therefore, we can say that EHEA discourse disregards the actual use of languages/language habits in different academic contexts, which perhaps does support the claim that the EHEA has a disregard for the linguistic implications of its policies (Hultgren et al., 2015)

Finally, it is important to emphasize that the two EU actors take very different stances to general language policy. The analysis revealed important differences in the frequency, focus, associations, and stances to language by the two actors. These differences can be characterized along economic and social dimensions with the ERA

explicitly foregrounding the instrumental nature of language and the EHEA

foregrounding cultural dimensions. Moreover, while the EHEA adopts a positive stance

to the culturist dimension, the ERA adopts a negative stance to language-related matters.

These different stances may well be reflective of the actors' institutional footing within

the EU. EHEA policy is still largely determined by powerful individual nation-state

interests, which are more inclined, on the surface at least, to foreground the benefits of

linguistic and cultural diversity. On the other hand, the ERA's more technocratic, and

arguably less powerful institutional footing may influence the instrumental approaches

yet negative stances to language. In other words, the ERA have a closer connection to on-

the-ground practices but less power to affect change. In sum, this study adds nuance to

our understanding of how the absence of EHEA language policy might be interpreted

while also revealing novel perspectives to intra-institutional tensions in the EU —

between actors and across portfolios — on the role of English in HE and research

contexts.

To what extent the ideological positions and stances of the EHEA and ERA

actually affect doctoral student experience, is beyond the scope of this paper, and

scholars have noted a disconnect between top-down ideology and on-the-ground practices

(Hultgren, 2014). An example of this disconnect is perhaps illustrated by the fact that

while the EHEA's Erasmus Programme (European Commission, 2022) is touted to

promote linguistic diversity and language learning "so that students may achieve their full

potential for European identity, citizenship, and employability" (Berlin Communique,

2003), the research scientists who acted as informants for the next chapter of this

dissertation reported that their participation in the Erasmus Programme, in which they spent time studying in another non-Anglophone European country, actually helped them to learn English – and English alone. This doesn't quite reflect the spirit of the Berlin Communique's linguistic diversity rhetoric and it also suggests that speakers are not only using but also acquiring English in ELF contexts. Future studies that seek to link ideological positions and on-the-ground practices of ELF in Europe might investigate the language experiences of post-doctoral students with current EHEA and ERA led initiatives and programmes. Alternatively, interviews could be carried out with European researchers to investigate how language policies impact their experiences as early-careers scientists.

In the next chapter, I report findings from my own semi-structured interviews with a group of internationally mobile research scientists from the EU. While the interviews did touch on the language experiences of these scientists in EU funded initiatives, the analysis in Chapter III more closely focuses on the construction of their ELF and disciplinary identities. Moreover, as I designed the study I began to think carefully about how my own language ideology might inform the methodological approach I took to the interviews. In the next chapter, I outline how the CA concepts of action and understanding provided the methodological lens to focus the research interviews as a social interaction.

CHAPTER III

MESO PERSPECTIVE: ACTION AND UNDERSTANDING IN THE SEMI-STRUCTURED RESEARCH INTERVIEW: USING CA TO ANALYSE EUROPEAN RESEARCH SCIENTISTS' ELF AND DISCIPLINARY IDENTITIES

3.1 Introduction

In Chapter II, I discussed how semi-structured research interviews with doctoral researchers and early-career scientists might provide useful insights to possible disconnects between language ideology and on the ground language practices of their discourse communities. Interview is commonly used in social science research to gain so-called emic or insider perspectives to the norms and practices of various discourse communities (Olson, 2016). Within the EAP literature, the semi-structured research interview is often used within a process of triangulation as a method that can enrich and extend understandings (i.e. Li et al., 2018; Zhang et al., 2020) or allow the researcher to gain special insights or insider views (i.e. Davis, 2019; Neumann et al., 2019; Hynninen & Kuteeva, 2017. Moreover, in a recent longitudinal (2005-2016) analysis of research trends in the broader field of Applied Linguistics, the semi-structured interview was found to have noticeably increased as a research method during the study's analysis period (Lei & Liu, 2018).

While the semi-structured research interview has become a central method of inquiry in EAP research, recent qualitative methodologies' scholarship has raised concerns that simplistic approaches to the research interview are pervasive and that interview data is both over-used and under-analysed (Silverman, 2017; Whitaker &Atkinson, 2019). These concerns are by no means new. Holstein and Gubrium (1995) critiqued the 'neutral' interviewer/answer-vessel' respondent model of interviewing in their characterization of the 'active interview' (1995). From an active interview perspective, the interview is a site for the production of meaning in which active subjects construct versions of reality interactionally (Gubrium & Holstein, 2001). Talmy (2010) raised concerns about approaches to interview in Applied Linguistics when he questioned whether the field's conceptualization of interview as a research instrument really did provide emic perspectives and discussed the advantages of reconceptualizing the research interview as a social practice. As an approach more centrally concerned with the social and interactional aspects of research interviews, Conversation Analysis (CA) (Sacks et al., 1974) has proven to be an important analytical approach in reconceptualizing research interviews as instances of social action rather than "neutral or unmediated conduits into participants' inner psychological worlds" (Prior, 2018, p. 491). Its analytical focus on the engagement of interview participants in interactional activities and with the overall structure of interactions (Wooffitt & Widdicombe, 2006), led it to become a productive site for interview research in Applied Linguistics (e.g. Gardner, 1994; Kasper & Wagner, 2004; Prior, 2018; Prior & Talmy, 2019; Roulston, 2011; Sert & Seedhouse, 2011). However, despite being an active area of scholarship in the sister field of Applied

Linguistics, the interactional nature of the research interview remains largely over-looked in EAP/ESP research. Few recent studies (e.g. Davis, 2019; Hynninen & Kuteeva, 2017; Kaufhold & McGrath 2019; Stenglin & Cléirigh 2020; Xu & Zhang, 2019) examine interview data in its interactional context and even fewer studies use CA to transcribe and analyse interview data (e.g. Han & Hyland, 2019). At a time when EAP/ESP researchers are being asked to more fully engage with producers of texts through interview, it seems appropriate to give renewed consideration to the benefits of adopting more interactionally grounded methods of analysis. This is particularly true for studies that seek to examine issues of identity and self, which are "handled in the turn-by-turn production of activities in interaction." (Wooffitt & Widdicombe, 2006, p. 49).

Therefore, the purpose of this study is to examine how the CA concept of *action and understanding* provides methodologically robust and insightful approaches to data from a series of semi-structured interviews with European research biologists who write in English for research publication purposes. A CA approach is particularly appropriate for this data set given that the interviews discuss the scientists' identities as ELF users and their attitudes to writing in their disciplinary communities.

3.1.1 Research Questions

This study is guided by the following research questions:

i.   In what way can CA be beneficial in the analysis of EAP research interview data?

ii. In what way does an interactionally grounded approach to interview data provide insight into the ELF and disciplinary identities of a group of European research

scientists?

3.2 Research Interviews and Interactional Context

The next section of this paper (i) defines more closely the speech event of the research interview, (ii) elaborates on the central role of questions in that speech event, (iii) discusses why the conversation analytical lens of action and understanding is appropriate for the analysis of interview data, and (iv) introduces the significance of intersubjectivity in the analysis of interview data.

3.2.1 Semi-structured Interviews

Interview is a distinct type of speech event (Moder & Halleck, 1998) with a distinct organisational structure and a distinct relationship between participants. Semi-structured interview is commonly used in applied linguistics' research (Lei & Liu, 2018). In these interviews, structured/ pre-planned questions are used to elicit specific information, but the interviewer is also said to allow for a responsive exploration of "the emerging worldview of the respondent, and to new ideas on the topic" (Merriam, 1998, p. 74). This orientation for pre-planned questions and emerging ideas, results in a distinct organisational pattern. For instance, in a semi-structured interview, the interviewer might avoid question and answer patterns characteristic of a more structured interview in order to elicit longer turns from the interviewee (Schiffrin, 1994). Indeed, Merriam (1998) suggests polar (yes/no) questions should be avoided in the semi-structured interview in favour of other question designs that prompt 'fuller' responses. In addition, there is usually an asymmetry of power and/or knowledge between participants (Schiffrin, 1994),

with the interview frame itself constraining the interactional roles of interviewer and interviewee (Moder & Halleck, 1998). Thus, the semi-structured interview might be said to function as a hybrid or mixed speech event that exhibits characteristics similar to both institutional talk *and* conversation (Schiffrin, 1994). As a result, analyses of semi-structured interview data focusing solely on respondents' turns as "telling quotations" (Silverman, 2017) reveals a rather distorted picture of the speech event and ignores the central role of questions in the ongoing action of the semi-structured interview.

3.2.2 Questions in Semi-Structured Interviews

Qualitative methods' textbooks discuss question design and sequencing in some detail (Flick, 2009; Merriam, 1998). However, emphasis is often placed on questions as stimuli that produce desired responses for subsequent analysis, rather than activities in an ongoing social interaction. From a CA point of view, however, questions do more than request a response; they implement a range of social actions that make a response relevant (Stivers & Rossano, 2010). More specifically, questions advance two agendas: a topical agenda (i.e. what is being talked about) and an action agenda (i.e. what the speaker is doing with the question) (Hayano, 2013). Action agendas have received much attention in the CA literature (Pomerantz & Heritage, 2013), particularly in relation to question preference (Sidnell, 2010). On the other hand, topical agendas, as they relate to interviews, have received less attention. In semi-structured research interviews, those asking the questions set the topical agenda; therefore, the interviewer's access to the epistemic domains of the interviewee becomes extremely important in the ongoing interaction. This is particularly relevant to EAP/ESP researchers (community outsiders)

aiming for insight into a discourse community's ways of knowing and doing by interviewing community insiders. However, this aspect of interview is not taken into account if we view questions as stimuli, which produce free and open responses as opposed to vehicles of social interaction, which constrain recipient responses through action and topical agendas. The next section of the paper will discuss in more practical terms how the CA lens of action and understanding might help EAP/ESP researchers take a more insightful and systematic approach to aspects of their interview data.

3.2.3 Action and Understanding in Semi-Structured Interviews

CA provides a useful approach to the study of interview data, as it focuses on the structural underpinnings of naturally occurring language and assumes that language use is orderly "at a minute level of detail" (Sidnell & Stivers, 2013, p. 36). In other words, while other analyses may focus on what the discourse is *about*, CA takes a more structural approach to analyse what the participants are *doing*. Indeed, CA views any single fragment of conversation as "the unique product of multiple, intersecting organizations of practices" (Sidnell, 2010, p. 249), which can actually reveal more about the beliefs and motivations of the speaker than what is said alone. It follows, therefore, that interview data should be transcribed and analysed in such a way that allows for the detailed analysis of the interactional work of participants. However, in other more circumscribed approaches to interview data, these features are often disregarded from the analysis as 'noise' or unremarkable features of the talk (Sidnell, 2010). Second, CA is concerned with the "knowledge claims that interactants assert, contest and defend in and through turns-at-talk" (Heritage, 2013, p. 370). A key concept here is intersubjectivity, or

the mode by which understanding is produced and owned by participants *in* interaction

(Heritage & Stivers, 2013). In other words, in talk-in-interaction such as the semi-

structured interview, "each utterance displays a hearing or analysis of a preceding one"

(Sidnell, 2010, p.12). Consequently, question and answer adjacency pairs provide an

ideal framework to analyse the intersubjective understandings of interview participants as

they exhibit both prospective and retrospective dimensions (Sidnell, 2010), with the

second pair part "displaying a speaker's understanding of the first to which it responds"

(Sidnell, 2010, p. 66). It follows, therefore, that questions should not be disregarded as

stimuli for response but integral to the intersubjective analysis of interview as a social

action.

3.2.4 Intersubjectivity in Semi-Structured Interviews

Benveniste (1971) theorized that the intersubjective nature of language made

linguistic communication possible. Central to this is the notion of subjectivity.

Subjectivity is the ability of speakers to posit themselves as 'subjects' and in expressing

themselves as *I* in the discourse, they construct an interior-exterior opposition between

themselves and the other person who, as the echo of the subject, becomes at once the *you*

for the *me* and the *I* for themselves (Benveniste, 1971). Closely related to subjectivity is

the concept of stance as speakers use linguistic subjectivity "to express their perceptions,

feelings, opinions and evaluations in discourse" (Baumgarten & House, 2007, p. 195).

The most basic and prototypical source of subjectivity in spoken discourse are subject-

predicate combinations that allow speakers to personalize their talk, to mark attitude,

evaluation and empathy (Scheibman, 2002). In her monograph of epistemic stance in

English conversation, Kärkkäinen (2003) noted the strong tendency toward the routinization of subject-predicate combinations such as *I think*, *I don't know if* and *I guess*. Kärkkäinen (2003) goes on to argue that these subject-predicate combinations are developing into discourse markers as their semantic meanings cannot be aligned with their interactional and pragmatic functions. Schiffrin (1987) defines discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31) thus providing "contextual coordinates" (Schiffrin, 1987, p. 41) for the ongoing interaction. Subject-predicate discourse markers such as *you know* and *I mean* are prolific in everyday conversation and, through their expanded pragmatic use, have undergone a type of sematic bleaching as part of a process of grammaticalization (Clayman & Heritage, 2021). However, contrary to previous characterisations of these discourse markers as randomly placed filler phrases lacking any communicative value, Fox Tree and Schrock (2002) suggest that the prolific use of these discourse markers in spontaneous speech can be attributed to each marker's basic meaning. Various accounts have been provided for basic meanings of these discourse markers. Ostman (1981) suggests that a speaker uses *you know* as they "strive toward getting the addressee to cooperate and/or to accept the propositional content of his utterance as mutual background knowledge" (p. 17) while Jucker and Smith (1998) suggest that *you know* invites addressee inference. More recently, Clayman and Raymond (2021) have sought to provide a unified account of how *you know* functions in ordinary conversation. They suggest that in its basic form, *you know* works as an intersubjective or affilliative token of alignment in environments where understanding and affiliation have become salient or potentially problematic (Clayman

46

and Raymond, 2021).

In her analysis of *I mean,* Schiffrin (1987) suggests that it forewarns the hearer of upcoming adjustments, by way of expansions on or explanations of, prior talk. In this sense, it functions as a particular type of self-repair that displays the speaker's intention to shift focus (Schiffrin, 1987). Scholars also suggest that native speakers and non-native speakers of English use these intersubjective discourse markers differently; thus suggesting ELF-specific practices of stance taking (Baumgarten & House, 2007 & 2010; House, 2009 & 2013). In an analysis of *I think* and *I don't know* in English L1 and ELF conversation, Baumgarten and House (2010) found that the discourse markers exhibited only partially overlapping functional profiles, with ELF users more likely to use them in their prototypical sense to mark opinion. Moreover, House (2009) suggests that *you know* has lost its 'other oriented' meaning and has been re-interpreted in ELF contexts as a special type of self-referencing connective used to preface formulation difficulties and aid semantic coherence between propositions. Finally, rather than a marker of reformulation/clarification, it is suggested that in ELF discourse *I mean* signals an emotional involvement and acts as a "focalizing device in the speaker's contribution which serves as the point of departure for an explicit expression of a subjective evaluation" (Baumgarten & House, 2007, p. 208). Whether we view these discourse markers as alignment tokens, inference invitations, or forewarnings, it seems clear that rather than dismiss them as filler phrases in the data transcription process, *how* interview participants use them could provide richer insights into *what* interview participants say. Moreover, while much of the existing research on *I think*, *you know* and *I mean* has

discussed their functional import in everyday conversation, their use in research interview environments remains underexplored. In summary, through the detailed transcription and analysis of the interactional work done by interview participants in question and answer adjacency pairs, the CA lens of *action* and *understanding* and the concept of intersubjectivity could help EAP/ESP researchers defend their methodological practices from existing critiques and also enrich their access to the worldviews of their participants.

3.3 Method

In this study, I use CA to analyse question and answer adjacency pairs from semi-structured interviews with six European research biologists. Although CA is a method more traditionally used in the analysis of ordinary conversation, it is also a fitting way to work with institutional interactions among laypersons and professionals (Sidnell & Stivers, 2013), such as recorded data from semi-structured research interviews. Moreover, while other CA approaches take an agnostic stance on participants' opinions, beliefs, and attitudes (Potter & Hepburn, 2012), my aim in this study is to illustrate how an analysis of the structure of the talk can enrich an understanding of the meaning of the talk. Of course, I recognise that one cannot inhabit the mind of the interviewee and that the theoretical underpinnings of CA dictate that utterances should *first* be understood as activities. However, understanding an utterance *first* as an activity in interaction does not prevent the analyst from engaging with that utterance's "wider meaningful dimensions of human experience" (Woofit & Widdicombe, 2006, p. 49) particularly as they relate to matters of identity and agency. In the following section, I provide further information on the participants, data selection, transcription, and analysis procedures.

3.3.1 Participants

All participants in the study are European-based or European-background research scientists who work in the biological sciences. As noted in the introduction, the dynamic and contested nature of identity is especially relevant for individuals who find themselves in different socio-cultural contexts as they move across geographical and psychological borders (Block, 2007). Therefore, I used personal contacts and snowball sampling to recruit European research biologists who had experienced high degrees of mobility. Each of the participants recruited for the study completed their education to MS level in their respective countries, but have since experienced high levels of research mobility and have worked in both Anglophone (i.e. UK and north-American) and non-Anglophone contexts (i.e. continental Europe). Table 3.1 provides an overview of each of the participants. All names are pseudonyms selected by the participants.

Table 3.1: Participant Overview

Participant	Nationality	PhD Location	Years post PhD	Place of employment	Academic Position
Amelia	Italian	UK	11	USA	Research Professor
Maria	Spanish	Spain	-1	Spain	PhD Student
Emma	German	Spain	7	Germany	Research Fellow
Renzo	Italian	UK	13	USA	Associate Professor
Tom	Italian	USA	11	Italy	Professor
Walter	French	France	6	Switzerland	Research Fellow

The interview guide (see Appendix 3.1) draws on Burgess and Ivanic's (2010) framework for the discoursal construction of writer identity which examines how the autobiographical self of the writer (individual influences) and the writer's possibilities for

self-hood (social influences) impact the discoursal construction of author identity. Participants were asked to choose two of their published research articles to be discussed in the interview and these articles were used as the basis of a talk-around-text approach, which is a widespread method in EAP studies (Lillis, 2009). I conducted the first three interviews over a 10-day period in April 2019 and the final three interviews in November 2021 and early 2022. Each interview lasted approximately one hour, was audio-recorded, and subsequently broadly transcribed. All participants in this IRB approved study (AS-18-98) signed the consent form (Appendix 3.2). Member checking was carried out as participants were asked to review the accuracy of the transcripts and provide additional commentary when necessary.

