18

DISCOURSE OF ACADEMIA FROM A MULTIDIMENSIONAL PERSPECTIVE

Tony Berber Sardinha

CATHOLIC UNIVERSITY OF SÃO PAULO

Introduction

This chapter seeks to demonstrate how the discourses that circulate in the texts of a scientific discipline, creating the discipline's identity, can be (1) identified using a particular corpus-based approach known as Multi-Dimensional (MD) Analysis, and (2) drawn upon to construct a history of the discipline. MD Analysis is a framework for the analysis of variation in language use introduced by Douglas Biber in the 1980s (Biber, 1988); it has since been widely used for the study of register variation (Berber Sardinha, & Veirano Pinto, 2014, 2019). The MD analytical framework is part of what Biber (2012, 2019) terms the text-linguistic approach to the study of register variation: "the text-linguistic approach to register variation uses quantitative methods to describe the linguistic characteristics of each text, as the basis for comparing the patterns of register variation across texts" (Biber, 2019, p. 43), with registers being "named, culturally-recognized categories of texts" (Biber, 2019, p. 44), such as conversation, research articles, and cooking recipes.

In this chapter, a variant of the MD Analysis framework is employed that draws exclusively on lexical units—namely in the analyses presented here, lemmas, which are the base form of words. Lexical MD Analyses have been conducted for a range of different purposes, including classifying registers through bigrams (Crossley & Louwerse, 2007), producing lists of collocations for EAP teaching (Zuppardi & Berber Sardinha, in press), analyzing register variation (Berber Sardinha, 2017), and describing representations of national identity (Berber Sardinha, 2019). In this chapter, the lexical dimensions comprise sets of lexical items that occur in similar groups of texts at particular time periods, with the lexical MD Analysis serving as a tool for detecting historical disciplinary discourses.

The primary goal of a lexical MD Analysis of discourse is to describe discourse variation; that is, how discourse varies systematically according to context. The variation is modeled from a multidimensional perspective, meaning that each text is seen as being shaped **simultaneously** by the incidence of the various discourses represented by the

various dimensions. As their names suggest, the text-linguistic approach in general and MD Analysis in particular both have the text as the centerpiece of the analytical framework, which means that the corpus must be designed and analyzed with the texts as the primary focal units.

This centrality of the text in corpus design and as a research analytical unit in MD Analysis contrasts with what Biber (2019) calls the "corpus-linguistic approach," in which the analysis is carried out using the whole corpus (or a section of it) as the unit of observation. Unlike in MD Analysis, in the corpus-linguistic approach, "texts are not recognized as relevant constructs" and consequently "there is only one observation per register" (Biber, 2019, p. 54). Similarly, a corpus-linguistic approach to discourse analysis would typically require only one observation per discourse or context and, as a result, the analysis would consider the discourse characteristics that occur in the corpus regardless of text dispersion. For both register analysis and discourse analysis, the problem with this approach is the same: because the variation among the texts with respect to the incidence of particular features is not a primary concern, the characteristics that occur in a small percentage of the texts may end up being generalized to the whole corpus. For example, a non-MD, "corpus-linguistic" investigation of misogynistic discourse would typically include counting the frequency of features associated with this particular discourse across the corpus as a whole, without taking into account the actual distribution of the features across the texts. As a result, if the distribution is skewed in such a way that features under investigation occur in a small subset of the texts with high frequency, the analysis might mistakenly conclude that this discourse is widespread. In contrast, a text-linguistic investigation of this corpus would find that variation exists across the texts with respect to the spread of misogynistic discourse and would look deeper into the high- and low-frequency groups of texts to determine some of the reasons why a discrepancy occurs in the incidence of those discourse features.

As the chapters in this handbook show, the literature presents different understandings of the concept of discourse (Baker & Ellege, 2011; Gee & Handford, 2012). In this chapter, we see discourse as a historically situated, meaning-making social activity that can be described in part through lexical choices. We seek to identify these lexical choices by applying a lexical MD Analysis to a corpus of academic texts to determine the sets of words that co-occur at particular times in the history of applied linguistics. This working definition is based on four interrelated understandings of discourse. The first understanding is as "ways of looking at the world, of constructing objects and concepts in certain ways, of representing reality" (Baker & McEnery, 2015, p. 5). Second, in this study discourse is seen as "the set of meanings, … representations, … statements and so on that in some way together produce a particular version of events" (Burr, 1995, p. 48). Third, it is "an ensemble of ideas, concepts, and categorizations that are produced, reproduced and transformed in a particular set of practices and through which meaning is given to physical and social realities" (Hajer, 1995, p. 44). Finally, it is:

> a group of statements which provide a language for talking about—i.e. a way of representing—a particular kind of knowledge about a topic. When statements about a topic are made within a particular discourse, the discourse makes it possible to construct the topic in a certain way. It also limits the other ways in which the topic can be constructed.
>
> *Hall, 1992, p. 201*

Of particular interest to us here is this notion that how ideas are constructed vary in many different ways, such as according to the general editorial policy of the particular journals in which the articles are published or, over time, reflecting how concepts, theories, methods, and so on are introduced, debated, and reconceptualized historically.

Historical perspectives or core issue and topics

We can identify two major strands in the study of academic discourse from a corpus perspective: those having the overall goal of describing particular linguistic characteristics and those having the overall goal of identifying representations and constructions in academic writing. The studies reported in this chapter fall within the second strand. Studies from both strands have been implemented through different methods. Studies pursuing the first goal are far more numerous and can be split into two subtypes, according to Biber, Connor, and Upton (2007): those that focus on the functions and distribution of surface linguistic features and those that describe patterns of discourse organization. Each of these types can be further categorized as applying either top-down or bottom-up methods; in the first case, the features of interest are defined ahead of time, whereas, in the second case, they emerge from the analysis of the corpus itself.

Studies of individual linguistic features focus on a sample of words or grammatical categories and proceed to count and analyze these features by searching for their occurrence and categorizing their uses across the corpus or within each sub-corpus. An example of a top-down study of features of academic writing is Harwood (2005), which looked at the uses of the pronouns "I" and "we" as self-promotional devices in research articles across four different fields: physics (hard, pure science), economics (soft, pure), computer science (hard, applied), and business and management (soft, applied). By looking at the use of these pronouns in the corpus, the study showed a range of different functions performed by these pronouns, such as engaging in self-promotion via disputation, stressing methodological innovation, and avoiding methodological pitfalls. In contrast, an example of a bottom-up study of individual features is Cortes (2008), which identified and classified the most salient lexical bundles (frequent four-word sequences) in history articles in both English and Spanish (thereby beginning an exception to the norm of English-centric studies of discourse). The categorization of the lexical bundles revealed a number of different structures and functions. For instance, the bundles perform various referential functions, like time-event, place-event, quantifying, and identification, in addition to stance and discourse organization functions. The analysis showed a high degree of similarity between the functions performed by equivalent and semi-equivalent bundles, suggesting that academic discourse cuts across language barriers to form a kind of language-independent mode of communication.

In contrast to analyses of lexical and structural features that operate independently of text-internal units, studies of discourse organization are concerned with discovering, categorizing, and describing the discourse-level units that exist within individual texts. Generally, these units consist of sequences of adjoining sentences, and each sequence performs a different textual or rhetorical function. An example of a top-down analysis of academic discourse organization is Kanoksilapatham (2007), which focused on the identification of the moves (functional units of text organization) in biochemistry research articles; a total of 15 such move types were hand-coded in the corpus. The analysis showed a wide variation in the use of the move types across the article sections, with

moves like "establishing a topic" being present in all introductions, whereas "detailing equipment" occurred in only 10% of the methods section. Finally, an example of a bottom-up analysis of discourse organization is Biber and Jones (2007), which used an automated procedure ("TextTiling"; Hearst, 1997) to segment each text in a corpus of biology research articles into Vocabulary-Based Discourse Units (VBDUs). Each VBDU is a contiguous lexically cohesive portion of text that can perform a variety of discourse functions. An MD Analysis unveiled the four major dimensions, or discourse functions, performed by the VBDUs: (1) evaluation of possible explanations, (2) current state of knowledge versus past events and actions, (3) procedural presentation of actions/events versus elaborated description, and (4) abstract/theoretical discussion of concepts.