3.3.2 Data Selection, Transcription and Analysis

For the purposes of this study, I conduct an analysis of two question-answer adjacency pairs from the interviews. The analyses use the CA concept of *action* and *understanding* to examine the researchers attitudes toward writing in their disciplinary communities and their ELF identities. The analytical focus on these two question-answer adjacency pairs is twofold. First, the level of detailed transcription and analysis required in CA limits the amount of data that can be analysed in any one paper. Second, because issues related to identity and self are "handled in the turn-by-turn production of activities in interaction." (Wooffitt & Widdicombe, 2006, p. 49), questions related to the scientists' attitudes and self-perceptions seemed optimal for conversation analytical treatment. Moreover, both adjacency pairs are initiated with a polar (yes/no) question. Despite, the tendency in qualitative research to eschew polar questions in semi-structured interview

formats, yes/no questions are particularly interesting to examine from an *action* and

understanding perspective because, not only do they tightly constrain participant

responses; they also embed presuppositions (Hayano, 2013). As such, the audio-

recordings of two question-answer adjacency pairs were more closely transcribed using

conventions from the Jeffersonian (2004) system and the Handbook of Conversation

Analysis (Hepburn & Bolden, 2013) (see Appendix 3.3). The audio-recordings were

listened to multiple times in order to capture various features of the delivery of the talk

that are, from a CA perspective, basic to how interlocutors "build actions and respond to

the actions of others" (Hepburn & Bolden, 2013, p. 104). Indeed, rather than ignoring

features such as 'fillers', pauses, and false starts, these features were paid close attention

to in the transcription process. Audio files were also converted to WAV files and run

through the speech analysis software Praat (Boersma & Weenink, 2021) to help verify

my manual transcription of pauses, emphasis, and intonation. Although manual or

auditory analysis is more common within CA, the combination of auditory (careful,

repeated listening) and acoustic (machine-generated) analysis is becoming more

widespread (Walker, 2013)

In approaching the data from a CA perspective, Sacks (1984) suggests that

observation is a basis for theorizing. This means that the interview data should not be

approached by pre-specified analytic goals but by "noticings of initially unremarkable

features of the talk or of other conduct"(Sidnell, 2010, p. 130). It was through the process

of listening to and closely transcribing the audio-recordings that instances of orientation

to and renegotiation of some of the premises of my questions became apparent. This was

particularly evident in their renegotiation of some of the presuppositions in my polar

questions. These instances were visible to me because the refutation of presuppositions

and dispreferred responses takes interactional work (Hayano, 2013). It was these

noticings of the interactional work to refute or renegotiate presuppositions, including the

use of subject-predicate discourse markers such as you know and I mean, which I then

began to focus on for further analysis across all six interviews. In short, the analysis is

driven by initial noticings within one interview, which then led to further observations as

I looked across all six interviews.

3.4 Analysis

In this section, I illustrate the benefits of applying CA approaches to EAP/ESP

interview data. I also show how attention to the interactional organization of the talk can

enrich understanding of the linguistic and disciplinary identities of a group of European

research scientists.

3.4.1 Analysis 1: Disciplinary identities

In this first analysis, I use data from the 2019 interviews to illustrate how the CA

concept of *action* and *understanding* can be used to track the epistemic stance of the

interviewer and provide enriched understanding of interviewee attitudes than might be

possible in content-based analyses alone. In each of the interviews, I asked participants if

they admired anyone as a scientific *writer*. Following the field's more mainstream

approach to the reporting of interview data, I might list this question in an interview

guide as '*Is there anyone you admire as a scientific writer in your field of expertise?*' or,

if I was summarizing, I might state that I asked each of the interviewees about *'admiration of science writers'*. However, a CA approach based on action and understanding provides an exact account of the question, not as it was intended, but as it was enacted in each of the 2019 interviews (see excerpts 1 – 3).

(1) **Renzo** *(4/4/19)*

```
1    FJ       e:mm I wonder is there anyone you particularly admi::re
2             ehh in=as a scientific writer in your field of expertise=
3             =this can be someone that you've worked with personally
4             <or> someone that yo:u read.


5    Renzo    (1.5) ((tongue click)) (1.1) I like several e:mm several authors
6             (0.7) mmm ((tongue click)) °to be honest° I'm not (1.2) part-
7             I uh I I don't know=
8             =but I don't know if I I like someone because of his <writing style>
```

(2) **Tom** *(4/9/19)*

```
1    FJ        .h emm (0.2) is there anyone <you admire>(.)
2              not as a scie::ntist but  as a scientific writer
3              in your field of expertise=
4              =you mentioned that you started readin:g thi:s anthropologis:t #or#


5    Tom       yeah no he's not really eh although (1.0) u:h no=
6              =I would say that eh there are a a couple of u:h author that e:hh s-
7              that >you know< they write (0.7) <<popular scientific book>> right?,
8              and e:hh °I don't remember names now°
9              one of them is u:hh ((provides name of author))
10             he's a British guy ((spells author name)) ^I [think^]


11   FJ                                                    [m:m


12   Tom       ((repeats author name))=


13   FJ        =but you like the way he he writes
```

```
14    Tom      I like the way he writes yes yes.
15             u:hh I was trying to think=
16             =there is (1.0) there is someo:ne
17             °that I was actually reading a paper a couple of weeks ago°
18             °°and I said wo:w°°
19             because sometime you you can you can see
20             >>you know<< people that they write
21             and uh oh yes ((provides name of author)) ((spells author last name))
22             he has a beautiful book that just came out (provides name of book))
23             very very well-written
24             =but he's also an amazing u:hh teacher
25             so he has all his u:hh lectures from Stanford online
26             it's it's fun to watch
```

(3) **Walter** *(4/9/19)*

```
1     FJ       e:mm I wonder
               is there anyone you admi:re as a scientific writer
2              (.) in your field of expertise,
3              >so not necessarily< the sci:ence but (.)
4              papers that you've read that you think WOW
5              this guy or(.)girl can really put this together
6              =and it's fine if the answer is no.


7     Walter   ehhh I don't know (.) actually we always have ((throat clear))
8              yeah there are some researcher
9              but it's more like
10             people being director of this big unit somewhere (.)
11             with like 200 publications=
```

As previously discussed, in each of the polar questions there is a presupposition that the interviewee will have some sort of attitudinal stance toward scientific writers. This

assumption is embedded in the question from both a topic and action viewpoint as the question is designed to invoke a yes/no response. In the following sections, I show that (i) the responses offered by the scientists change my epistemic stance to the question over time and (ii) the scientists renegotiate the topical agenda of the question to reveal attitudinal stances to their disciplinary identities.

3.4.1.1. Tracking interviewer stance

In my first interview with Renzo, I add the word '*particularly*' to the core question and I stress the words '*admire*' and '*writer*'. In the second interview, however, I do not use '*particularly*' but rather I front '*not as a scientist*' and stress the word '*scientist*'. This fronting, is potentially driven by my updated understanding of '*admiration of science writers*' as revealed in the first interview in which Renzo renegotiated the topical agenda of the question from *writing* to *research*. Finally, in the third interview with Walter, the core question is enacted in its most basic form, as it might have been written in the interview guide.

In addition to asking the polar question, however, I did other interactional work in the first-pair part (FPP) of the question-answer adjacency pair on each occasion. This interactional work might be omitted in broader approaches to interview data reporting, but it plays an important role in understanding the ongoing interaction *within* and *across* the interviews. In each interview I latch what is called a pre-second insert expansion (Stivers, 2013). These types of expansions are telling because they are oriented to potential trouble with the doing of a preferred second-pair part (SPP) (Sidnell, 2010). In

the first and second interviews, I latch expansions that offer candidate answers. However, in the third interview, my insert expansion is much more complex. As is shown in excerpt 3, I begin by latching '*not necessarily the science but*', which mirrors the fronting I did *within* the polar question in excerpt 2. I then also provide Walter with a candidate answer in '*papers that you've read that you think wow this guy or girl could really put this together*'. It is interesting to note here that my candidate answer contains the idea of '*thinking wow*', which is actually taken directly from a response Tom provided in his interview a few days earlier. Finally, I finish my insert expansion with '*it's fine if the answer is no*'. With this, I am not only signaling that Walter might have trouble giving a preferred SPP, but I am also giving him permission to offer a disprefered response to my polar question. In short, across the three interviews I do not ask the 'same' question, but rather I unconsciously use my updated understandings of the topic from each interaction to do different interactional work and thus impose varying levels of constraints on how the interviewees might do their answers. Using CA in this way, to reveal changes in *interviewer* stance through an analysis of questioning across a series of interviews, could provide a novel analytical approach to EAP/ESP interview data.

3.4.1.2. Renegotiating disciplinary identities

In addition to revealing changes in interviewer stance, excerpts 1 – 3 also illustrate the ways in which each of the scientists (i) attempts to provide a preferred response to my polar question (ii) displays trouble with the question, and (iii) renegotiates the question's topical agenda. As shown in excerpts 1–3, each interviewee attempts to progress the action of the interview by providing some kind of preferred response to my

polar question. However, what I suggest here is that these responses may be more related to the strong conversational preference for agreement and progressivity rather than the interviewees' attitudes to *'admiration of science writers'*. As Sidnell (2010) notes, "structural preferences that organize much of talk-in-interaction are quite independent of … individual, psychological preferences" (p. 76). In other words, the agreeing responses to my positively framed *'is there x'* polar question may simply be a structural feature of the talk rather than a true reflection of the interviewees attitudinal stance towards the presupposition embedded in the question. Indeed, there is further evidence that the interviewee's responses are indicative of structural rather than psychological preference. In the excerpts, we see the interactional work, commonly omitted in broader transcription, around these aligning utterances. This interactional work shows that each speaker initially has trouble progressing the action. Given that there is a strong conversational preference for agreement and progressivity, this non-conforming interactional work or 'trouble' is produced for cause (Hayano, 2013). Renzo produces several tongue clicks and long pauses before and after his agreeing response. Tom produces a yeah/no utterance and uses a conditional expression 'I would say' to respond. Finally, Walter makes a false start and produces a throat clear before his response. Pomerantz (1984) suggests that these hesitations, silences, or prefaces to responses with weak agreement such as the *yeah/no* construction are, in fact, indicative of disagreeing responses. Therefore, this interactional work is important as it provides us with evidence that the scientists are having trouble with the knowledge claim embedded in the polar question. Through this analysis, we begin to see that they are engaging in the

interactional work needed to refute or renegotiate the embedded claim in my question (Hayano, 2013). Indeed, after giving an agreeing response, all three attempt to renegotiate the presupposition with reformulations such as *but I don't know if*, *but he's also*, and *but it's more like*. Renzo begins his renegotiation by questioning whether he likes someone because of '*writing style'*. He then works to focus his response on *admiration of research*. It is possibly this renegotiation, which leads me to front '*not as a scientist'* in my polar question to Tom a few days later. Both Tom and Walter also work to renegotiate my admiration of writer to '*amazing teacher'* or '*director'*. Interestingly, even in Walter's initial aligning utterance, he reformulated *writer* to *researcher*. This shift in terminology works to re-set the topical agenda to *research* not *writing*.

3.4.2 Analysis 2: ELF identities

In this next analysis, I illustrate how *action* and *understanding* can be used to examine a topic of interest in EAP/ESP research. The issue of whether multilingual scholars are inherently disadvantaged compared to native speakers of English in writing for publication continues to be a topic of considerable debate in EAP/ESP scholarship (Hyland, 2016; Flowerdew; 2018). On one hand, scholars suggest that multilingual writers face significant additional linguistic burdens — particularly as novices — when writing for publication (Curry & Lillis, 2019; Flowerdew, 2018). Moreover, in the context of the European Union, debates around the spread of English have led to the suggestion that native speakers of English have a free-ride by enjoying all the benefits of ELF while not contributing to its production (Van Parijs 2011). On the other hand, it is argued that academic English is no-ones first language and that all scholarly writers

regardless of their L1 must acquire this specialized competency (Hyland, 2016). Indeed, Ferguson (2018) argues that the native-speaker advantage orthodoxy fails to recognise that Standard English is not automatically acquired by the English (or the Scots, the Welsh or the Irish) but is, in fact, acquired through lengthy formal education that does have attendant costs (Ferguson, 2018). Moreover, scholarly writing is a complex situated practice within which linguistic (dis)advantage (particularly defined along native/non-native lines) may not be the primary barrier faced by multilingual scholars (Hultgren, 2020; Hyland, 2016; Kuteeva & Mauranen, 2014).

During each of the interviews, I asked the research scientists a polar question that embedded the presupposition of non-native speaker linguistic disadvantage. Appendix 3.4 shows the complete transcript of the six question-answer adjacency pairs in which I embed the FPP native advantage/non-native disadvantage presupposition. Each time I ask the question, the presupposition is also combined with the epistemic stance construction 'by the fact that'. For instance, I ask: Have you felt *disadvantaged in writing for publication* **by the fact that** *English is not your first language*? According to Levinson (1983), presuppositions can be used as resources to convey challenging messages. In this case, the scientists have the choice to either provide a relevant answer thus accepting the advantage/disadvantage proposition or they can bring "the presupposition to the surface of the interaction" (Levinson, 1983, p.396). In the following analysis, I show how each of the scientists use intersubjective discourse markers in their responses to (i) strive for affiliation in renegotiating the presupposition in the question; (ii) boost coherence and aid formulation difficulties; (iii) forewarn of shifts in focus; and (iv) modulate stance.

3.4.2.1 You know as token of affiliation with topical renegotiations

   In this section, I show that irrespective of the participants' initial acceptance or

rejection of the presupposition embedded in the FPP of the adjacency pair, each of them

employs *you know* in the SPP to preface renegotiations or downgrade part or all of the

linguistic disadvantage proposition. Amelia uses '*or you know'* constructions to preface

actions that pivot the topical agenda from the *linguistic* disadvantage proposition to a *not*

good enough proposition. She uses '*or you know'* to preface the notion that her ideas, as

opposed to her language, are the disadvantage (excerpt 4). She uses the construction a

final time in her answer to preface another renegotiation — this time a pivot to the topic

of *convoluted vs. linear* writing styles (excerpt 5). She then finishes her response by

returning to an aligning utterance of "*but yeah all the time"*.

(4) **Amelia**
```
1       (£)even now(£) I always feel like I'm not (0.5) good enough
2   >   |o::r you know my idea are not| good enough
3       |or my °English°is not good enough|.
```

(5) **Amelia**
```
1   >   |o:r you know| some time us also a w::ay to express a sent:ence that is
2       not super linear,
3       mine is very (0.4) convoluted and I need to put things (£) linearly (£)
4       (heh heh heh .h)
```

Tom also uses the discourse marker *you know* to preface a renegotiation of the linguistic

disadvantage proposition. He uses a '*so you know'* construction to preface doubt toward

the proposition, which he has just accepted. He moves from "*WELL YEAH OF COURSE I mean there is a disadvantage if*" to "*you know I never understood if [there is a linguistic disadvantage]*". He then advances *personal motivation* (excerpt 6) and *socio-economic background* (excerpt 7) as alternative factors in an advantage/disadvantage proposition.

(6)   **Tom**

```
1   >   so you know (0.7) I always never understood if it was because
2   >   |you know| these are guys these are people that have been (0.5)
3       they had the experience and the motivation to learn English
4       when they were younger
5   >   |you know in| high sch:ool
```

(7)   **Tom**

```
1   >   maybe you know there's also people that their parents (0.5) were either
2       professors or doctors
3   >   so you know maybe they've been exposed to that.
```

Unlike Amelia and Tom, Walter and Maria do not initially align with the disadvantage proposition posed in the FPP — Walter produces a weak "*yeah*" + negative construction and Maria produces an "*okay so*" followed by laughter. Both participants, do however use *you know* to preface renegotiations. Walter uses *you know* in turn initial position to preface the proposition that learning and using English is part of the job of a scientist (see excerpt 8).

(8)   **Walter**

```
1   >   ˚you know it's˚ it's also a part of our job t:o (0.5) well (1.0)
2       learn English and (0.5) being able to speak English and write in English
```

Maria also uses *you know* to construct her answer but collocates the discourse

marker with *just* which seeks to downgrade the nature of the disadvantage to losing time

(excerpt 9) and also in a turn final position with "*you know that's it*" (excerpt 10).

(9)   Maria

```
1           okay so [heh heh .h ] (0.5) I feel that I'm losing a to:n of ti:me=
2      >    you know (0.1) just looking for words as synonyms
3      >    or you know like they will not ((connection lost))
```

(10)   Maria

```
1           but how to put it in |#a#| (0.3) ((tongue click)) formal wa:y
2           and I'm not repeating myself
3      >    °you know° (.) that's it
```

Rather than being disregarded from the interview transcript, the discourse marker *you*

know can help identify action environments in research interviews where affiliation has

become salient and potentially problematic (Clayman & Raymond, 2021). In these

examples, the interviewees renegotiate the linguistic disadvantage proposition and use

you know as an affilliative token to strive for the interviewer's acceptance of the

renegotiated propositional content of their responses.

3.4.2.2 *You know as a coherence booster / formulation aid*

You know is also employed to accomplish other actions by interview participants.

In addition to its use as token of affiliation, *you know* was used by Tom and Maria to

connect ideas and assist in planning problems (House, 2009). Tom frequently uses the

discourse marker *you know* in this way. In excerpt 11, we can see its use as a coherence

booster when Tom rapidly produces the discourse marker after a hesitation, false start and

micropause sequence to connect the ideas that the books in English were related to the

writer's job as a scientist.

(11) **Tom**

```
1       this guy this my PI from (mid-western city)
2       he's an amazing wri:ter he wrote a
3       book or two books in English (0.5)
4    >  uh s- >>you know<< related to his job.
```

Maria also uses the discourse marker to aid in production difficulties. The marker is

produced the first time with smiley voice after an audible intake of breath within a

syntactic boundary *(excerpt 12)* and again with a markedly lower pitch.

(12) **Maria**

```
1    >     s:o like how to put the verb .h and the (£) you know and the(£)
2          (1.4) subject and everything=
3    >     =|you know| like the structure I know how to do it?
```

Prosodic clues (i.e. quicker rate of production, smiley voice, sharp fall in pitch) as well as

the marker's placement within a syntactic boundary set, what House (2009) calls, self-

referencing connectives, apart from the 'other oriented' affilitative *'or you know'* and *'so*

you know' constructions described in 3.4.2.1. Clayman & Heritage (2021) also note that

while affilliative *you know* is most frequently deployed in turn initial position, self-repair

you know is most commonly deployed in turn medial position and within the repair space,

which is the case in both excerpt 11 and excerpt 12.

3.4.2.3 I mean to shift focus

As previously discussed, Tom produces a strongly aligning utterance to the

linguistic disadvantage proposition embedded in the polar question. He also goes further

by explicitly confirming one part of the presupposition with "*there is a disadvantage*".

However, he prefaces this confirmation with the intersubjective discourse marker *I mean*

and then attempts to qualify who is disadvantaged with "*especially if you are*", which he

then abandons for "*if you don't know English*" (excerpt 13).

(13)   **Tom**

```
1    >   WELL YEAH OF COURSE I mean there is a disadvantage
2        eh eh especially <if you are> (0.5) don't know English,
```

As previously noted, *I mean* can be used to forewarn of upcoming adjustments and shifts

in attention by way of expansions on or explanations of prior talk (Schiffrin, 1987). Tom

uses it here to shift attention to the disadvantage of *not knowing* English, which is closely

related to, but not equal to non-native speaker status. He then expands on this proposition

by giving examples of two non-native speaker colleagues who "did an amazing job

writing". In shifting focus from native speaker status to level of proficiency, he echoes

Kuteeva and Mauranen's (2014) claim that EAP debates have moved beyond a native vs.

non-native dichotomy. Like Tom, Renzo strongly aligns with the disadvantage

proposition with an initial alignment of *"well for sure"*. He also uses the intersubjective

discourse marker *I mean*, but this time to preface the length of time it took him to get to

"the level I am now" (excerpt 14).