Although MD Analysis has been used for the study of discourse organization, its main application has been the study of register variation in contemporary academic writing (e.g., Biber, 2006; Hardy, 2015). More rarely, MD Analysis has been used to describe historical changes in scientific writing. One example is Atkinson (1992), which analyzed the evolution of medical writing in a corpus of articles from the *Edinburgh Medical Journal* dating from 1735 to 1985. The study traced the temporal variation in the corpus using the dimensions of register variation found by Biber (1988). The study detected major shifts in the discourse of medical writing, such as a steady move from "involved" to "informational" discourse, a loss of "narrativization," and a greater reliance on covert persuasion.

These studies of the uses of particular linguistic characteristics in academic writing contrast with the focus of this chapter, which is to explore corpora for forms of "constructing objects and concepts in certain ways" (Baker & McEnery, 2015, p. 5) through a multivariate, multidimensional approach. A study with a similar goal and method to the one reported in this chapter is Fitzsimmons-Doolan (2014), which used factor analysis to identify the ideologies underlying a corpus of language policy texts from the Arizona Education Department. The corpus, which comprised 389 texts totaling *c*. 1.4 million words, was analyzed by retrieving the collocates of the node words "language," "literacy," and "English." The counts of the collocates were entered into a factor analysis, which revealed six factors, each consisting of collocates that tended to co-occur with the same node words in the same texts. The factors were considered ideologies to the extent that they indexed "attitudes toward language both affectively … and cognitively" (Fitzsimmons-Doolan, 2014, p. 65). The six factors were interpreted as the following ideologies: "written language as measurably communicative," "language acquisition as systematically metalinguistic and monolingual," "academic language as standard and informational," "language acquisition as a process of decoding meaning," and "nativeness of language skills as marking group variation."

Another study with a similar goal (but a different method) is Pülzl, Kleinschmitt, and Arts (2014), which examined the historical development of the discourse of bioeconomy. The authors conducted searches on Google Scholar, Scopus, and Web of Science for words such as "global," "forest," and "policy." Based on the documents retrieved, they identified seven major discourses, including three related to global forestry policy (i.e., neoliberalism, civic participation, and global governance discourse) and four related to the environment (i.e., modernity discourse, limits to growth discourse, ecological modernization, and sustainable development discourse). Each of these discourses were found to "shape actors' views, influence their behavior, impact on their beliefs and interests and can cause institutional change in a given society" (Pülzl et al., 2014, p. 386).

Focal analyses

The goal of the focal analyses presented here is to detect some of the major discourses of applied linguistics and, in doing so, map out some of the history of this academic discipline. This is showcased through the analysis of diachronic corpora of applied linguistics, but the underlying principles guiding the analysis can be equally applied to corpora of different kinds (e.g., synchronic corpora), or to corpora of different fields of hard or soft, pure or applied sciences.

Research questions

The research questions examined in this chapter are: (1) What are the major discourses of applied linguistics? (2) How do these discourses shift over time? (3) What historical periods can be discerned based on these discourse shifts? To answer these questions, two corpus-based analyses of large corpora of texts published in flagship journals were carried out. The first analysis consisted of the investigation of a large register-diversified corpus of texts published in the following applied linguistics journals: *Applied Linguistics*, *TESOL Quarterly*, *ELT Journal*, *Language Learning*, and *IRAL* (Berber Sardinha & Reppen, 2019). The second analysis focused on a modified version of the *TESOL Quarterly* sub-corpus from the larger corpus, restricted to the research articles only (Berber Sardinha, 2016). The first corpus seeks to represent the discourses of applied linguistics in general, since its beginning in the 1940s, as disseminated through its leading journals. The second corpus, in contrast, aims to capture a subset of the discourses of applied linguistics, namely the discourse of language teaching and learning. A comparison of the lexical dimensions resulting from these two analyses was carried out, with the goal of determining whether and how discourses migrate between the "mother" field of applied linguistics and its "daughter"/subsidiary field of language teaching.