(14)   ***Renzo***

```
1       >well< <for su:re>
2    >  I mean (1.0) I (0.5)
3       it took me (0.7) a long time to get to the level (0.5) I am now=
3    >  =and I think I'm still learning.
```

The use of the marker here again shifts attention away from a fixed non-native vs. native dichotomy. However, rather than prefacing explanations of or expansions on prior talk, *I mean* acts as a point of departure for Renzo's subjective evaluation (Baumgarten & House, 2007) of his growth as an ELF user on a novice to expert continuum. Renzo then uses the discourse marker *I think* to introduce the metaphor *steeper learning curve* (excerpt 15), which aligns with the notion that novice multilingual writers experience additional (but not unique) burdens when writing for research purposes (Curry & Lillis, 2019; Flowerdew, 2018). And, as Kärkkäinen (2003) notes, *I think* also serves as a point of departure for the verbalization of the speaker's personal perspective, which may co-occur with topic shifts.

(15)   ***Renzo***

```
1        yeah
2    >   I I ^think^ compared to a native speaker I had a um
3        (0.7) a steeper learning curve to begin with,
```

3.4.2.4 I think and I guess to modulate stance

Finally, Emma was the only interviewee to produce a dispreferred response to the linguistic disadvantage question (excerpt 16). Like Renzo, she constructs her response with the intersubjective discourse marker *I think*. However, unlike Renzo who uses *I think* to introduce a new perspective on the topic, Emma uses *I think* to mitigate her

disprefered response. Kärkkäinen (2003) notes that *I think* is used in this way to do "face-work at points of trouble in interaction" (p. 183). Kärkkäinen (2003) also notes that *I think* employed to accomplish facework receives strong prosodic emphasis, which is consistent with it deployment in excerpt 16 at a distinctly higher pitch than the rest of the answer.

(16) **Emma**

```
1   >   (2.0) u:mm I ^think^ not not per:sonally,
```

Moreover, Emma goes on to discuss how she may have seen disadvantage – not from her subject position as a writer, but from the position of reviewer. This speaks to the notion that multilingual writers themselves might hold biases in gatekeeping positions such as editing and reviewing. However, Emma builds her response with a number of subject-predicate constructions indicating mental states such as *I feel, I imagine, I hope* and *I guess*. In excerpt 17, Emma uses *I guess* to finish her turn. *I guess* in this position functions as an "intersubjective stance frame that organizes the stance taking activity between conversational co-participants" (Kärkkäinen, 2007, p. 184). In using *I guess* Emma hedges her prior proposition that things will "balance out".

(17) **Emma**

```
1   >   well but I think (1.0) e:::h (1.0)
2   >   <well I hope> (0.5) that most like
3       since=since there's so many non native (0.3) e:h researchers
4       non native e:h writers (1.0).h
5   >   that that it balances out (0.7) |°I guess°|
```

To sum up, in the transcribed data from one question and answer adjacency pair across six semi-structured interviews, intersubjective discourse markers were used by each of the interview participants to perform a number of interactional and pragmatic functions. An examination of the intersubjective and affilliative functions of these markers helped to enrich my understanding of how the participants worked to position themselves to the propositional content of the question. Moreover, the context of their deployment underscores that the research interview cannot be characterised as a research instrument in which questions are deployed as stimuli for response, but rather a social action where participants' collaboratively shape the interaction in explicit and subtle ways (Prior, 2018).

3.4.3 European researchers' perspectives on disciplinary and linguistic identity

In the previous section, I illustrated *how* an interactionally focused analysis of EAP interview data could be conducted. I also highlighted the methodological benefits of such an approach. In this section, I more specifically focus on *what* the analysis reveals about the disciplinary and linguistic identities of the European research scientists who participated in this study. The detailed analysis of the structure, rather than content of the talk in Analysis 1 provides us with insight into how Renzo, Tom, and Walter think about their disciplinary identities. In particular, it reveals the way in which each of them demonstrates trouble with the *scientist as author* persona and renegotiates to focus different 'subject positions' (Foucault, 1988). I argue here that these renegotiations are influenced by what they perceive to be socially available possibilities for self-hood (Burgess & Ivanič, 2010) within their discipline. As noted in Chapter 1, possibilities for

self-hood are relevant when examining disciplinary identity construction, as it is within

the constraints of distinct discourse communities that these European research scientists

shape disciplinary identity. Hyland (2001) suggests that practical agent personas are

foregrounded above other subject positions in the hard sciences and it would appear that

Renzo, Tom, and Walter favor more prototypical practical agent positions in their

discipline as they renegotiate *author* to *researcher*, *teacher*, and *director*. Moreover, their

difficulty with the *scientist as author* persona might also be influenced by strong realist

traditions within the hard sciences, which have sought to separate science from rhetoric

and social processes (Myers, 1990). However, rhetorical awareness and the development

of authorial agency are key dynamics in scholarly success across the disciplines (Hyland,

2016). As such, the separation of science and rhetoric is problematic. Reconsidering this

type of separation is relevant in the context of dominant trends in European HE.

European universities have largely mirrored this separation (i) through the use of English

as a means to deliver content (EMI) or (ii) through the use of English in academic writing

(EAP) classes with a focus on language (Airey, 2016). Indeed, Airey (2016) suggests that

pedagogical approaches which promote the integration of language *and* content would

better help European graduate students develop what he terms "disciplinary literacy"

(p.79). A focus on disciplinary literacy is also likely to heighten rhetorical awareness,

which in turn could focus scientific communication as a central disciplinary *practice*.

Turning attention to the construction of linguistic identity, Analysis 2 illustrates

how Amelia and Tom challenge the primacy of language in the linguistic disadvantage

orthodoxy by suggesting that socio-economic background, motivation, or intellectual

ability might be more salient disadvantages. Their stances echo Hyland's (2016) assertion

that physical, scholarly, and financial disadvantages may be greater than linguistic ones

for multilingual scholars. Moreover, although Walter and Maria acknowledge that

language is a factor, they downgrade the disadvantage part of the proposition, by

reformulating *disadvantage* to an *aspect of the job* or as something that *takes more time*.

Tom and Renzo also take up the theme that, for multilingual writers, it is time rather than

native speaker status, which plays the biggest factor in a disadvantage hypothesis. Renzo,

in particular, emphasizes this as he discusses his growth as an ELF user on a novice to

expert continuum. His description aligns with the notion that novice multilingual writers

experience additional (but not unique) burdens when writing for research purposes (Curry

& Lillis, 2019; Flowerdew, 2018). Importantly, all of the topics, which form the

scientists' topical renegotiations, are issues foregrounded by scholars who argue that

debates must move beyond a neat native vs. non-native dichotomy to recognize the more

complex and dynamic aspects of scholarly writing as a situated practice (Hultgren, 2020;

Hyland, 2016; Kuteeva & Mauranen, 2014). Indeed, in the context of the continued

commodification and marketization of the European HE sector, decreasing social

mobility and increasing economic inequalities, rather than linguistic disadvantage, have

been proposed as the most significant marginalizing factors facing graduate students and

early career scientists (Velayutham, 2021). In sum, attention to the trouble the scientists

displayed with the *scientist as author* persona and their topical renegotiations of the

linguistic disadvantage proposition allowed for rich insights into their linguistic and

disciplinary identities.

3.5 Discussion

A central aim of this study was to illustrate how the analysis of interview data as activities in interaction can help the analyst engage with that data's wider dimensions of human experience. The CA lens of *action* and *understanding* illustrates how the relationship between the interviewer and the interviewee is a continually updated sequence of intersubjective understandings (Sidnell, 2010) in which the epistemic stance and actions of *both* participants have an impact on the ensuing interaction (Hayano, 2013). Dominant approaches to research interview data transcription and analysis in EAP research places great emphasis on emic perspectives, but often disregard important insights that can be revealed in the structure of the talk itself. Moreover, the dominance of thematic analyses or direct reportage of respondent turns is based on a very simplistic account of human behaviour (Silverman, 2017; Talmy, 2010). This might be characterised as a type of introspection illusion, which assumes a "direct link between the language of people's accounts and their past and present psychic states" (Silverman, 2017, p. 144).

As was shown in *Analysis 1*, my polar questions and insert expansions are more than stimuli for response. Within each interview, they control the interaction by imposing varying levels of constraints on the interviewees from both an action and topical viewpoint (Clayman & Heritage, 2002; Hayano, 2013). Despite Talmy (2010) highlighting the co-constructed nature of interview a decade ago, these interactional aspects continue to be largely over-looked in EAP research in favour of a dislocated list of 'clean' questions and claims to 'open' responses. Moreover, *Analysis 1* illustrates how

each interaction updates my own understanding of the interviewees' understanding of possibilities for self-hood in their disciplinary field. Over the course of the three 2019 interviews, my polar question becomes less constrained as my epistemic stance to the question changes from *'the interviewee will have a particular candidate'* in excerpt 1 to the *'interviewee will have no candidate'* in excerpt 3. In addition, my pre-second insert expansions become more complex, which signals my understanding of 'trouble' with the question. It is unclear whether a broader thematic analysis would have captured these subtle but important renegotiations of disciplinary identity and the difficulty the participants had with the 'scientist as author' claim posed in the polar question.

Analysis 2 also shows how features of the talk that might be considered as filler phrases like *you know* and *I mean* actually provide additional interactional clues in the analysis of the talk as they reveal what the interviewee's are *doing* with their responses. Each of the scientists interviewed in this study used discourse markers in ways previously identified in both English and ELF conversational contexts to affiliate, plan, shift focus, and mark stance to construct their responses. Notably, however, Clayman & Raymond (2021) did not identify *you know* as an affilliative token deployed with topic renegotiations in their analysis of 200 tokens of the discourse marker. In their study, *you know* is used as an affiliative token with negative assessments, requests that have met prior resistance, and with misdeeds (Clayman & Raymond, 2021). Context of use may well be at play here. Their study analysed *you know* in face-to face and telephone interactions between family and friends, whereas the use of *you know* with topical renegotiations is identified here in research interview contexts. In other words, the

relationship between participants affects deployment of the resource.

In conclusion, the claim that semi-structured interviews inherently allow free and open exploration of interviewees' worldviews denies the researchers own agency, epistemic access, and intervention in the interview context. Moreover, disregarding the important role of intersubjectivity in the research interview as a social action is problematic in EAP research where issues of epistemic access, authority, and responsibility (Bolden, 2018) are enormously important issues. This is particularly true when, as community outsiders, we attempt to gain insight (through questioning) into conventions and aspirations of another discourse community. In the next chapter, I more closely interrogate the discoursal conventions of self-mention and authorial stance in the biological sciences by using a case-study approach to trace the development of Renzo's authorial voice over a period of nine years.

CHAPTER IV

MICRO PERSPECTIVE: FINDING VOICE IN BIOLOGY: A DIACHRONIC
ANALYSIS OF SELF-MENTION AND EVALUATIVE-*THAT* IN THE
DISCUSSIONS OF A MULTILINGUAL SCHOLAR

4.1 Introduction

In Chapter III, I illustrated how the CA concept of *action and understanding*

provided a methodologically robust and insightful approach in an analysis of the identity

positions adopted by European research biologists in turns-at-talk. In this chapter, I turn

my attention to the construction of authorial identity in written academic discourse,

which is frequently investigated through notions of voice in synchronic and textually

focused EAP studies (Matsuda, 2015). A large number of these studies focus on the

incidence of self-mention, such as first-person pronouns (Stock & Eik-Nes, 2016) or

linguistic features such as evaluative *–that* clauses which allow writers to evaluate,

attribute and even depersonalize propositions (e.g. Groom, 2005; Hewings & Hewings,

2002; Hyland & Tse, 2005a; Hyland & Tse, 2005b; Thompson, 2009). While synchronic

studies have added much to our knowledge of voice in academic writing, research that

takes a diachronic approach to capture the embodied expression and development of voice features over time (Dressen-Hammouda, 2014; Fogal, 2019;, 2020; Pérez-Llantada, 2009) are far less prolific. Moreover, EAP studies are oftentimes too textually focused (Swales, 2019) and driven by large-scale analyses, which pay little attention to contexts and variation. This paucity of contextual information seems particularly puzzling given the centrality of context and community to voice construction (Tardy, 2012). The problem here is that textually focused and synchronic studies fail to recognise the historical embeddedness of the writer and the text. As a result, they are unable to capture the dynamic and complex ecology in which voice features might be motivated within the same discipline, rhetorical situation, and writer.

Therefore, the purpose of this study is to more closely investigate the expression and development of authorial voice in the work of a multilingual scholar. This is achieved by adopting a case-study approach to diachronically examine how a European research biologist, who writes in English as an additional language for research publication purposes, uses self-mention to index agency and *that* clauses to index evaluative stance in his research articles (RAs) over a period of nine years. Case study provides a powerful methodological approach for in-depth observation (Casanave, 2002; Yin, 2014) and situated and holistic analysis (Dressen-Hammouda, 2008), thus avoiding the limitations of decontextualized and disembodied analyses of voice construction. Moreover, this study adopts a usage-based linguistics (UBL) approach, which posits that cognitive representations for language are built as language users experience, encode and categorize utterances based on form, meaning and context (Bybee, 2013). Using a UBL

approach means that form-meaning correspondences, known as constructions, are the basic units of language analysis (Goldberg, 1995). In other words, the focal writer's (FW's) language use represents the mental organisation (and subsequent reorganisations) of his language experiences as a scientist across timescales and contexts (Burgess & Ivanic, 2010). In adopting this approach, I not only recognise that writer identity is dynamic and fluid (Matsuda, 2015), but also that "a single act of writing involves the coordination of multiple processes that exist on different timescales" (Burgess & Ivanic, 2010, p, 229).

4.1.1 Research Questions

As argued above, there is a lack of EAP research that captures the embodied expression and development of voice in closely defined contexts across timescales. Based on this identified research gap, the current study addresses the following research questions:

i.   With what frequency does the focal writer use self-mention and evaluative-*that* constructions at different points in time?

ii.   What types of self-mention and evaluative-*that* constructions does the focal writer use at different points in time?

iii.   How do self-mentions and evaluative-*that* clauses construct authorial voice at different points in time?

iv.   What influence does the discourse context have on the construction of authorial voice?

4.2 Finding Voice in biology

Before embarking on the analysis, however, it is first necessary to review (i) self-mention in EAP research; (ii) the function of evaluative-*that* in academic writing; (iii) why voice features are appropriate to examine from a UBL perspective; and (iv) why voice is a particularly interesting concept to examine in the biological sciences. The next section of the paper deals with each of these points in turn.

4.2.1 Self-mention in EAP research

Following Hyland's (2005) model of interaction in academic discourse, this study defines self-mention as the use of first-person pronouns (i.e. *I* and *we*) and possessive adjectives (i.e. *my* and *our*). First-person pronouns can perform a wide array of functions in academic writing (Harwood, 2005) and have been linked to authorial presence (Hyland, 2002) and portrayed as devices that can help writers create a promotional tenor (Harwood, 2005; Walkova, 2019). Possessive adjectives are also used to project discoursal presence (Hyland, 2005) and allow writers in disciplines such as biology to stake proprietary claims to new discoveries (Halloran, 1984). Verbal collocates of *we* are key in constructing the character of the writer as either a practical, cognitive or discursive agent (Hyland & Tse, 2005)

As previously noted, self-mention has been a productive site of investigation (Stock & Eik-Nes, 2016) in voice studies. Arguably, this focus has been driven by the inclusion of self-mentions in Hyland's (2005) model of stance and engagement and earlier research on identity in written discourse (Cherry, 1988; Tang & John, 1999). Studies investigating

self-mention have included research on cross-disciplinary (e.g. Harwood, 2005; Hyland

& Jiang, 2017) and cross-linguistic differences (e.g. Flottum, 2010; Mur-Duenas, 2007;

Walkova, 2019; Shleykina & Junnier, 2021) in frequency, form and function of self-

mention in a number of academic genres. As a result, various taxonomies have been

proposed (Walkova, 2019) which assign greater or lesser promotional power to self-

mentions based on their form, verbal collocations and rhetorical purpose in the text.

4.2.2 Evaluative-*that* in academic writing

Evaluative-*that* clauses enable academic writers to "thematize attitudinal

meanings and offer an explicit statement of evaluation by presenting a complement

clause within a super-ordinate clause" (Hyland & Tse, 2005a, p. 123). The extent to

which the statement of evaluation is made explicit is, however, very much dependent on

the source of that evaluation. Indeed, the choice of subject in the matrix clause allows the

writer to shift, accept, or conceal responsibility for the information provided in the *that*

clause (Hyland & Tse, 2005a). Writers can attribute evaluations to human subjects, such

as themselves or other researchers (i.e. we established that … /they established that …);

they can use abstract entities, such as an experimental phenomena, the data, or the results

in the subject position (i.e. the assay shows that … /the data show that …/the results show

that …); or they might conceal the source of the evaluation altogether by using the

dummy-subject *it* (It is shown that …) (Hyland & Tse, 2005a). Evaluative source is

important in the construction of authorial voice because it acts as a contextual frame for

what follows in the scope of the evaluation (Hyland & Tse 2005b). For example, using an

inanimate source rather than a human source can help writers frame the evaluation as an

objective truth based on an observable phenomena, or, in the case of the dummy-subject *it*, can frame the evaluation as impersonal and without personal bias (Hewings & Hewings, 2002; Hyland & Tse, 2005b). In sum, *that*-clauses are commonly used in academic writing and can help academic writers modulate evaluative stance.

4.2.3 Voice from a UBL perspective

Recently, Swales (2019) has criticised the over-exploration of circumscribed textual studies of voice in the EAP field. However, work by Dressen-Hammouda (2008, 2012, 2014) in geology, is an exemplar of how a more contextualised *and* diachronic analysis can avoid this circumscription to reveal much greater insight into the ways in which features of voice might emerge or develop in different contexts in different writers over time. Of particular interest to this study's usage-based lens is her description of the ways in which linguistic features of voice emerge over time as writers develop specialist knowledge frames (Dressen-Hammouda, 2008). In other words, writers come to develop a repertoire of prototypes of particular frames (Paltridge, 1997) to interpret their disciplinary identity and to formulate their own messages of self-representation through self-mention. In short, the EAP field has been dominated by large-scale frequency analysis of 'big (corpus) data'. As a result, less emphasis has been given to the "detailed study of small phenomena" (Sacks, 1984, p.24) at the discourse level, which acknowledges that language use/acquisition is a make-do solution embedded in and sensitive to specific contexts (Larsen-Freeman, 2006).

4.2.4 Voice in the hard sciences

The hard sciences have been the focus of much of the EAP research on authorial presence (e.g. Dressen-Hammouda, 2014; Hyland, 2001; Williams, 2012). This interest in identity construction in the hard sciences is often framed in respect of their reputation as author evacuated and their proclivity to objectify research in order to "construct texts in which the physical world seems regularly to speak, and sometimes to act, for itself" (Gilbert & Mulkay, 1984, p. 56). However, authorial presence is actually more prevalent in sciences, such as biology, than has previously been characterized. Indeed, Halloran (1984) argues that the rhetorical/ethical style adopted by scientists in the post-war period in the, then, new and emerging field of molecular biology was more entrepreneurial than empiricist and signaled a break from author-evacuated writing styles. Moreover, recent longitudinal corpus-driven research indicates that over the past 50 years, the biological sciences have been the site of significant shifts in self-mention usage – with a 213% increase in the use of first-person pronouns in RAs published in the same leading journals in biology (Hyland & Jiang, 2017). Theories have subsequently been proposed to account for this increase in self-mention, including their use as self-promotional tools in an increasingly competitive disciplinary environment (Hyland & Jiang, 2017). However, caution is needed here. The construction of authorial-self and voice is a complex and strategic choice (Williams, 2012) connected tightly to disciplinary and sub-disciplinary conventions (Samraj, 2005) and within each sub-discipline is strongly influenced by other factors such as rhetorical purpose (Conrad, 2018), study type (Williams, 2012) and disciplinary expertise (Dressen-Hammouda, 2014). Moreover,

Walkova (2019) posits that the degree of promotional 'power' of self-mentions is governed by grammatical form, rhetorical function and co-text. Therefore, generalizations about the increased use of first-person pronouns in the hard sciences as being driven by self-promotion, purely based on frequency counts alone, seem premature. Such claims require a closer contextualised analysis of the actual form, meaning and function of self-mentions at the discourse level. Moreover, given that authorial agency and development of expertise are "key dynamics" in publication success (Hyland, 2016, p. 66) an analysis of the kind undertaken here is a timely exercise and one which is of wider interest within the EAP field, particularly for novice multilingual scientists seeking to find their own voices in their increasingly international and interdisciplinary discourse communities.

4.3 Method

The following sections detail the selection of the focal writer; the focal writer and baseline textual data sampling processes; and the data analysis.

4.3.1 Focal writer

Renzo was chosen as the focal writer for this case-study. The decision to focus on this particular researcher's trajectory is twofold. Firstly, I know him personally, and, having proofread many of his manuscripts over the past 13 years, I have been intrigued by his developing textual identity. Secondly, his career trajectory and language experience is common to many science graduates in Europe who, like him, have become internationally mobile through their careers and publish solely in English-medium journals. In a semi-structured interview forming part of a larger research project on writer

identity, I asked Renzo to define his disciplinary profile. In addition to navigating physical boundaries, his work traversed disciplinary boundaries too. In his own words,

> My undergraduate training is in *molecular biology*. I have a degree in *biotechnology*, and after that I moved onto *biochemistry* and *comparative physiology*, keeping the core of *molecular biology* as part of my research. Basically, my PhD was in *ecotoxicology*, so *I would say that I could define myself as a molecular ecotoxicologist.*

(Renzo, Interview, 4 April 2019)

A native Italian speaker, prior to his arrival in the UK in his late twenties, Renzo's most recent formal English instruction had taken place in high school. His English, academic or otherwise, was very limited. Moreover, he had received no formal writing instruction in his L1 at the university level. In fact, his M.S thesis, which was composed in Italian, was written using models he was able to access from colleagues. As he stated,

> So *instructions was pretty poor*, the only basically thing I have done I have *read other people masters dissertations in the lab* or for example, … the graduate student in the lab gave me her Master's thesis and then I read some others and that was it, and then, that was my instruction basically to basically use that format and *write my own out of that.*

(Renzo, Interview, 4 April 2019)

However, from 2003 onwards Renzo worked, wrote, and, in 2008, published in English. It is also relevant to note here that he did not take any academic writing courses in the UK, but rather, as he informed me, learned to write for publication in English by reading research articles and through the "back and forth" re-drafting/editing of manuscripts between him and his PhD supervisors.

And then **at the PhD level**, when I moved to the United Kingdom, then *that is when I really started learning because I was reading a lot of papers, and my PhD Advisor was reviewing my manuscripts,* which were probably terrible to begin with. And by learning from my mistakes, I probably improved that way. But I never received a formal instruction on how to write for academic purposes. *Maybe 99% of the learning process of writing a scientific paper came from reading papers and from the back and forth between my main advisor and my secondary PhD Advisor.*

(Renzo, Interview, 4 April 2019)

4.3.2 Data

I use the following main data sources (i) four research articles (RAs) published by the Renzo at three-year time intervals over a period of nine years; (ii) baseline data consisting of ten randomly selected discussions published in the same journals in which the focal writer published at each examined time point (40 discussions in total); and (iii) the transcript of an audio-recorded text-based interview with the focal writer.

4.3.2.1 Selection of focal writer textual data.

Once Renzo agreed to participate in the study and had signed the relevant consent forms, I asked him to select research articles, which would be used for the textual analysis and text-based interviews. I asked that Renzo follow two criteria in selecting the research articles for this study. First, he should be the first-author and/or have been primarily responsible for writing the manuscript. Second, given that this study takes a diachronic approach, I asked that he choose research articles from work at different time-points in his career. Although Renzo quickly came back to me with four research articles that met the above criteria, he expressed some concern that the "primarily responsible for

writing" criterion did not apply equally to all part-genres of the four articles. In particular, he pointed out that the abstracts and introductions of the articles might not meet this criterion, but that the methods, results, and discussion/conclusions generally did.

As is the norm in Renzo's discipline, each of the research articles was multi-authored, and, in each project a senior scientist and other collaborators would have advised on, edited and re-composed sections of the paper to greater and lesser degrees. We agreed, therefore, that it might be best to narrow our focus in the more detailed textual analyses and the interviews to sections of the articles that Renzo felt were the 'truest' representations of his written work. At this point, I began to think about focusing the study on a specific part-genre of the RA. With reference to the literature, it became clear that in addition to the methodological considerations discussed with Renzo, there were persuasive grounds for focusing the detailed analyses on discussion writing.

Discussion writing is reported as presenting a considerable challenge for EAL science writers seeking to publish their work in English-medium journals (Martin et al., 2014). Moreover, from a rhetorical perspective, the discussion necessitates some form of authorial-positioning as it requires writers to take an evaluative and interpretive stance on results (Cotos et al., 2016). In short, a decision to focus the detailed analyses of self-mention constructions in RA discussions was driven by (i) Renzo's concerns over authorship, (ii) the reported difficulty this section presents for science writers and, (iii) the rhetorical nature of the part-genre. While I understood that this focus would result in a certain smallness to the focal writer textual data set (7,248 words), it allows for the fine-

grained usage-based analysis of the form, meaning and function of self-mentions and evaluative-*that* clauses in specific communicative events across contexts and timescales.

4.3.2.2 Focal writer textual data.

The RAs used in the textual analysis were published at three-year intervals over a nine-year time period in which the focal writer moves from PhD student to Assistant Professor across linguistic, disciplinary and physical boundaries. The articles are referred to as CBP 2008, BBA 2011, AT 2014 and CBT 2017 throughout the paper[1]. Each RA had between three and five authors, and, as is customary in biology (Clement, 2014), the senior investigator is listed as the 'last author'. It is also important to mention here that on each project, Renzo (as first author) worked with a different senior investigator, who also acted as the chief co-writer/editor. Table 4.1 provides an overview of the publication contexts.

Table 4.1: Focal writer textual data manuscript publication contexts

Publication Year	Journal Name	No. of years post PhD	Academic position	Study type	Study focus
2008	Comparative Biochemistry and Physiology	-1	PhD Student	experimental	cloning and characterization of genes
2011	Biochimica and Biophysica Acta	2	Postdoctoral Researcher	experimental	cloning and characterization of genes
2014	Aquatic Toxicology	5	Research Fellow	experimental	testing of model
2017	Cell Biology and Toxicology	8	Assistant Professor	experimental	characterization and testing of model

[1] Not the official abbreviations of the journals

4.3.2.3 Baseline data.

A small-scale corpus of discussions was compiled from original experimental studies published in the same volumes of CBP, BBA, AT and CBT in which Renzo published his 2008, 2011, 2014 and 2017 research articles. The rationale behind this was to build situated baseline corpora providing highly specific and contextualized data on the frequency and type of self-mention and evaluative-*that* constructions in those journals at those specific points in time. The use of a small, specialized corpus is beneficial on two fronts. First, Flowerdew (2005) advocates an approach in corpus-based genre studies where the analyst is also the compiler of the corpus and more importantly has familiarity with the wider socio-cultural context in which the text was created. Second, from a usage-based perspective, this approach is particularly fitting given a corpus including the types of language that a language user is commonly exposed to, serves as a appropriate benchmark for language experience (Kyle & Crossley, 2017). This is particularly noteworthy given the way in which Renzo learned to write for publication. From a usage-based perspective, in addition to developing his content knowledge through extensive reading, Renzo would have developed exemplars of prototypical self-mention and evaluative-*that* constructions in his discipline through his experience and categorization of these tokens in the input. Great care was taken to ensure that the discussions in each baseline were comparable with the focal writer's work. For example, based on the journals' own categorizations, only original research articles with an experimental focus were included for selection. Furthermore, only those articles, which mirrored the focal writer's rhetorical organisation, were included for selection. This meant that for the

2008, 2011, and 2017 baselines, the articles had a distinct discussion section, which also concluded the article. In the 2014 baseline, on the other hand, only articles, which had separate discussion and conclusion sections were selected to reflect the fact that the focal writer used this arrangement in his 2014 article. Once these articles were identified in the relevant volumes, 10 articles from each were randomly selected in order to form the baseline corpus (50,606 words).

4.3.3 Data Analysis

In order to calculate frequency of self-mentions and evaluative-*that* clauses, the data gathered for the focal writer (excluding figures and tables) were converted into plain text files and the freeware corpus analysis toolkit, Antconc v.3.4.4. (Anthony, 2014), was used to search for use of self-mention and evaluative-*that*. Given that all the RAs are multi-authored, '*we*' and '*our*' are the only tokens of self-mentions found in the focal writer's RAs. Concordance lines were extracted and manually checked to ensure that the identified linguistic features were true representations of their category. Following this initial control, frequency was calculated and then normalized per thousand words (ptw). The same frequency calculation procedure was also followed for the baseline data, which consisted of 40 randomly selected discussions from the four time points.

In addition to the calculation of frequencies and mean use per thousand words as detailed above, each linguistic feature in the baseline data is also shown as a range based on Standard Deviation (SD) (Dressen-Hammouda, 2012, p. 3). The range for each linguistic feature is calculated as one SD above and one SD below the baseline mean and is used to represent the "range of 'normal', or most common occurrences, of that

variable" (Dressen-Hammouda, 2012, p.3). Moreover, excluding the 2014 possessive adjective data, all other baseline data are normally distributed as determined by the D'Agostino & Pearson normality test ($p < 0.05$). Therefore, the baseline data can be used as a measure to track how the focal writer's use of self-mention moves closer to or further from prototypical use by his disciplinary peers.

The data were then analysed from a UBL perspective, which takes form-meaning pairings known as constructions as the basic unit of analysis (Goldberg, 1995). This kind of analysis is particularly useful in uncovering patterns in the expression and development of language use over time (Li et al., 2014). Moreover, I take the constructions into the discourse realm by analyzing them as "tripartite form-meaning-function constellations" (Ostman & Trousdale, 2013, p.486). In other words, the discussion data were analysed at different levels to look at the (i) form (collocations), (ii) meaning (frame semantics) and (iii) function (discourse context) of the self-mention and evaluative-*that* constructions. More specifically:

i.   tokens of self-mention and evaluative-*that* were coded and described at each time point in the focal writer and baseline data with a focus on verbal collocations for self-mention and source of evaluation for evaluative-*that* clauses

ii.   to elaborate on the prototypical or core frames of voice in the data, the most frequent construction types were then identified and described in more detail, specifically in relation to the baseline corpora and Renzo's use compared to his disciplinary peers.

iii.   the core frame of self-mention used by Renzo was analysed in more detail with a

specific focus on the function of the construction at the discourse level, thus taking into consideration the ways in which self-mention constructions and the frames they embody might influence or be influenced by the surrounding discourse and the effect this has on the promotional power of the internal features of constructions at the sentence level.

4.4 Findings and discussion

This diachronic case study reveals that Renzo's inventory of self-mention and evaluative-*that* constructions become more productive over time. There is a continued use of some core constructions, the discontinued use of others, and importantly the emergence of new ones over the nine-year time period. Importantly, frequency counts alone only tell a partial story. As will be illustrated, they do not reveal the emergence of new construction types or how these construction types create a more varied authorial voice over time.

4.4.1 Frequency of self-mention

Twenty-six tokens of self-mention were identified in Renzo's four discussions and 254 tokens were identified in the four baseline corpora. As is shown in Table 4.2, in 2008, 2011 and 2014 Renzo uses self-mention to a lesser extent than the baseline averages and to a greater extent than the baseline average in the 2017 discussion. However, at each time point he remains within the 'normal' range of use. Table 4.2 also shows that in the randomly selected discussions, which form the baseline corpora for this study, there is considerable variation in the frequency with which authors choose deliberately or otherwise to use self-mention in their discussions. This supports Dressen-

Hammouda's (2012) assertion that actual genre use is "characterized by substantial, but equally valid, variation from perceived norms" (p. 2) and that the range based on SD is a useful analytical method to capture this natural variation.

Table 4.2: Focal writer and baseline self-mention frequency (raw & ptw)

Publication Year	Focal Writer		Baseline		
	raw	ptw	raw	av. use (ptw)	range of use
2008	7	2.8	39	3.5	1.8 – 4.8
2011	6	3.4	93	7.9	3.2 – 12.2
2014	2	1.6	61	3.7	0.0 – 7.9
2017	11	6.4	61	5.4	1.8 – 10.7

4.4.2 Frequency of evaluative-*that*

Forty-six token of evaluative-*that* were identified in Renzo's four discussions and 364 tokens were identified in the four baseline corpora. As is shown in Table 4.3, there is an overall decrease in Renzo's usage of evaluative-*that* clauses over time. In 2008, he used the linguistic feature 8.8 ptw and in 2017 this fell to 4.1 ptw. Table 3 also shows Renzo's use of evaluative-*that* over time compared to baseline usage. At each time point, he uses evaluative-*that* to a lesser extent than his peers and, by 2017, his decreasing use of the feature drops him out of the SD Range of use in the 2017 CBT baseline corpus.

Table 4.3: Focal writer and baseline evaluative-*that* raw and ptw frequency

Publication Year	Focal Writer			Baseline	
	raw	ptw	raw	av. use (ptw)	range of use
2008	22	8.8	102	9.4	5.3 – 13.4
2011	11	6.2	118	10.1	5.0 – 15.1
2014	6	4.8	104	6.4	1.6 – 11.2
2017	7	4.1	139	13.3	7.3 – 19.3

4.4.3 Type of self-mention constructions across timescales

The next section of the paper explores the concept of variation, not in frequency, but in the types of self-mention constructions used by the focal writer in comparison to the baseline corpora across timescales.

4.4.3.1 Self-mention in 2008 – the PhD student

The use of *we* constructions in Renzo's 2008 discussion falls into two categories. He uses the pronoun most often (80%) in conjunction with verbs indexing what Hyland & Tse (2005) call 'research acts' or 'actions in the real world' + NP (example 1) and secondly with a reporting verb + *that* (example 2).

(1) 2008 FW   Consequently, **we** **characterised** *the cDNA* for the sea bream
               homologue of Ctr1, saCtr1…

(2) 2008 FW   In this study, **_we_ _found_ that** _adult sea bream_ Ctr1 mRNA was also widely expressed, but most highly in intestine with level …

These constructions are also seen in the 2008 baseline data. In addition to these uses, first-person pronouns are used by his disciplinary peers in a construction with expressions of ability + research act + NP (example 3) and cognitive verbs + _that_ (example 4).

(3) 2008 Baseline   Using the continuous Percoll density gradient (density gradient of 1.03 to 1.10) in the present study, **_we_ _were able to isolate_** _the specific cell types_ and estimate their numbers.

(4) 2008 Baseline   **we** **_speculate_** _that_ the basic mechanism of inhibition is similar in the two enzyme sources …

In 2008, Renzo uses possessive adjectives to compare his and his co-authors results and findings with other studies in compare and contrast constructions (examples 5 & 6). This compare and contrast construction, which juxtaposes 'our' _vs_. 'other' is, however, only seen in three of the 22 tokens of _our_ constructions in the 2008 baseline. More common usages in the 2008 baseline include _our_ in a locative phrase, such as _in our study_ (6 tokens); _our_ + research entity as an abstract rhetor, such as _our data strongly suggest_ (5 tokens); and _our_ collocating with self-promotional expressions, such as _our study represents the first_ …(4 tokens).

(5) 2008 FW   Indeed **_previous studies_** … provides* evidence for … and **_our_ _finding_** that… **_supports this contention_**.

(6) 2008 FW   **_Our results suggest_** that … **_In contrast_**, ...

4.4.3.2 Self-mention in 2011 – the post-doc researcher

In the 2011 discussion, Renzo uses similar *we* constructions to 2008 — four out of 6 tokens are used in *we* + research act verb + NP constructions — but he extends his use to include a *we* + expression of ability construction (example 7).

(7) 2011 FW

```
The elongase Elovl2 was not consistently changed by
any treatment and thus could not be included in any
of the groups of genes to which we were able to infer
a possible regulatory mechanism.
```

The 2011 baseline differs from the 2008 baseline not only in terms of frequency of use but also in the type of constructions used. The verbal collocation in the 2011 baseline which is most frequently used is *we* + reporting verb in an evaluative-*that* clause rather than *we* + research act verb + NP. A number of different verbs such as *found*, *show*, *demonstrate* and *confirm* are used in this reporting construction. The data were also checked for *that* deletion and only one token was found. There were also two tokens used in chunked phrases with reporting verbs i.e. *as we reported previously / as we have reported previously*.

Although Renzo extends *we*-type constructions in 2011, he does not use any possessive adjectives in his discussion. This is in contrast to the 2011 baseline, which shows the greatest use of *our* constructions (2.7 ptw) across all of the baselines. Moreover, in the 2011 baseline, *our* in a compare and contrast construction (previously used by Renzo in 2008) is commonly used (11 out of 33 tokens) along with the other uses of *our* with abstract rhetors (16 tokens) and in a locative phrase (4 tokens).

4.4.3.3 Self-mention 2014 – the Research Fellow

As previously shown, the 2014 discussion has the lowest frequency (0.8ptw) of first-person pronoun use across the nine-year period. Moreover, it also differs in the type of construction used (example 8).

> (8) 2014 FW
> ```
> The criteria of using an insert with a TEER in >5kQ
> was based on our previous work (Walker et al., 2007,
> 2008) and also had a practical component because for
> each study we have a number of different batches of
> cells that are all growing at slightly different
> rates.
> ```

Clearly, there is a change in both frequency and type of first-person pronoun use in this 2014 discussion compared with Renzo's previous manuscripts. Moreover, the type of *we* construction used is not seen in the 2014 baseline or indeed any of the other baselines. In fact, the 2014 baseline mirrors the other baselines in that we see authors use first-person pronouns most frequently to describe their actions in a *we* + research act verb + *NP* construction or report their claims in a *we* + reporting verb evaluative-*that* construction.

In 2014, we also see Renzo use an *our-* type construction uncommon across the baselines (example 9). In fact, of the 110 tokens of *our* across the four baselines, only one is used in this *based on our*-type construction.

> (9) 2014 FW
> ```
> The criteria of using … was based on our previous
> work and also had a practical component because ...
> ```

In short, Renzo's 2014 usage deviates from baseline use as well as his own prior use of *we* and *our* constructions. When I asked Renzo to account for this variation in 'style', his

immediate response was to suggest that the manuscript had probably been "re-worked" by the senior scientist on the project (personal communication, 20 April 2019). Indeed, it is interesting to note that the 2014 discussion uses passive constructions more often than the other texts and that the single idiosyncratic uses of *we* and *our* in the 2014 discussion section were, in fact, the only tokens of self-mention in the entire manuscript. It is not totally implausible that, as Renzo suggested, the final edit re-workings of the senior collaborator led to this different stylistic fingerprint.

4.4.3.4 Self-mention 2017 – Assistant Professor

In 2017, Renzo uses *we* constructions in similar ways to his 2008 and 2011 discussions and to his disciplinary peers by placing the first-person pronoun in what appears to be his 'core frame' of *we* + research act + NP. Importantly, however, in 2017, he also uses the pronoun with verbs that do more than index action or report on findings (examples 10 & 11). In fact, 25% of occurrences of *we* in the 2017 discussion were with verbs performing a discourse function (Hyland & Tse, 2005) in an evaluative-*that* clause. This type of use was not seen in his previous discussions, but is seen in the 2017 baseline corpus and, in fact, across the baselines.