Corpora

The composition of the two corpora used for this study is shown in Tables 18.1 and 18.2.

Methodology

As mentioned above, the method consisted of a lexical MD Analysis, a derivation of MD Analysis (Biber, 1988) that uses lexical features as data (Berber Sardinha, 2019). Unlike

Table 18.1 Applied linguistics corpus

Journal	*Interval*	*Texts*	*Word count*
Applied Linguistics	1980s–2010s	1,370	7,928,757
ELT Journal	1940s–2010s	4,215	10,184,001
IRAL	1960s–2010s	842	5,489,055
Language Learning	1940s–2010s	1,596	11,128,611
TESOL Quarterly	1960s–2010s	3,423	12,034,356
Total		11,446	46,764,780

Table 18.2 TESOL Quarterly articles corpus

Interval	*Texts*	*Words*
1960s–2010s	1,239	8,251,665

a "traditional" MD Analysis, which uses mostly grammatical features and whose general goal is to identify functional dimensions of register variation, a lexical MD Analysis uses lexical units (words, n-grams, collocations) to identify the dimensions. Briefly, the procedures included tagging the texts for part of speech; lemmatizing the word forms; retrieving all nouns, verbs, and adjectives; counting and selecting the most frequent items of these classes; and norming the frequencies to a rate per 1,000 words. These steps were carried out through computer scripts designed for this study. The resulting lexical counts were entered in a factor analysis (in SAS University Edition) that yielded six factors interpreted qualitatively as dimensions based on their discourse characteristics (for a fuller description of the statistics and computational procedures involved in an MD Analysis see Cantos Gómez (2019), Friginal and Hardy (2014), and Egbert and Staples (2019)). Due to space constraints, only a partial listing of the loadings for the first three factors is provided below. Each text was subsequently scored on each factor by adding up the Z-score of each feature loading on the factor. This provided a measure for marking each text on each dimension. Mean scores were computed for each time period and journal, which were then compared through ANOVAs.

Factor 1

Positive pole: result (.63), score (.59), study (.59), measure (.57), effect (.52), high (.52), finding (.50), proficiency (.49), test (.49), correlation (.48), group (.46), variable (.45), ability (.42), difference (.42), factor (.40), level (.40), performance (.39), variance (.38), sample (.37), compare (.37).

Negative pole: bring (.17), call (.17).

Factor 2

Positive pole: education (.56), educational (.49), school (.45), community (.42), social (.38), cultural (.38), attitude (.35), curriculum (.35), professional (.35), policy (.34), educator (.34), country (.33), experience (.32), economic (.32), challenge (.31), program (.31), culture (.30), academic (.30), minority (.29), world (.28).

Negative pole: sentence (.40), verb (.37), syntactic (.32), structure (.32), grammatical (.31), type (.31), tense (.28), simple (.28), meaning (.27), word (.26), example (.26), phrase (.25), object (.25), noun (.25), semantic (.25), lexical (.24), correct (.24), present (.23), past (.22), construction (.21).

Factor 3

Positive pole: focus (.47), classroom (.46), interaction (.45), activity (.39), learner (.37), opportunity (.34), interactional (.33), talk (.32), role (.32), context (.31), participant

(.31), perspective (.29), feedback (.28), research (.28), practice (.27), understanding (.27), explicit (.27), discourse (.26), turn (.25), outcome (.24).

Negative pole: consonant (.40), vowel (.39), sound (.38), phoneme (.35), phonetic (.35), phonemic (.33), English (.31), final (.29), pronunciation (.28), syllable (.28), stop (.25), foreign (.25), phonology (.24), difficulty (.23), stress (.23), initial (.22), tongue (.22), position (.21), phonological (.20), native (.20).