(10) 2017 FW   **We therefore propose** that these cells are better able to eliminate silver

(11) 2017 FW   In conclusion, **we here present** development and initial characterization of the first in vitro intestinal barrier model for fish.

In addition, Renzo uses *our* constructions most often (1.7 ptw) in 2017 and again

uses them in diverse ways. For example, he uses them to compare and contrast findings
(example 12), collocate with promotional expressions (example 13) and as an abstract
rhetor presenting a finding (example 14).

(12) 2017 FW   **Our** results are **in support of** this proposal;
moreover, in contrast ...

(13) 2017 FW   **Our** proof-of-concept experiment **is encouraging** for
the use of ...

(14) 2017 FW   **Our** results **demonstrate** that …

In summary, Renzo's use of self-mention construction types becomes more
productive and varied over time. In 2008 and 2011, he predominantly uses first-person
pronouns to emphasize his practical agency through *we* + research act verb constructions.
However, in 2017 we see a greater and more varied use of first-person pronouns and
possessive adjectives performing research, discoursal, comparison, and promotional
functions.

4.4.4 Type of evaluative-*that* constructions across timescales

The next section of the study explores the concept of variation, not in frequency,
but in the types of evaluative-*that* constructions used by Renzo in comparison to the
baseline corpora across timescales. More specifically, I discuss the types of evaluative
sources he uses at each time point, how his use aligns with or deviates from the baseline,
and how his choice of evaluative sources contextually frame evaluations thus
constructing different authorial voices over time.

4.4.4.1 Evaluative-*that in 2008 – the PhD student*

In 2008, Renzo most commonly uses the dummy-subject *it* (example 15) and measurable experimental phenomena (example 16) in the subject position of evaluative-*that* clauses. These two source types account for 17 of the 22 evaluative-*that* clauses. Other evaluative sources include generalized findings from his own study (four tokens) and one human source in the form of the FPP *we*.

(15)   For example, in freshwater (low Na+) **it is well established that** copper uptake can occur through the ENaC sodium transporters Na+ of gill, although in conditions of high concentrations, such as the seawater used here, these transporters are thought unlikely to play a significant part in either gills or gut (citation).

(16)   Thus, the **relatively high level of MT and GR mRNA observed under normal commercial dietary conditions indicates that** Zn levels may be excessive in this diet.

In this early-career discussion, Renzo relies more heavily on the dummy-subject *it* to act as the source of his evaluations than his peers. The construction accounts for 41.2% of all evaluative sources used by Renzo in 2008, but only 27.4% of sources in the baseline corpus. As previously noted, the dummy-subject *it* performs important hedging and depersonalising functions (Thompson, 2009) and previous research has noted its more frequent use by novice academic writers (Hewings & Hewings, 2002). Following an analysis of this feature's function in the discourse, it is suggested here that Renzo uses this feature as a novice academic writer to conceal the subjective nature of his evaluations (Groom, 2005).

4.4.4.2 Evaluative-*that* in 2011 & 2014 – the post-doc researcher/research fellow

Unlike the quite marked stylistic change in Renzo's use of FPPs in 2014, his use
of evaluative-*that* appears to be more consistent over time. In 2011 and 2014, Renzo
again uses the dummy-subject *it* as an evaluative source, but does so to a lesser extent
than 2008 and it now accounts for 29.4% of evaluative sources across the discussions.
The most common sources for evaluative-*that* clauses are now his own generalized
findings (example 17) and once again measurable experimental phenomena (example
18), which account for 12 of the 17 types of sources used.

(17) 2014   **The results from the current study show that** a primary fish
             gill cell culture system, can withstand unfiltered and
             filtered

(18) 2011   Furthermore, **tissue expression profiling showed that** both
             salmon SREBP genes were expressed at the same level in each
             of the tissues examined.

Renzo continues to commonly use measurable experimental phenomena as the
source of his evaluations across timescales. In 2011 and 2014 they make up 35.2% of the
evaluative sources compared to 26.4% of evaluative sources in the 2011 and 2014
baselines. As a post-doctoral researcher, these sources are likely to be phenomena he
himself has witnessed and has become expert in interpreting. He uses these phenomena
as "abstract rhetors" (Halloran, 1984, p. 75) to speak for him, thus suppressing his own
agency and authorial presence in the discourse. Renzo's continued tendency to use these
observable phenomena in the subject position to contextually frame what follows in the

complement clause is noteworthy in that it aligns with the practical agent persona he constructs using self-mention.

4.4.4.3 Evaluative-*that* 2017 – Assistant Professor

Although Renzo doesn't use evaluative-*that* clauses as frequently in 2017, the types of evaluative sources used more closely align with baseline use. In 2017, Renzo draws on a wider range of sources and uses abstract rhetors to a lesser extent than in previous years. Indeed, in 2017 he attributes evaluations to other scholars (example 19) and places himself in the subject position to promote his own agency in the matrix clause of the evaluation (example 20). In doing so he allows himself, and other scholars in his disciplinary field, to take responsibility for evaluative stances. These human subject sources are also seen in the baseline, accounting for 20.1% of sources across all baseline data. Notably, Renzo also uses the depersonalising function of the dummy-subject *it* to a much lesser extent and it now accounts for only 14.3% of evaluative sources in 2017.

(19) 2017   **Kawano et al.** (2011) ***proposed that*** this intestinal cell line may have…

(20) 2017   **We** therefore ***propose that*** these cells are better able to eliminate silver

In summary, although Renzo uses evaluative-*that* clauses less frequently over time, the source types he uses within those clauses become more varied. From a UBL perspective, we can say that there is an overall increase in the degree of creativity of evaluative-*that* construction types over time as the type-token ratio increases from 0.18 in 2008 to 0.71 in 2017 (see Table 4.4)

Table 4.4: Type-token frequencies for focal writer's evaluative source use

Year	Type	Token	Ratio
2008	4	22	0.18
2011	3	11	0.27
2014	3	6	0.50
2017	5	7	0.71

In 2008 he frequently used the dummy-subject *it* to conceal his evaluations. In 2011 and 2014, he continues to use his core evaluative source of measurable experimental phenomena in line with the practical agent persona we see developing in his use of self-mention constructions. Finally, in 2017 he uses evaluative-*that* clauses least often in terms of frequency, but shows his most varied use of source types – commonly using human subjects as the source of his evaluations. Moreover, although his frequency of use deviates from the baseline by taking him below the SD range in 2017, his varied use of source types more closely aligns with the different ways in which his disciplinary peers attribute their evaluations.

4.4.5 Constructing authorial voice

As has been shown, Renzo's use of self-mention and evaluative-*that* constructions becomes more creative over time. The analysis also shows that he embodies the authorial identity of a practical agent across time to a greater degree than the baseline averages and that he reduces his reliance on concealed evaluative sources in favor of animate and inanimate subjects, which allow him to act as a textual moderator. In this next section, I look more closely at these two core frames of authorial voice – the practical agent and the textual moderator.

4.4.5.1 Author as practical agent

The most common usage of *we* by the focal writer and a common usage in each of the baselines is *we + research act verb + NP*. In essence, Renzo uses the first-person pronoun to construct the identity of 'doer'. Hyland (2001) calls this the 'practical agent' who uses self-mention to "reassure us of … professional credentials through a demonstrable familiarity with disciplinary research practices." (p. 220). Interestingly, Renzo tends to construct a practical agency in his discussions to a greater degree than is seen in the baseline (see Table 4.5). Although this tendency decreases over time, this use of self-mention allows him to "adopt a particular stance and disciplinary-situated authorial identity" (Hyland, 2005, p. 181) as the practical agent. Drawing on Dressen-Hammouda's notion (2008) of "specialized embodied frames" (p.236), it appears that, at an early stage, Renzo embodies "the idea that knowledge is built on experiment, induction, replication, and falsifiability" (Hyland, 2011, p.194).

Table 4.5: Construction of practical agent frame by focal writer and baselines over time (%)

		Focal Writer	Baseline
we + research act verb	2008	80	41
	2011	67	27
	2014	-	41
	2017	50	40

As Renzo also commented, in the early stages of his career his enjoyment was *doing* rather than *writing* the science.

> Well, at the beginning was very difficult because I had never done anything like that before, and yeah I mean I enjoyed the ***research***

100

side of things, and I feel, felt passionate back then and like I am still am now to a certain extent. *The enjoyment was to do with the science rather than the writing the science back then*.

(Renzo, Interview, 4 April, 2019.)

4.4.5.2 Author as textual moderator

The most marked change over time in Renzo's use of evaluative-*that clauses* is his dropping of concealed evaluative sources in favor of non-animate sources of evaluation in 2011 and 2014 and animate and non-animate sources of evaluation in 2017 (see Table 4.6).

Table 4.6: Construction of moderator frame by focal writer and baselines over time (%)

Year	Non-animate		Animate		Concealed	
	FW	B	FW	B	FW	B
2008	54.5	57.8	4.5	14.8	40.9	27.4
2011	72.7	56.0	-	29.6	27.2	14.4
2014	66.6	61.5	-	15.4	33.3	23.1
2017	42.9	62.8	42.9	19.2	14.3	18.0

Despite this increased use of animate sources in 2017, he continues to place specific measurable phenomenon as abstract rhetors in the subject position of *that*-clauses thus suppressing his own agency and authorial presence more generally in these types of

evaluations (Halloran, 1984). These textual findings are consistent with his characterization of himself as a scientific writer who let's data speak, As he states;

> Yeah, I'm thinking, probably something that changed from my 2008 writing style to my more current writing style, *I've tried to let the data speak for me*. I don't want to express my opinion.

<p style="text-align:center">(Renzo, Interview, 4 April 2019)</p>

4.4.6 Discourse function core frame – the practical agent

Renzo's core frame of self-representation is clearly the practical agent (*we +* research act verb +NP), which he uses in 12 out of the 20 tokens of *we* across the four discussions. This final section of the analysis looks at the different discourse functions this practical agent is positioned in. These are characterised as (i) the reasoning practical agent, (ii) the limited practical agent, (iii) the problem-solving practical agent and (iv) the contributing practical agent. The arrows in Figures 4.1–4.4 represent the direction or flow of influence either to or from the agent in each discourse function. So, for example, if Renzo's descriptions of his actions are being influenced by the surrounding discourse, there will be an arrow *towards* the practical agent but if Renzo's actions influence the surrounding discourse, then the arrow will move outwards *from* the practical agent.

4.4.6.1 Reasoning practical agent (5 tokens)

The most frequent way that Renzo positions himself as the practical agent (PA) is in a stretch of discourse that explains the purpose or aim of the research and his actions in achieving those purposes and aims (Figure 4.1).

| study purpose or aim | → | reasoning PA |

Figure 4.1: Reasoning Practical Agent

This discourse function is most frequently seen in the 2008 and 2011 discussions (see examples 21 & 22). This function is introduced by phrases such as *in order to* or by conjunctive adverbs such as *therefore* and *consequently*.

(21) FW 2008

In attempting to understand the mechanisms of Cu uptake in fish, whether under conditions of deficiency, sufficiency or excess, **it is essential to consider known Cu-specific transport proteins.** Consequently *we characterised* **the cDNA** for the sea bream homologue ...

(22) FW 2011

Previous studies have indicated the role of SREBPs in regulating genes such as those of cholesterol and LC-PUFA biosynthesis involved in the adaptation to vegetable oil-based diets in salmon liver [10,13–17]. **In order to study this process in more detail <u>we searched for a cellular model</u>** for salmon lipid homeostasis.

4.4.6.2 Limited practical agent (4 tokens)

In 2008, 2011 and 2017 Renzo uses self-mention to describe the limitations of his practical agency (Figure 4.2).

| current knowledge | → | limited PA |

Figure 4.2: Limited Practical Agent

In each of these uses there is a negative construction such as "we did not" or "we could not" followed by a research act verb within a stretch of text that mitigates those limitations by reinforcing the authors' knowledge of disciplinary findings (examples 23 & 24). What is notable in the 2017 example is that he is able to use a self-citation, thus mitigating practical limitations with existing knowledge from his own work.

(23) FW 2008   In freshwater rainbow trout, renal excretion of excess Cu in fish is negligible compared to hepatobiliary routes (citation) and **although _we did not measure_ copper levels in the kidney, marine fish, in contrast to freshwater species, can accumulate Cu** in the kidney following brachial exposure, although at considerably lower levels than liver (citation).

(24) FW 2017   **Although _we could not localize_** ATP7A in RTgutGC due to the lack of a fish specific antibody, **the ATP7A gene is present in fish as well as other vertebrates (self-citation).**

4.4.6.3 Practical agent as problem-solver   (2 tokens)

In 2017, the practical agent is also positioned as a problem solver (Figure 4.3).

| problem-solving PA | → | disciplinary limitation |

Figure 4.3: Problem-Solving Practical Agent

In this function, Renzo uses his practical agency to mitigate existing disciplinary or methodological limitations and issues with phrases such as _to account for this_ and _even though_ (examples 25 & 26). Unlike the functions described in 4.4.6.1 and 4.4.6.2, this particular positioning of the practical agent is only seen in the 2017 discussion and is the

first time we see the influence arrow move outwards from the practical agent to the discourse.

(25) FW 2017   Thus, less silver per cell might per se be available. However, **to account for this possibility, _we_ _normalized_** the levels of silver measured in the cells to total protein in the different culture set-ups.

(26) FW 2017   **_We measured the NKA_ enzyme activity for the first time in a rainbow trout cell line even though it was not yet possible to do this in RTgutGC cells** taken from insert cultures because of the high amount of material needed.

4.4.6.4 Practical agent as disciplinary contributor   (1 token)

A final way in which Renzo positions his practical agency to influence the surrounding discourse in 2017, is to connect his actions as the practical agent to support existing disciplinary knowledge (Figure 4.4).

| contributing PA | ➡ | disciplnary knolwedge |

Figure 4.4: Contributing Practical Agent

As is shown example 27, he uses this construction to first communicate to the reader the current thinking in the field, and then in the main clause uses the authorial exclusive *we* to show how he, and his co-authors, contributed to existing knowledge.

(27) 2017   ***In further support of*** the maturation of a polarized epithelium in vitro, *__we__* *__observed the development of__* a profound actin network in the cells

In summary, although Renzo's core frame of self-mention across the nine-year period is the practical agent, he positions his agency within the discourse in diverse ways. Indeed, situating the sentence-level self-mention constructions in the discourse realm reveals the more nuanced aspects of his developing authorial voice. More significantly, it is argued here that his later uses of self-mention as the problem-solving and disciplinary contributing practical agent, are, in fact, more self-promotional than previous uses of the practical agent construction. It is his actions as supporting or accounting for existing disciplinary knowledge and experimental problems, which reveal his identity not just as 'doer' but as a practical agent influencing disciplinary knowledge construction. This insight into the positioning of practical agency in biology discussions provides a new perspective and level of detail to previous analyses and taxonomies of the function of first-person pronouns in academic writing (Walkova, 2019), which place greater emphasis on the functions and rhetorical power of self-mentions collocating with reporting, discourse and cognitive verbs. As we have seen, while important, the focal writer in this study and the biologists in the baseline corpora spend more time talking about what they are *doing* rather than thinking or proposing and that phraseological patterns in the surrounding discourse help frame the nature and promotional tenor of this practical agency.

CHAPTER V

CONCLUSION

Drawing on theories from EAP, Academic Literacies, and ELF research, the goal of this doctoral dissertation was to bring together European perspectives on *language*, *identity*, and *voice* in an examination of EILS. At the macro level, the comparative critical discourse analysis of English as a lingua franca in European higher education and research contexts revealed intra-institutional disparities in European Union language policy discourse and provided nuance to claims that language policy is absent from EHEA discourse. At the meso level, I showed that the lens of Conversation Analysis (CA) (Sacks et al., 1974) can be used to enrich understandings of how European research scientists construct linguistic and disciplinary identities in semi-structured research interviews. Finally, at the micro level, the Usage-Based Linguistics (UBL) study revealed new insights into the textual development of multilingual writer's authorial voice across timescales and contexts.

In this final chapter, I will first outline the significance and implications of each of the studies. I then conclude with some broader comments on limitations, future directions, and the contribution of the dissertation to our understanding of EILS as a complex dynamic system.

5.1 English as a lingua franca in the European Union

The aim of Chapter II was to use the discourse of two EU policy actors as entry points to uncover the ideological positions and potential *intra*-institutional differences in European ELF policy at the supranational level. More specifically, I conducted a comparative analysis of ELF discourse in the published declarations and communications of the European Higher Education Area (EHEA) and the European Research Area (ERA). The absence of explicit ELF discourse over the institutional lifetimes of these education and research actors signifies the politically sensitive nature of language policy and the continuance of a top-down/bottom-up conceptual gap at the heart of EU language policy-making identified over two decades ago (Seidlhofer, 2001).

A very large body of scholarly work has documented the rapid expansion of EILS in Europe and the dynamic ways English is used in research and education contexts as a lingua franca. However, EHEA and ERA policy ignores the increasing dominance of English as a communicative medium of choice and, more significantly, seems apparently unaware that ELF can be appropriated and adapted to meet the specific needs and contexts of its users in education and research contexts. Moreover, the two actors conceptualise language in very different ways. The ERA conceptualizes language in neo-liberal terms as a culture-free tool that might be used to advance the KBE agenda, while EHEA rhetoric takes a socio-cultural perspective that ties language to a community's values, culture and traditions. Ultimately, these opposing positions fail to recognise that debates around EILS in Europe have transcended the binaries of cultural vs. economic, native vs. foreign, global vs. local, and centre vs. periphery (Pérez-Llantada, 2015). In

short, both ERA and EHEA policy is disconnected from the nuanced reality of language use in education and research contexts. Both actors adopt reductionist stances to language and language users, which disregard the interplay of elements and agents that constitute ELF in Europe. Moreover, the incorporation of the ELF paradigm into ERA policy does not look likely in the immediate future. The most recent policy agenda of the ERA "European Research Area Policy Agenda: Overview Of Actions For The Period 2022-2024", which due to its very recent publication did not form part of the ERA corpus for this study, does not contain any discussion of language-related matters. Again, given the increasing centrality of EILS in European research contexts, this absence is striking.

5.2 Action and understanding in the semi-structured research interview

The goal of Chapter III was to illustrate how the CA concept of *action and understanding* provides methodologically robust and insightful approaches to data from a series of semi-structured interviews with European research scientists who write in English for research publication purposes. The analysis shows that EAP researchers can achieve richer insights into disciplinary conventions and discoursal aspirations by focusing on "the stubborn details of real events" (Sidnell, 2010, p.35) through a more detailed transcription and analysis of the interview data. For example, I was able to illustrate how the three scientists' attitudes to writing were inextricably bound to other disciplinary identities by using evidence from the structure of their talk. Moreover, the study illustrated how intersubjective discourse markers were used to renegotiate presuppositions and adopt attitudinal stances to the linguistic disadvantage orthodoxy. In addition, although researchers are told to avoid *yes/no* question formats in semi-

structured interviews (Merriam, 1998), this analysis shows how carefully crafted polar questions, which require the interviewee to engage in interactional work to accept or reject presuppositions, might actually serve as vehicles to test existing assumptions about disciplinary conventions and discoursal aspirations.

It is important to note, however, that a CA approach will not be suited to all interviewing contexts or research questions and the level of analysis required clearly limits the amount of interview data that can be analysed in any one study. However, if EAP researchers are to lay claim to emic perspectives through interview, then they must not erase their own agency as researchers *or* the messy details of the talk from the analysis. To this end, I offer the following practical suggestions to researchers, who might consider using, or partially incorporating, CA when working with interview data:

- ***Insider lens***: CA is particularly appropriate in some areas of inquiry. Interview questions concerning the cognitive states of interviewees i.e. identity, emotion, and attitudinal stance etc. may be better served by a CA approach rather than content-based analysis.
- ***Self-reflexive tool:*** CA can be used as a self-reflexive tool to track the epistemic stance of the *interviewer* over the course of a research project. This could add another dimension of analysis to an existing project, or the data could be used to generate an additional publication.
- ***Hypothesis tester:*** CA can be partially incorporated by researchers using other analytical methods for their interview data. Polar questions can be intentionally placed in targeted sections of an interview and then analysed using a CA approach

110

to test existing disciplinary hypotheses

5.3 Finding voice in biology

The UBL analysis in Chapter IV shows that the focal writer's developing authorial voice is inseparable from the timescales and contexts from which it emerges. This observation is consistent with Lei and Hu's (2019) study, which found that scholarly publishing is a culturally mediated and socially distributed activity in which novice researchers use cultural artifacts and social others to overcome linguistic difficulties. Renzo employed numerous resources and multiple processes to successfully write for academic purposes. The texts analysed for Chapter IV are years of doing and thinking work; the discussions are the products of months of daily edits, revisions and conversations about science and the writing of the science. As was shown, Renzo develops a creative inventory of form-meaning pairings of self-mention and evaluative-*that* constructions over time. Secondly, it was shown that the core frame of self-mention developed over time by Renzo is that of practical agent. Moreover, the examination of how his core frame of practical agency functions at the discourse level shed light on how he positioned himself in the text in more and less self-promotional ways across time even though he continued, at the sentence level at least, to use the same *we* + research act verb + NP construction. The diachronic study also highlights the ways in which circumscribed textual analyses can miss the vital clues as to why scientific writers might choose (deliberately or otherwise) a range of features to index authorial voice. It is only through the investigation of the "tripartite form-meaning-function constellations" (Ostman & Trousdale, 2013, p.486) and the contexts "for and out of which" (Tardy, 2012, p.39) they

are produced that we can begin to truly understand and assign any meaning potential to the increased, or for that matter decreased, use of linguistic features indexing voice by academic writers in specific contexts.

5.4 Limitations and future directions

In Chapter 2, I discussed the top-down/bottom-up disconnect at the heart of EU language policy-making (Seidlhofer, 2001). With its focus on the top-down ideological positions of EU actors, the study in Chapter 2 did not attempt to bridge that gap. It remains far from clear how the positions of the EHEA and the ERA actually impact doctoral student experience. As previously noted, however, the research scientists who acted as informants for this dissertation reported that, far from promoting linguistic diversity, participation in the EHEA's Erasmus Programme (European Commission, 2022) solidified the role of English as *the* lingua franca of European academic life. Future directions for research in this area could examine the extent of the top-down/bottom-up disconnect in EU language policy-making. This could be achieved by analyzing policy documents published by EU education and research initiatives such as the Erasmus Programme (European Commission, 2022) and by conducting in-depth interviews with Erasmus participants on their actual language experiences within the programme. Interview-based studies could also be used to further examine the range of factors (i.e. economic, scholarly, physical, and linguistic) that affect the ability of European scholars to publish internationally. Indeed, only a very small portion of the interview data I collected for this dissertation was anlaysed in Chapter 3. Future plans for the interview data include a study on the linguistic disadvantage orthodoxy. In that study, I also aim to more fully develop a methodological

approach, which harnesses the analytical power of both thematic analysis and CA. Within

EAP interview research, only Han & Hyland (2019) have attempted to combine these

approaches in a study on emotional reactions to written corrective feedback. While I

recognize that CA and thematic analysis have very different theoretical backgrounds, I

believe that the combination of methodologies could potentially provide a powerful

methodological synergy for EAP research akin to Baker et al's. (2008) work in combining

CDA and CL in discourse studies. Finally, it should be noted that most studies

investigating authorial voice in EAP contexts use published RAs as textual data. Moreover,

those that adopt diachronic analyses tend to use large units of time (i.e. decades/years). My

micro-level study followed these norms, and I provide justifications for these research

design choices in Chapter 4. However, an original goal of this research project was to

include a study on the development of voice over smaller timescales (weeks/months) in

successive drafts of an RA manuscript. While Fogal (2017) was able to adopt a similar

approach in his investigation of authorial voice in a TOEFL preparation class, the more

sensitive nature of manuscript preparation and co-author dynamics in the biological

sciences meant that this level of data was unattainable to me. A considerable amount of

relationship building between the researcher and the research participants would be needed

to access such data rich sources.

5.5 Closings

European scholarship examining the role of English as the global language of

science has been instrumental in dismantling traditional dichotomies (i.e. global vs. local,

native vs. non native, and centre vs. periphery) associated with EAP research. Critical

approaches, Ac Lits scholarship, and the growing ELF research agenda on the continent, have highlighted the need to explore the complexity of scholars' goals, interests, disciplines, and sociopolitical contexts beyond these binaries (Curry & Lillis, 2019). Using the theoretical framework of DST, this dissertation examined language as a dynamic system, involving cognition, embodiment, human interaction, and society. The epistemological approaches undergirding this dissertation, which foreground interaction and context, expanded understanding of the rich and dynamic system that is EILS. In the introduction, I proposed that the macro, meso, and micro levels of analysis were convenient abstractions that might be better conceptualized as interdependent complex systems nested one within the other. At the micro level, the development of Renzo's authorial voice is nested within and shaped by the conventions and constraints of his disciplinary community. For example, we saw how he adopted stronger authorial positions over time to more closely mirror the discourse of his disciplinary peers. In turn, the conventions and constraints of his disciplinary community are nested within and shaped by socio-political and market forces. For example, increased use of self-mention in the biological sciences has been attributed to technological and societal innovations in the post-war period (Halloran, 1984). In short, the three levels of analysis reflect the fundamental DST concept of complete interconnectedness by recognizing that variables impact other variables in and across systems (de Bot, Lowie & Verspoor, 2007). The levels of analysis also reflect the DST theory that agency is spatially-temporally situated as individuals orient to social and material affordances (Larsen-Freeman, 2019). In conclusion, using the frame of DST and the epistemological approaches of CDA, CA, and UBL, I hope to have shown that the individual writers, the socio-rhetorical communities,

and the socio-political contexts examined in this dissertation both shape and are shaped by

one another in a dynamic and complex interplay of language, identity, and voice.

REFERENCES

Airey, J. (2016). EAP, EMI or CLIL?. In K. Hyland & P. Shaw (Eds.), The Routledge handbook of
English for academic purposes. (pp.71-83.) Routledge.

Anthony, L. (2014). AntConc (Version 3.4.4) [Macintosh OS X]. Tokyo, Japan: Waseda University.
Available from http://www.laurenceanthony.net/software

Ammon, U. (2011). *The dominance of English as a language of science: Effects on other languages
and language communities*. De Gruyter Mouton. https://doi.org/10.1515/9783110869484

Aston, G. (1997). Small and large corpora in language learning. In B. Lewandowska-Tomaszczyk
and P. Melia (Eds.), *Practical applications in language corpora* (pp. 51–62). Lodz
University Press.

Baisa, V., Michelfeit, J., Medved, M. & Jakubíček M. (2016). European Union language resources in
Sketch Engine. In *the proceedings of tenth international conference on language resources
and evaluation (LREC'16)*. European Language Resources Association (ELRA).

Baker, W. (2015). *Culture and identity through English as a lingua franca: Rethinking concepts and
goals in intercultural communication*. De Gruyter Mouton.
https://doi.org/10.1515/9781501502149

Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008).
A useful methodological synergy? Combining critical discourse analysis and corpus
linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse
& society*, *19*(3), 273-306.

Banchoff, T. (2002). Institutions, inertia and European Union research policy. *Journal of Common
Market Studies*, *40*(1), 1-21.

.

Baumgarten, N., & House, J. (2010). I think and I don't know in English as lingua franca and native English discourse. *Journal of Pragmatics*, *42*(5), 1184-1200

Baumgarten, N., & House, J. (2007). Speaker stances in native and non-native English conversation. In Thije, J. D. T., & Zeevaert, L. (Eds.), Receptive multilingualism : Linguistic analyses, language policies and didactic concepts (pp.195-216). John Benjamins Publishing Company.

Benesch, S. (2001). *Critical English for academic purposes: Theory, politics, and practice*. Routledge.

Benesch, S. (2009). Theorizing and practicing critical English for academic purposes. *Journal of English for Academic Purposes*, *8*(2), 81-85.

Bennet, K. (2015). Towards an epistemological monoculture: Mechanisms of epistemicide in European research publication. In R. Plo Alastrué & C. Pérez-Llantada (Eds.), *English as a scientific and research language: Debates and* discourses (pp. 9-36). Walter de Gruyter GmbH & Co KG.

Benveniste, E. (1971). Subjectivity in language. *Problems in general linguistics*, 1, (pp. 223-30). University of Miami Press

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes, 5*(2), 97-116.

Björkman, B. (2014). Language ideology or language practice? An analysis of language policy documents at Swedish universities. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, *33*(3-4), 335-363.

Block, D. (2007). The rise of identity in SLA research, post Firth and Wagner (1997). *The Modern Language Journal*, *91*, 863-876.

Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.38, retrieved 2 January 2021 from http://www.praat.org/

Bolden, G. B. (2018). Speaking 'out of turn': Epistemics in action in other-initiated repair. *Discourse Studies*, *20*(1), 142–162.

Brumfit C. (2004) Language and higher education: Two current challenges. *Arts and Humanities in Higher Education*, *3*(2),163-173.

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse studies*, *7*(4-5), 585-614.

Burgess, A., & Ivanič, R. (2010). Writing and being written: Issues of identity across timescales. *Written Communication*, 27(2), 228–255.

Bybee, J. L. (2013). Usage-based theory and exemplar representations of constructions. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 49–69). Oxford University Press.

Canagarajah, A. S. (2006). The place of world Englishes in composition: Pluralization continued. *College composition and communication*, *57*(4), 586–619. http://www.jstor.org/stable/20456910

Casanave, C. P. (2002). *Writing games: Multicultural case studies of academic literacy practices in higher education*. Lawrence Erlbaum Associates.

Cheng, A. (2006). Understanding learners and learning in ESP genre-based writing instruction. English for Specific Purposes, *25*(1), 76–89. https://doi.org/10.1016/j.esp.2005.07.002

Cherry, R. D. (1998). Ethos versus persona: Self-representation in written discourse. *Written Communication*, *15*(3), 384–410.

Clayman, S. E., & Heritage, J. (2002). Questioning presidents: Journalistic deference

and adversarialness in the press conferences of US Presidents Eisenhower and Reagan. *Journal of Communication*, *52*(4), 749–775.

Clayman, S. E., & Raymond, C. W. (2021). You know as invoking alignment: A generic resource for emerging problems of understanding and affiliation. *Journal of Pragmatics*, *182*, 293-309.

Clement, T. P. (2014). Authorship matrix: a rational approach to quantify individual contributions and responsibilities in multi-author scientific articles. *Science and Engineering Ethics*, *20*(2), 345–361.

Cogo, A., & Jenkins, J. (2010). English as a lingua franca in Europe: A mismatch between policy and practice. *European Journal of Language Policy*, *2*(2), 271.

Coleman, J. A. (2006). English-medium teaching in European higher education. *Language Teaching*, *39*(1), 1-14.

Conrad, S. (2018). The use of passives and impersonal style in civil engineering writing. *Journal of Business and Technical Communication*, *32*(1), 38–76.

Cotos, E., Link, S., & Huffman, S. (2016). Studying disciplinary corpora to teach the craft of Discussion. *Writing & Pedagogy*, *8*(1), 33–64. https://doi:10.1558/wap.v8i1.2766

Curaj, A., Deca, L., & Pricopie, R. (2018). *European higher education area: The impact of past and future policies*. Springer Nature.

Curry, M. J., & Lillis, T. (2004). Multilingual scholars and the imperative to publish in English: Negotiating interests, demands, and rewards. *TESOL Quarterly*, *38*(4), 663-688.

Curry, M. J., & Lillis, T. M. (2010). Academic research networks: Accessing resources for English-medium publishing. *English for Specific Purposes*, *29*(4), 281-295.

Curry, M. J., & Lillis, T. (2019). Unpacking the lore on multilingual scholars publishing in English: A discussion paper. *Publications*, *7*(2), 1-14.

Davis, M. (2019). Publishing research as an EAP practitioner: Opportunities and threats. *Journal of English for Academic Purposes*, *39*, 72–86.

De Bot, K., Verspoor, M., & Lowie, W. (2005). Dynamic systems theory and applied linguistics: The ultimate "so what"?. *International Journal of Applied Linguistics*, *15*(1), 116-118.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, *10*(1), 7-21.

De Elera, Á. (2006). The European research area: On the way towards a European scientific community? *European Law Journal*, *12*(5), 559-574.

Dimova, S., Hultgren, A. K., & Jensen, C. (Eds.). (2015). *English-medium instruction in European higher education*. Walter de Gruyter GmbH & Co KG.

Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, *100*(S1), 19-47.

Dressen-Hammouda, D. (2008). From novice to disciplinary expert: Disciplinary identity and genre mastery. *English for Specific Purposes*, *27*(2), 233–252. doi:10.1016/j.esp.2007.07.006

Dressen-Hammouda, D. (2012). Measuring the construction of discoursal expertise through corpus-based genre analysis. *Corpus-Informed Research and Learning in ESP: Issues and Applications*, 193–214.

Dressen-Hammouda, D. (2014). Measuring the voice of disciplinarity in scientific writing: A longitudinal exploration of experienced writers in geology. *English for Specific Purposes*, *34*(1), 14–25.

Duguid, A., & Partington, A. (2018). You don't know what you're missing. Or do you? In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 38-59). Routledge.

European Commission (2021, November 13). *What is ERA?* European research area. https://ec.europa.eu/info/research-and-innovation/strategy/era_en

European Commission (2022, April 22). Important characteristics of the Erasmus+ Programme. Erasmus+. https://erasmus-plus.ec.europa.eu/programme-guide/part-a/important-characteristics-of-the-erasmus-programme

Evans, L. (2018) Reshaping the EHEA after the demise of neoliberalism: A UK informed perspective. In A. Curaj, L. Deca & R. Pricopie (Eds.), *European higher education area: The impact of past and future policies* (pp. 23-42). Springer Nature.

Fairclough, N. & Wodak, R. (1997). Critical discourse analysis. An overview. In T. A. van Dijk (Ed.), *Discourse as social interaction* (pp. 258-284). Sage.

Ferguson, G. (2015). Introduction: Attitudes to English. In A. Linn, N. Bermel, G. Ferguson (Eds.), *Attitudes towards English in Europe* (pp. 3-23). De Gruyter Mouton.

Ferguson, G. (2018). European language policy and English as a lingua franca: a critique of Van Parijs's 'linguistic justice'. In Z. Tatsioka, B. Seidlhofer, N. Sifakis & G. Ferguson (Eds.), *Using English as a lingua franca in education in Europe: English in Europe* (pp. 28-48). De Gruyter Mouton. https://doi.org/10.1515/9781501503115-003

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, *81*(3), 285-300.

Flick, U. (2009). *An introduction to qualitative research* (4th ed.). Sage Publications Limited.

Fløttum, K. (2010). Linguistically marked cultural identity in research articles. *Linguistic Insights Studies in Language and Communication, 125*, 267–280.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes, 24*(3), 321–332

Flowerdew, J. (2019). The linguistic disadvantage of scholars who write in English as an additional language: Myth or reality. *Language Teaching*, *52*(2), 249-260.

Fogal, G. G. (2017). Tracking microgenetic changes in authorial voice development from a complexity theory perspective. *Applied Linguistics*, 1–25.

Fogal, G. G. (2019). Investigating variability in L2 development: Extending a complexity theory perspective on L2 writing studies and authorial voice. *Applied Linguistics*, *41*(4), 1–27.

Foroni, M. (2015). Bridging education, research and innovation: The pivotal role of doctoral training in the European Higher Education Area, In A. Curaj, L. Matei, R. Pricopie, J. Salmi & P. Scott (Eds.), *The European higher education area: Between critical reflections and future policies* (pp. 541-544). Springer Nature.

Foucault, M. (1988). Technologies of the self. In L.H. Martin, H. Gutman & P.H. Hutton (Eds.), *Technologies of the self: a seminar with Michel Foucault*. University of Massachusetts Press.

Fox Tree, J.E. & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics, 34*(6), 727–747. https://doi.org/10.1016/S0378-2166(02)00027-

Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996-2005. *Journal of English Linguistics*, 36(1), 5–38. https://doi.org/10.1177/0075424207311247

Gardner, R. (1994). Conversation analysis: Some thoughts on its applicability to applied linguistics. *Australian Review of Applied Linguistics*. Supplement Series, 11(1), 97-118.

Giannoni, D. S. (2008). Medical writing at the periphery: The case of Italian journal editorials. *Journal of English for Academic Purposes*, 7(2), 97-107.

Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge University Press.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Groom, N. (2005). Pattern and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes*, 4(3), 257–277. https://doi.org/10.1016/j.jeap.2005.03.002

Gubrium, J. F., & Holstein, J. A. (2001). *Handbook of interview research: Context and method*. Sage Publications.

Guillén, M. F. (2007). Complex organizations and power. *International Public Management Journal*, 10(2), 167-171.

Halloran, S. M. (1984). The birth of molecular biology: An essay in the rhetorical criticism of scientific discourse. *Rhetoric Review*, 3(1), 70-83.

Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes*, 38, 1–13.

Harwood, N. (2005). "Nowhere has anyone attempted ... In this article I aim to do just that" A corpus-based study of self-promotional I and we in academic writing across four disciplines. *Journal of Pragmatics*, 37(8), 1207–1231. doi.org/10.1016/j.pragma.2005.01.012

Hayano, K. (2013). Question design in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 392–407). Wiley-Blackwell.

Hepburn, A. & Bolden, G. (2013). The Conversation Analytic Approach to Transcription. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 104–112) Wiley-Blackwell.

Heritage, J. (2013). Epistemics in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 613–636). Wiley-Blackwell.

Heritage, J., & Stivers, T. (2013). Conversation Analysis and sociology. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 613–636). Wiley-Blackwell.

Hewings, M., & Hewings, A. (2002). "It is interesting to note that...": A comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes*, *21*(4), 367–383. https://doi.org/10.1016/S0889-4906(01)00016-3

Holstein, J. A., & Gubrium, J. F. (1995). *The active interview*. Sage publications

House, J. (2009). Subjectivity in English as lingua franca discourse: The case of you know. *Intercultural Pragmatics*, *6*(2), 171-193.

House, J. (2013). Developing pragmatic competence in English as a lingua franca: Using discourse markers to express (inter) subjectivity and connectivity. *Journal of Pragmatics*, *59*, 57-67.

Hultgren, A. K. (2014). English language use at the internationalised universities of Northern Europe: Is there a correlation between Englishisation and world rank?. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, *33*(3-4), 389-411.

Hultgren, A. K. (2020). Global English: From "Tyrannosaurus rex" to "red herring". *Nordic Journal of English Studies*, 19(3), 10-34.

Hultgren, A. K., Gregersen, F., & Thøgersen, J. (Eds.). (2014). English in Nordic universities: Ideologies and practices (vol. 5). John Benjamins Publishing Company.

Hultgren, A. K., Jensen, C., & Dimova, S. (2015). English-medium instruction in European higher education: From the north to the south. In S. Dimova, A. K. Hultgren & C. Jensen (Eds.), English-medium instruction in European higher education (pp 1-15). Walter de Gruyter GmbH & Co KG.

Hyland, K. (2001). Humble servants of the discpline?: Self-mention in research articles. *English for Specific Purposes*, 20, 207–226.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, *34*(8), 1091–1112.

Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, *7*, 173–192. doi:10.1177/1461445605050365

Hyland, K. (2011). Disciplines and discourses: Social interactions in the construction of knowledge. *Writing in knowledge societies*. *Perspectives on Writing*, 193-214.

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing*, *31*, 58–69. doi:10.1016/j.jslw.2016.01.005

Hyland, K. (2018). Sympathy for the devil? A defence of EAP. *Language Teaching*, *51*(3), 383-399.

Hyland, K., & Guinda, C. S. (Eds.). (2012). *Stance and voice in written academic genres*. Palgrave Macmillan

Hyland, K., & Jiang, F. (2017). Is academic writing becoming more informal? *English for Specific Purposes*, *45*, 40–51. https://doi.org/10.1016/j.esp.2016.09.001

Hyland, K., & Shaw, P. (Eds.). (2016). *The Routledge handbook of English for academic purposes*, Routledge.

Hyland, K., & Tse, P. (2005a). Hooking the reader: A corpus study of evaluative that in abstracts. *English for Specific Purposes*, *24*(2), 123 –139. doi:10.1016/j.esp.2004.02.002

Hyland, K., & Tse, P. (2005b). Evaluative that constructions: Signaling stance in research abstracts. *Functions of Lanugage*, *12*(1), 39–63.

Hynninen, N., & Kuteeva, M. (2017). "Good" and "acceptable" English in L2 research writing: Ideals and realities in history and computer science. *Journal of English for academic purposes*, *30*, 53–65.

Iksra, K. A. (2022) *Language Policy*. Factsheets on the European Union. https://www.europarl.europa.eu/factsheets/en/sheet/142/language-policy

Jefferson, G. (2004). Glossary of transcription symbols with an introduction. In G. H. Lerner (Ed.), Conversation Analysis: Studies from the first generation (pp. 13–23). John Benjamins.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching*, *44*(3), 281-315.

Jenkins, J., Baker, W. & Dewey, M. (Eds). (2017) *The Routledge handbook of English as a lingua franca*. Routledge.

Jucker, A. H., & Smith, S. W. (1998). Interactive aspects of reference assignment in conversations. *Pragmatics & Cognition*, 6(1-2), 153-187.

Junnier, F. (2020). Finding voice in biology: A diachronic analysis of self-mention in the discussions of an L2 scholar. *Journal of English for Academic Purposes*, *47*, 100889.

Kärkkäinen, E. (2003). *Epistemic stance in English conversation. A description of its interactional functions, with a focus on I think*. John Benjamins

Karkkainen , E. (2007) The role of I guess in conversational stancetaking. In R. Englebretson (Ed.), *Stancetaking in discourse*. John Benjamins Publishing Company.

Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes*, 9(1), 3-20. https://doi.org/10.1111/j.1467-971X.1990.tb00683.x

Kasper, G., & Wagner, J. (2014). Conversation analysis in applied linguistics. *Annual Review of Applied Linguistics*, 34, 171-212.

Katsarova, I. (2019). *Multilingualism: The language of the European Union*. https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/642207/EPR_BRI(2019)642207_EN.

Kaufhold, K. (2017). Tracing interacting literacy practices in Master's dissertation writing. *London Review of Education*, *15*(1), 73-84.

Kaufhold, K., & McGrath, L. (2019). Revisiting the role of discipline in writing for publication in two social sciences. *Journal of English for Academic Purposes*, *40*, 115–128.

Koch Christensen, K. (2005). *Doctoral programmes for the European knowledge society*. Bologna Seminar. https://c3qa.iqaa.kz/en/project-documents/salzburg-principles

Krzyżanowski, M., & Wodak, R. (2011). Political strategies and language policies: The European Union Lisbon strategy and its implications for the EU's language and multilingualism policy. *Language Policy*, *10*(2), 115-136.

Kuteeva, M. (2015). Academic English as 'nobody's land': The research and publication practices of Swedish academics. In R. Plo Alastrué & C. Pérez-Llantada, (Eds.). English as a scientific and research language: Debates and discourses (pp. 261-280). Walter de Gruyter GmbH & Co KG.

Kuteeva, M., & Mauranen, A. (2014). Writing for publication in multilingual contexts: An introduction to the special issue. *Journal of English for Academic Purposes*, *13*(1), 1-4.

Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, *34*(4), 513–535.

Lehman, I. M., & Anderson, R. (2017). Academic Identities and Literacy Practices: A Few Remarks on the Influence of EAP Instruction on the Construction of Disciplinary Identities of Italian Tertiary Students. In D. Gabryś-Barker, D. Gałajda, A. Wojtaszek, P. Zakrajewski (Eds.), *Multiculturalism, multilingualism and the self. Second language learning and teaching*. Springer. https://doi.org/10.1007/978-3-319-56892-8_9

Larsen-Freeman, D. (1997). "Chaos/complexity science and second language acquisition." *Applied Linguistics 18*(2), 141–165.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, *27*(4), 590–619.

Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics?. *Language Teaching*, *45*(2), 202-214.

Lasagabaster, D. (2015). Language policy and language choice at European Universities: Is there really a 'choice'?. *European Journal of Applied Linguistics*, *3*(2), 255-276.

Lea, M. & Street, B. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education 23*(2), 157-172.

Lea M. & Street, B. (2006) The "Academic Literacies" Model: Theory and Applications. *Theory Into Practice*, 45(4), 368-377.

Lei, L., & Liu, D. (2019). Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics*, 40(*3*), 540–561.

Levinson, S.C. (1983). *Pragmatics*. Cambridge University Press

Li, P., Eskildsen, S. W., & Cadierno, T. (2014). Tracing an L2 learner's motion constructions over time: A usage-based classroom investigation. *The Modern Language Journal*, *98*(2), 612–628.

Li, Y., Flowerdew, J., & Cargill, M. (2018). Teaching English for research publication purposes to science students in China: A case study of an experienced teacher in the classroom. *Journal of English for Academic Purposes*, 35, 116–129.

Lillis, T. (2009). Bringing writers' voices to writing research. In A. Carter, T. Lillis & S. Parkin (Eds.), *Why writing matters: Issues of access and identity in writing research and pedagogy* (pp. 169-187). John Benjamins Publishing Company.

Lorés-Sanz, R. (2016). ELF in the making? Simplification and hybridity in abstract writing. *Journal of English as a Lingua Franca*, *5*(1), 53–81.

Macaro, E., Hultgren, A. K., Kirkpatrick, A., & Lasagabaster, D. (2019). English medium instruction: Global views and countries in focus. *Language Teaching*, *52*(2), 231-248.

Martín, P., Rey-Rocha, J., Burgess, S., & Moreno, A. I. (2014). Publishing research in English-language journals: Attitudes, strategies and difficulties of multilingual scholars of medicine. *Journal of English for Academic Purposes*, *16*, 57–67.

Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second language writing*, *10*(1-2), 35-53.

Matsuda, P. K. (2015). Identity in written discourse. *Annual Review of Applied Linguistics*, *35*, 140-159.

Mauranen, A., Hynninen, N., & Ranta, E. (2010). English as an academic lingua franca: The ELFA project. *English for Specific Purposes*, *29*(3), 183-190.

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Merriam, S. B. (1998). Qualitative research and case study applications in education. Jossey-Bass Publishers.

Moder, C. L., & Halleck, G. B. (1998). Framing the language proficiency interview as a speech event: native and non-native speakers' questions. In R. Young & A.W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 117-148). John Benjamins Publishing.

Montgomery, S. (2004). Of towers, walls, and fields: perspectives on language in science. *Science*, *303*(5662), 1333-1335.

Mur Dueñas, P. (2007). 'I/we focus on…': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, *6*(2), 143–162.

Mur-Dueñas, P. (2013). Spanish scholars' research article publishing process in English-medium journals: English used as a lingua franca?. *Journal of English as a Lingua Franca*, *2*(2), 315-340.

Myers, G. (1990). Writing biology: Texts in the social construction of scientific knowledge. University of Wisconsin Press

Neave, G., & Veiga, A. (2013). The Bologna Process: inception, 'take up' and familiarity. *Higher Education*, *66*(1), 59-77

Neumann, H., Leu, S., & McDonough, K. (2019). L2 writers' use of outside sources and the related challenges. *Journal of English for Academic Purposes*, *38*, 106-120.

Olson, K. (2016). *Essentials of qualitative interviewing*. Routledge.

Östman, J. O. (1981). *'You know': A discourse-functional study*. John Benjamins Publishing.

Ostman, J. O. & Trousdale, G. (2013). Dialects, discourse, and construction grammar. In T. Hoffmann & G. Trousdale, (Eds.), *The Oxford handbook of construction grammar* (pp. 476-489). Oxford University Press.

Paltridge, B. (1997). *Genre, frames and writing in research settings*. John Benjamins.

Pennycook, A. (1997). Vulgar pragmatism, critical pragmatism, and EAP. *English for Specific Purposes*, *16*(4), 253-269.

Pérez-Llantada, C. (2009). Shifting identities, textual responses and conflicting demands in knowledge construction processes. In M. Gotti (Ed.), *Commonality and individuality in academic discourse*, (pp.191–213). Peter Lang.

Pérez-Llantada, C. (2015). Teasing out the tensions between English monolingualism vs. plurilingualism in European academic and research settings in R. Plo Alastrué & C. Pérez-Llantada, (Eds.), *English as a Scientific and Research Language: Debates and Discourses* (pp. 353-362). Walter de Gruyter GmbH & Co KG.

Pérez-Llantada, C. (2018). ELF and linguistic diversity in EAP writing pedagogy: academic biliteracy in doctoral education. In Z. Tatsioka, B. Seidlhofer, N.C. Sifakis & G. Ferguson (Eds,). *Using English as a lingua franca in education in Europe*, (pp. 200-224). De Gruyter Mouton. https://doi-org.argo.library.okstate.edu/10.1515/9781501503115

Phillipson, R. (1996). Linguistic imperialism: African perspectives. *ELT Journal*, *50*(2), 160–167. https://doi.org/10.1093/ELT/50.2.160

Phillipson, R. (2006). English, a cuckoo in the European higher education nest of languages? *European Journal of English Studies 10*(1), 13–32.

Phillipson, R. (2015). English as threat or opportunity in European higher education in S. Dimova, A.K. Hultgren & C. Jensen (Eds.). *English-medium instruction in European higher education* (pp. 19 - 42). Walter de Gruyter GmbH & Co KG.

Phillipson, R. (2017). Myths and realities of European Union language policy. *World Englishes*, *36*(3), 347-349.

Plo Alastrué, R. (2015). Debates and discourses on English as an academic and research language in R. Plo Alastrué & C. Pérez-Llantada, (Eds.). *English as a scientific and research language: Debates and discourses* (pp. 3-8). Walter de Gruyter GmbH & Co KG.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shape. In J.M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 57–101). Cambridge University Press.

Pomerantz & Heritage, (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 234 – 249). Wiley-Blackwell

Potter, J., & Hepburn, A. (2012). Eight challenges for interview researchers. *Handbook of interview research*, *2*(1), 541-570.

Prior, M. T. (2018). Accomplishing "rapport" in qualitative research interviews: Empathic moments in interaction. *Applied Linguistics Review*, *9*(4), 487-511.

Prior, M. T., & Talmy, S. (2019). A discursive psychological approach to the analysis of talk and text in applied linguistics. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics*, (pp. 474-487). Routledge.

Randour, F., Perrez, J., & Reuchamps, M. (2020). Twenty years of research on political discourse: A systematic review and directions for future research. *Discourse & Society*, *31*(4), 428-443.

Roulston, K. (2011). Interview 'problems' as topics for analysis. *Applied Linguistics*, 32(1), 77-94.

Sacks, H. (1984). Notes on methodology. In J. M. Atkinson, & J. Heritage (Eds.) *Structures of social action: Studies in conversation analysis*, (pp. 21-27). Cambridge University Press.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turntaking for conversation. *Language*, 50, 696–735.

Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, *24*(2), 141-156.

Santulli, F. (2015). English in Italian universities: The language policy of PoliMi from theory to practice. In S. Dimova, A.K. Hultgren & C. Jensen (Eds,). *English-medium instruction in European higher education, (*pp. 269-290). Walter de Gruyter GmbH & Co KG.

Scheibman, J. (2002). *Point of view and grammar: Structural patterns of subjectivity in American English conversation*. John Benjamins Publishing.

Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press.

Schiffrin, D. (1994). *Approaches to discourse*. Blackwell Publishers.

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics, 11*(2), 133–158.

Seidlhofer, B., Breiteneder, A., & Pitzl, M. L. (2006). English as a lingua franca in Europe: Challenges for applied linguistics. *Annual review of applied linguistics, 26*, 3-34.

Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford University Press

Sert, O., & Seedhouse, P. (2011). Introduction: Conversation Analysis in Applied Linguistics. Novitas-ROYAL (Research on Youth and Language), 5(1), 1-14.

Shleykina, G., & Junnier, F. (2021). Translating authorial presence: First-person pronouns in research article abstracts. *Journal of English for Research Publication Purposes, 2*(2), 104-128.

Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. Routledge.

Sidnell, J. (2010). *Conversation Analysis: An introduction*, John Wiley & Sons.

Sidnell, J., & Stivers, T. (Eds.). (2013). *The handbook of Conversation Analysis*. Wiley-Blackwell

Silverman, D. (2017). How was it for you? The Interview Society and the irresistible rise of the (poorly analyzed) interview. *Qualitative Research, 17*(2), 144–158.

Soler-Carbonell, J., Saarinen, T., & Kibbermann, K. (2017). Multilayered perspectives on language policy in higher education: Finland, Estonia, and Latvia in comparison. *Journal of Multilingual and Multicultural Development, 38*(4), 301-314.

Spolsky, B. (2004). *Language policy*. Cambridge University Press.

Spolsky, B. (2019). A modified and enriched theory of language policy (and management). *Language Policy, 18*(3), 323-338.

Stenglin, M., & Cléirigh, C. (2020). Scientific grant application writing: Re/packaging text to enhance its impact. *Journal of English for Academic Purposes*, *44*, 100823.

Stivers, T. (2013). Sequence Organization. In J. Sidnell & T. Stivers (Eds.) *The handbook of Conversation Analysis*. (pp. 217–233). Wiley-Blackwell.

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on language and social interaction*, *43*(1), 3-31.

Stock, I., & Eik-Nes, N. L. (2016). Voice features in academic texts – A review of empirical studies. *Journal of English for Academic Purposes*, *24*, 89–99.

Swales, J. (1990). *Genre Analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. (2000). Languages for specific purposes, *Annual Review of Applied Linguistics*, *20*, 59-76. doi:10.1017/S0267190500200044

Swales, J. M. (2019). The futures of EAP genre studies: A personal viewpoint. *Journal of English for Academic Purposes*, *38*, 75-82.

Talmy, S. (2010), Qualitative Interviews in Applied Linguistics: From Research Instrument to Social Practice, *Annual Review of Applied Linguistics* 30, 128-148.

Tang, R., & John, S. (1999). The 'I' in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes*, *18*, S23–S39.

Tardy, C. M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of second language writing*, *15*(2), 79-101.

Tardy, C. M. (2012). Current conceptions of voice. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 34-48). Palgrave Macmillan, London.

Thompson, P. (2009). The writing of British undergraduates. In M. Gotti (Ed.) *Commonality and Individuality in Academic Discourse,* (pp. 53 – 82). Peter Lang.

Tudor, I. (2005). *The challenge of the Bologna Process for higher education language teaching in Europe*. Universite Libre de Bruxelles.

Van Dijk, T. A. (1999). Critical discourse analysis and conversation analysis. *Discourse & Society*, *10*(4), 459-460.

Van Dijk, T.A. (2001). Multidisciplinary CDA: A plea for diversity. In R. Wodak &

M. Meyer (Eds.), *Methods of Critical Discourse Analysis*, (pp. 95–120). SAGE.

Van Parijs, P. (2011). *Linguistic justice for Europe and for the world*. Oxford University Press

Velayutham, S. (2021). Understanding How the Marketisation of Higher Education Contributes to Increased Income Inequality and Decreased Social Mobility. In J. D. Branch & B. Christiansen (Eds,). *The Marketisation of Higher Education: Concepts, Cases, and Criticisms*. (pp. 371-395). Palgrave Macmillan.

Villares, R. (2019). The role of language policy documents in the internationalisation of multilingual higher education: An exploratory corpus-based study. Languages, 4, 1-9.

Walker, G. (2013) Phonetics and prosody in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 440–456). Wiley-Blackwell.

Walková, M. (2019). A three-dimensional model of personal self-mention in research papers. *English for Specific Purposes*, *53*, 60–73.

Williams, I. A. (2012). Self-reference in biomedical research article discussions: Further evidence for cross-cultural diversity in academic and scientific discourse. *International Journal of Corpus Linguistics*, *17*(4), 546–583.

Whitaker, E. M., & Atkinson, P. (2019). Authenticity and the interview: a positive response to a radical critique. *Qualitative Research*, *19*(6), 619–634.

Widdowson, H. G. (2002). The ownership of English. In V. Zamel & R. Spack (Eds,). *Enriching Esol pedagogy: Readings and Activities for Engagement, Reflection, and Inquiry* (pp. 381-392). Lawrence Erlbaum Associates.

Wodak, R. (2001). What CDA is about–a summary of its history, important concepts and its developments. In R. Wodak, & M. Meyer (Eds.), *Methods of critical discourse analysis* (pp. 1-13). SAGE Publications Ltd. https://dx.doi.org/10.4135/9780857028020.n1

Wodak, R. (2011). Critical linguistics and critical discourse analysis in J. Zienkowski, J.O. Östman & J. Verschueren (Eds,). *Discursive pragmatics* (pp. 50-70). John Benjamins Publishing.

Wooffitt, R. (2005). *Conversation analysis and discourse analysis: A comparative and critical introduction*. Sage.

Wooffitt, R., & Widdicombe, S. (2006). Interaction in interviews. In P. Drew, G. Raymond & D. Weinberg (Eds.), *Talk and interaction in social research methods* (pp. 28-49). Sage Publications.

Xu, L., & Zhang, L. J. (2019). L2 doctoral students' experiences in thesis writing in an English-medium university in New Zealand. *Journal of English for Academic Purposes*, *41*, 100779.

Yin, R. K. (2014). *Case study research: Design and methods*. CA: Sage Publications.

Zhang, X., Ardasheva, Y., & Austin, B. W. (2020). Self-efficacy and English public speaking performance: A mixed method approach. *English for Specific Purposes*, *59*, 1–16.

APPENDICES

APPENDIX A: EHEA Focal Corpus

Code	Year	Document Title		Words
EHEA_1	1998	Sorbonne Declaration	Joint declaration on harmonisation of the architecture of the European higher education system	855
EHEA_2	1999	Bologna Declaration	Joint declaration of the European Ministers of Education	960
EHEA_3	2001	Prague Communiqué	Towards the European Higher Education Area: Communiqué of the meeting of European ministers in charge of higher education	1758
EHEA_4	2003	Berlin Communiqué	"Realising the European Higher Education Area" Communiqué of the Conference of Ministers responsible for Higher Education	3048
EHEA_5	2005	Bergen Communiqué	The European Higher Education Area - Achieving the Goals Communiqué of the Conference of European Ministers Responsible for Higher Education	2371
EHEA_6	2007	London Communiqué	Towards the European Higher Education Area: responding to challenges in a globalised world	3126
EHEA_7	2009	Leuven Communiqué	The Bologna Process 2020 - The European Higher Education Area in the new decade. Communiqué of the Conference of European Ministers Responsible for Higher Education	2641
EHEA_8	2010	Budapest-Vienna Declaration	Budapest-Vienna Declaration on the European Higher Education Area	947
EHEA_9	2012	Bucharest Communiqué	Making the Most of Our Potential: Consolidating the European Higher Education Area	2889
EHEA_10	2015	Yerevan Communiqué	No title	1913
EHEA_11	2018	Paris Communiqué	Conférence ministérielle européene pour l'enseignement supérieur	2233
EHEA_12	2020	Rome Communiqué	Rome Ministerial Communiqué	3200

APPENDIX B: *ERA Focal Corpus*

Code	Year	Document Type	Title	Words
ERA_1	**2000**	Communication from the European Commission	Towards a European research area	10,862
ERA_2	**2000**	Communication from the European Commission	Making a reality of The European Research Area: Guidelines for EU research activities (2002-2006)	6,680
ERA_3	**2001**	Communication from the European Commission	Fulfilling the JRC's mission in the European Research Area	3,326
ERA_4	**2001**	Communication from the European Commission	The Framework Programme and the European Research Area: application of Article 169 and the networking of national programmes	2412
ERA_5	**2001**	Communication from the European Commission	Mobility Strategy For The European Research Area	5,822
ERA_6	**2001**	Communication from the European Commission	The International Dimension Of The European Research Area	7,519
ERA_7	**2001**	Communication from the European Commission	The Regional Dimension of the European Research Area	12,391
ERA_8	**2002**	Communication from the European Commission	The European Research Area: Providing New Momentum Strengthening Reorienting Opening Up Perspectives	8,321

ERA_9	2003	Communication from the European Commission	Researchers In The European Research Area: One Profession, Multiple Careers	10,124
ERA_10	2005	Communication from the European Commission	Building the ERA of knowledge for growth	6,339
ERA_11	2007	Green Paper	The European Research Area: New Perspectives	8,298
ERA_12	2012	Communication from the European Commission	A Reinforced European Research Area: Partnership for Excellence and Growth	5,385
ERA_13	2013	Report from the European Commission	1st Progress Report	4,110
ERA_14	2014	Report from the European Commission	2nd Progress Report	3,947
ERA_15	2016	Report from the European Commission	3rd Progress Report	5,023
ERA_16	2017	Report from the European Commission	The European Research Area: time for implementation and monitoring progress	4,593
ERA_17	2018	Report from the European Commission	4th Progress Report	4,670
ERA_18	2019	Report from the European Commission	The European Research Area: advancing together the Europe of	4,670
ERA_19	2020	Communication from the European Commission	A new ERA for Research and Innovation	9,427

APPENDIX C: Coding Sheet

1. CODING BY LANGUAGE POLICY DOMAIN

Spolsky (2004; 2019) suggests that language policy can be characterised by three independent but interconnected components

(CODE A) Language Practices

DESCRIPTION OF LANGUAGE USE: Discourse that refers to the actual use of different languages/language habits in different academic situations, that is, what languages people are using/language habits in use.

Example: *"The assessments are difficult to cross-reference because they are not carried out in various languages spoken in the institution"*

Example: *"Language practices are uneven between stakeholders resulting in …."*

(CODE B) Language Beliefs / Ideology

DISCUSSION OF LANGUAGE ATTITUDE: Discourse that refers to beliefs/attitudes about language and language use, as well as the speakers' language attitudes. Abstract discussion of language matters – no direct /no specific details given of what should be done *or* what is actually happening.

Example: *"We believe linguistic diversity is fundamental to the aims of the EU"*

"Language and culture form important aspects of our shared European tradition"

(CODE C) Language Intervention, Planning or Management,

PRESCRIPTON OF SPECIFIC LANGUAGE PLANS: Discourse that "refers to the formulation and proclamation of an explicit plan or policy, (Spolsky 2004: 11). The way in which stakeholders' attempt, either implicitly or explicitly, to shape language ideology and practices.

Example: *"Language practices should be simplified to streamline the unwieldy recruitment process for all candidates"*

"Linguistic competency requirements will be altered to account for this increased mobility of researchers"

2. CODING BY ATTITUDINAL STANCE

Stance can be expressed to differing degrees by using value-laden word choice and grammatical devices (Biber, 2006). Code the data according to semantically-marked attitudinal stance, classified along positive – negative dimensions with attention to markers such as stance nouns (i.e. cooperation/barrier), stance verbs (i.e. foster/block) attitudinal adjectives (i.e. unequalled/unwieldy) and attitudinal adverbs (i.e. importantly/worryingly).

(A) POSITIVE STANCE (CODE +)

> *Example:*   *"Fostering linguistic diversity is fundamental to the aims of the EU"*

> *Example*   *"Language and culture form important aspects of our shared European tradition"*

(B) NEGATIVE STANCE (CODE–)

> *Example: "Language barriers should be addressed to streamline the unwieldy recruitment process for all   candidates"*

> *Example: "Language practices are uneven between stakeholder resulting in obstacles f or…."*

(C) NEUTRAL STANCE (CODE N)

> *Example: "Diplomas should be issued in a widely spoken European language"*

> *Example: "Pension rights and language transferability are on-going topics"*

APPENDIX D: Interview Guide

PART ONE – INDIVIDUAL FACTORS

Literacy History

- o Can you tell me something about your experiences of writing for academic purposes (for example essays or reports) in your native language at either the high school or at university level?
  - Was it something you enjoyed?
  - Do you recall any particular type of instruction you received?

Language Acquisition

- o Can you tell me something about your experiences of learning to write in English for academic purposes?
- o Did you receive any kind of formal instruction in high school or at university? If so, what was it?

Attitude to Writing for Publication

- o Is there anyone you admire as a scientific writer in your field of expertise?
- o Have you ever read a published article and thought – this is really well written?
- o Have you ever felt disadvantaged, with regards to writing for publication, by the fact that English isn't your first/native language?
- o Do you write on a regular basis for academic purposes? For example, is time for writing part of your weekly routine?

PART TWO – SOCIAL FACTORS

** This next section of the interview will be conducted in relation to each of the research articles chosen by participants **

Institutional Position

- o Can you tell me something about where you were in your academic career when you wrote this research article?

Can you tell me about your investment in this particular line of research? For example, is/was it your main area of research; is/was it funded by an outside agency; was it your dissertation project?

Co-Authorship Dynamics

- o Can you tell me something about your research article co-authors – in terms of their language backgrounds, experience of writing for publication?
- o Can you tell me about the manuscript preparation process and how you wrote this research article? For example, the role each person took in writing the article, in revisions/ the length of time it took to write and have it published etc.

Disciplinary Norms

- o How often do you read other research articles? For example, is time for reading part of your weekly routine?
- o Scientific writing has been described as impersonal and factual. What words would you use to describe the style of writing in your discipline?
- o Can you recall a time when you thought that your style of writing or presenting your work was different to that of your peers?
- o Is it important to follow the conventions of your discipline when writing a research article for publication?

PART THREE – AUTHORIAL SELF

** This next section of the interview will be conducted in relation to the discussion section of each of the research articles chosen by participants **

- o When you write a discussion section what are your main goals?
- o Are you satisfied with this discussion – does it achieve the goals we just discussed?
- o Is there anything you notice about the way it's written that you would change now?
- o ** some excerpts will be discussed in relation to authorial presence i.e. I notice that you often use passive voice (provide examples)/ I noticed that you often use mitigating verbs (provide examples) etc**… what effect do these features have?
- o Do you think the reader of this discussion learns anything about you as a scientist when reading it?
- o Is there anything you'd like to add about any of the topics we discussed or any other issues you think are important which we haven't covered in the interview so far?

APPENDIX E: Participant Consent Form

ADULT CONSENT FORM

OKLAHOMA STATE UNIVERSITY

Study Title:
Disciplinary Identity Construction in Biology - Second Language Writers' Perceptions of
Authorial Self

Background Information:
You are invited to be in a research study about the ways in which scientists, who write in English
as an additional language for research publication purposes, think about their role as authors in
their published works, and how these perceptions are influenced by individual and social factors
such as literacy backgrounds, academic position, disciplinary norms, and co-authorship dynamics.

Investigators:
This project will be conducted by Frances Junnier, PhD Student TESL/Applied Linguistics,
Oklahoma State University under the direction of Dr. Carol Moder, TESL/Applied Linguistics,
English Department, Oklahoma State University.

Procedures:
If you agree to be in this study, I would ask you to do the following: 1) choose two of your
published research articles to be used in the study 2) participate in an audio-recorded interview
about your literacy background, your career path, your discipline's conventions and your
experiences of co-authorship as they relate to your role as a scientific writer and 3) review the
transcripts of those interviews and the subsequent analysis of those transcripts for their accuracy.

Participation in the study involves the following time commitment:
No more than 4 hours. This includes reading this consent form, scheduling the interview/s, the
interview/s, and review of the transcript/s and the analysis for their accuracy.

Risks of being in the study:
There are no known risks associated with this study, which are greater than those ordinarily
encountered in daily life.

Benefits to participation are:
Participation in the study may well serve as a self-reflective and awareness raising exercise. More broadly, this study may help the researcher learn more about the individual and social factors, which influence disciplinary identity construction in biology.

Compensation:
There is no compensation associated to this study.

Confidentiality
To protect your identity, pseudonyms will be used in any publications, presentations, or reports resulting from this study and references to geographical locations or specific academic institutions will be generalized. <u>However, as excerpts of your published work may be used alongside some interview data, I cannot completely guarantee that your identity will not be indirectly revealed.</u>   I will collect your information through audio recordings and field notes. The audio recordings will be transcribed. The recordings will be deleted after the transcription is completed and verified. This informed consent form will be kept for 3 years after the study is complete, and then it will be destroyed. Your data collected as part of this research project, may be used or distributed for future research studies.

Contacts:
If you have any questions about the research, please feel free to contact the PI of the study: Frances Junnier, Phone: or Email: frances.junnier@okstate.edu . If you have questions about your rights as a research volunteer, you may contact the IRB Office at 223 Scott Hall, Stillwater, OK 74078, 405-744-3377 or irb@okstate.edu

Participant Rights:

Your participation is voluntary and there is no penalty for refusal to participate. You are free to withdraw consent and participation in this study at any time, without penalty. During the interview you can skip any questions that make you uncomfortable and can leave the interview at any time.

Statement of Consent:

Indicate Yes or No:

I affirm that I am 18 years of age or older
___Yes ___No

I give consent to be audiotaped during this study.
___Yes ___No

I give consent for my data to be used in future research studies:
___Yes ___No

I have read and fully understand this consent form. I sign it freely and voluntarily. I have had the

opportunity to ask questions and have had my questions answered.
A copy of this form will be given to me. I hereby give permission for my participation in this study.

_____ _____

Signature of Participant   Date

I certify that I have personally explained this document before requesting that the participant sign it.

_____ _____

Signature of Researcher   Date

APPENDIX F: Institutional Review Board Approval Letter

Oklahoma State University Institutional Review Board

Date: 08/30/2018

Application Number: AS-18-98

Proposal Title: Disciplinary Identity Construction in Molecular
                               Biology - Second Language Writers' Perceptions
                               of Authorial Self

Principal Investigator: FRAN JUNNIER
Co-Investigator(s):
Faculty Adviser: CAROL MODER
Project Coordinator:
Research Assistant(s):

Processed as: Exempt

Status Recommended by Reviewer(s): Approved

The IRB application referenced above has been approved. It is the judgment of the
reviewers that the rights and welfare of individuals who may be asked to participate in
this study will be respected, and that the research will be conducted in a manner
consistent with the IRB requirements as outlined in section 45 CFR 46.

The final versions of any recruitment, consent and assent documents bearing the IRB
approval stamp are available for download from IRBManager. These are the versions that
must be used during the study.

.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be approved by the IRB. Protocol modifications requiring approval may include changes to the title, PI, adviser, other research personnel, funding status or sponsor, subject population composition or size, recruitment, inclusion/exclusion criteria, research site, research procedures and consent/assent process or forms.
2. Submit a request for continuation if the study extends beyond the approval period. This continuation must receive IRB review and approval before the research can continue.
3. Report any unanticipated and/or adverse events to the IRB Office promptly.
4. Notify the IRB office when your research project is complete or when you are no longer affiliated with Oklahoma State University.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact the IRB Office at 223 Scott Hall (phone: 405-744-3377, irb@okstate.edu).

Sincerely,

Hugh Crethar, Chair Institutional Review Board

APPENDIX G: Transcription Conventions
(adapted from Jefferson, 2004 & Hepburn & Bolden, 2013)

.	A period indicates a falling, or final, intonation contour, not necessarily the end of a sentence
?	A question mark indicates rising intonation, not necessarily a question
,	A comma indicates 'continuing' intonation, not necessarily a clause boundary
?,	A question mark followed by a comma indicates intonation rise between a comma and a question mark
:::	Colons indicate stretching of the preceding sound, proportional to the number of colons
word	underlining is used to indicate some form of stress or emphasis by increased amplitude and higher pitch
wo:rd	underlined element followed by a colon indicates an up-down intonation contour through the word
wor:d	A colon underlined indicates pitch movement sliding up through the word
^	Caret indicates a sharp rise in pitch
|	Pipe indicates sharp fall in pitch
> <	A combination of 'more than' and 'less than' symbols indicates that the talk between them is produced noticeably quicker than surrounding talk

< >	In the reverse order, they indicate that a stretch of talk is markedly slowed or drawn out
#	Hash indicates creaky voice
£	Pound sign indicates smiley voice
=	Equal sign indicates no break or delay between the words thereby connected
-	A hyphen after a word or a part of a word indicates a cut-off or self interruption word
WOrd	Upper case indicates loudness
°	Degree signs indicate talk that is markedly quiet or soft
.h	Hearable inhalation (in-breath) is shown by a period before an h-.
(())	Double parentheses enclose descriptions of conduct
(word)	When all or part of an utterance is in parentheses, this indicates uncertainty on the transcriber's part
()	Empty parentheses indicate that something is being said, but no hearing can be achieved
(0.5)	Numbers in parentheses indicate silence in tenths of a second
(.)	A dot in parentheses indicated a 'micropause', hearable but not readily measurable; ordinarily - less than 2/10 of a second
[	Separate left square brackets, one above the other on two successive lines with utterances by different speakers indicate a point of overlap onset
]	Separate right square brackets, one above the other on two successive lines with utterances by different speakers indicate a point at which two overlapping utterances both end or where one ends while the other continues

APPENDIX H: Interview Excerpts

Interviewer   em and I also wonder, have you ever felt disadvantaged with regards to writing for publication by the

Amelia   oy

Interviewer   fact that English isn't your first language.

Amelia   yeah every single time even now even now
I always feel like I'm not good enough or you know my ideas are not good enough or my English is not good enough

but at some point I need to make peace with my brain and just do it

and then I always have a couple of people that are native speakers so they can help me with English or you know some time is also a way to express a sentence that is not super linear mine is very convoluted and I need to put things linearly ((laughs))

but yeah all the time I think it will never change even if I am going to be 60 or 70 and I have done this job for forever it's just because there are people very confident and other people not and I am the second type ((laughs))

Interviewer   umm do you think that your em English well let me put it this way

Have you felt disadvantaged with regards to writing for publication by the fact that English isn't your first or native language

Maria   okay so [laughs] I feel that I'm losing at ton of time you know just looking for words that's synonyms or you know like they will not

((connection lost during interview 12:51 – 13:49))

Interviewer   hey sorry I don't know why I lost you

Maria	yeah me neither I'm sorry maybe it's the a bad connection.
Interviewer	yeah hopefully I don't happen again so sorry I you were you were basically saying that you feel like you just eh lose a ton of time in terms of of writing so looking for
Maria	yeah
Interviewer	synonyms things like that yeah.
Maria	yeah so that's it yep so that's basically it.
	like yeah so it's just that sometimes like that it's not like the structure of a sentence that it's going to be fine, but how to put it in a formal way I'm not repeating myself you know that's it
Interviewer	what do you mean it's not the structure of a sentence
Maria	so like how to put the verb and the you know and the subject and everything you know like the structure I know how to do it I know how to make passive and everything but sometimes just difficult to find like the formal way to do it.
Interviewer	okay so formal What do you mean by formal.
Maria	like how I will write it for a journal I will not I will not speak the same way to you that to you know to like in a congress, you know
	(inaudible cross talk)
Interviewer	okay. So we talked a little bit about when you see a paper that's badly written and it might be a case of translation.
	that obviously suggests that the authors we're talking about aren't English native speakers, right?
	you yourself are not a native speaker of English; have you ever felt particularly disadvantaged with regards to writing for publication by the fact that English is not your first language?
Renzo	well for sure, I mean it took me a long time to get to the level I am now and I think I'm still learning.
	at the beginning especially I had to learn first of all how to ...

yeah, I think compared to a native speaker I had a steeper learning curve to begin with, but then, after that initial struggle, I feel now I'm more or less at the same level as a native speaker when I'm writing for academic purposes.

Interviewer   so, just based from that, have you felt disadvantaged by the fact that you are not a native English speaker? Or..? English is your second language-

Tom   well, yeah, of course! I mean there is a disadvantage, especially if you are don't know English, because there are friends of mine that are Italian, they, they did an amazing job writing.

this guy, this my PI from [name of mid-western city], he's an amazing writer, he wrote a book or two books in English you know related to his job. And, a friend of mine, she's now back in Italy, but when she was in [name of mid-western city], her PI asked her to write grants for her because she was writing an amazing English.

so, you know, I always never understood if it was because, you know, these are guys, these are people that had the experience and the motivation to learn English when they were younger, you know, in high school.

maybe, you know, there's also people that their parents were either professor or doctors, so maybe they've been exposed to that.

Interviewer   have you ever felt disadvantaged with regards to writing for publication by the fact that English is not your first language?

Walter   yeah, I wouldn't feel disadvantaged, but I feel like the English speaker are advantaged. You know, it's a part of our job to learn English and being able to speak English and write in English. Of course, it's easier if it's your native language.

Interviewer   that's interesting.

you don't feel disadvantaged, but you think they are advantaged?

Walter   yeah.

Interviewer   interesting

Interviewer   And so have you ever felt disadvantaged with regards to writing for publication by the fact that English isn't your first or native language.

Emma   um I think not not personally

but I when so when I was reviewing things I think I I felt like I was biased

for example, if I read an article which the English wasn't very good I felt bias from the beginning and which I tried to like fight against.

so I can imagine that if if a reviewer reads reads one of my manuscript in it there's there's some weird German constructions in there and that that might happen as well, but I think.

Well, I hope that most since since there's so many non native researchers is non native writers that that it balances out I guess.

VITA

Frances M. Junnier

Candidate for the Degree of

Doctor of Philosophy

Dissertation: LANGUAGE, IDENTITY, AND VOICE: EUROPEAN PERSPECTIVES ON ENGLISH AS THE INTERNATIONAL LANGUAGE OF SCIENCE

Major Field: English

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in English at Oklahoma State University, Stillwater, Oklahoma in July, 2022.

Completed the requirements for the Master of Arts in Politics and International Relations at University of Aberdeen, Aberdeen, United Kingdom in 2001.

Experience:

Associate Director, Technical Writing Program, Department of English, Oklahoma State University, 2022 – 2022

Graduate Teaching Associate, Department of English, Oklahoma State University, 2017 - 2022