Interpreting the dimensions as discourses is subjective, because of the shifting nature of discourse itself. As Baker (2006, p. 4) points out, "where I see a discourse, you see a different discourse, or no discourse." The interpretation involves selecting a short, descriptive title that captures the essence of the dimension. The process of labeling dimensions in MD Analysis is notoriously difficult because of the issues involved in packaging a large set of features into a short descriptive phrase (Friginal & Hardy, 2019). Arguably this process becomes amplified with the analysis of discourse, because "discourses can … be difficult to pin down or describe—they are constantly changing, interacting with each other, breaking off and merging" (Baker, 2006, p. 4). In addition, there is no one-to-one relationship between a dimension and a discourse; rather, a single dimension can comprise several simultaneous discourses, and similar discourses can crop up in different dimensions.

Findings

Analysis of the applied linguistics corpus

The first factor in the MD Analysis of the applied linguistics corpus captures a quantitative, empirical research dimension that came to the forefront of the field in the 1990s (Figure 18.1), predominantly in journals such as *Language Learning* and *Applied Linguistics*. This points to a representation of "applied linguistics as an empirical/physical/natural science"—a field that depends on statistics, mathematics, empirical verification, and control-group experiments for its rigor, validity, and reliability. An example is Vercellotti's (2017) "The development of complexity, accuracy, and fluency in second language performance: A longitudinal study."

Sample of Dimension 1 (quantitative, empirical research)

Using these quantitative analyses, **high fluency**, and/or **high** lexical variety did not negatively impact **accuracy**. These trajectory patterns, therefore, indicate that **individual differences** in **performance** are more likely a **result** of developmental lag, not developmental deficit, at this stage of SLD.

The second factor reflects a distinction between lexis used to talk about social, cultural, literacy, and professional concerns in a broader context of education (on the positive pole) and lexis largely related to linguistic theory and analysis (on the negative pole). These two lexical sets occur in two separate time periods: linguistic analysis and theory in the early period (1940s through 1980s) and educational concerns in the more recent period (from the 1990s on; see Figure 18.2). This opposition points to a duality between the discourse of what we could call "the discourse of educational practice," on the one hand, and "the discourse of linguistic structure and theory" on the other, or "education" versus "linguistics" for short. This opposition reflects a duality existing in the field: one that sees applied

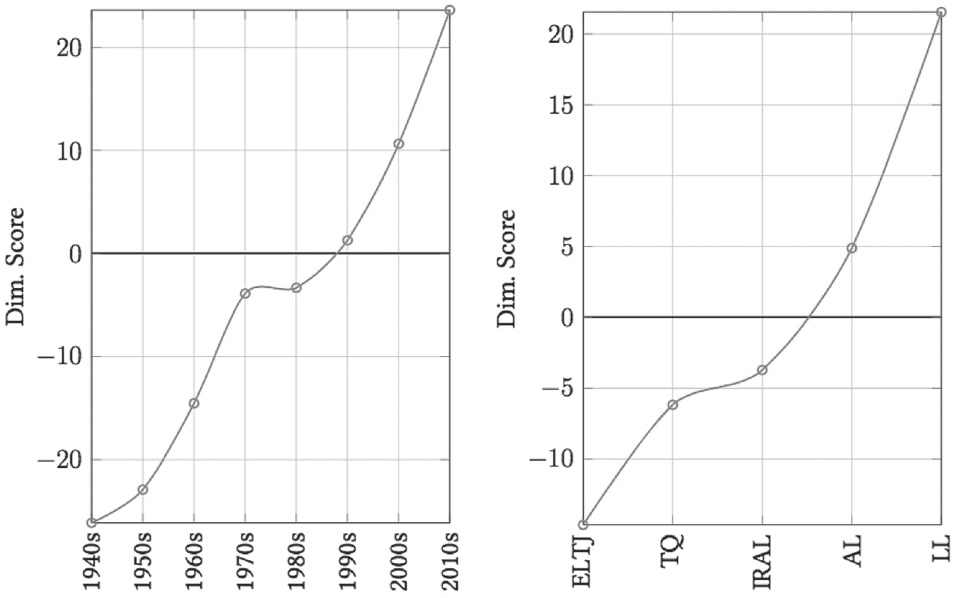

Figure 18.1 Mean Dimension 1 scores for decade and journal

linguistics as not being restricted to language teaching but also extending to education as a whole, and another that sees applied linguistics as a daughter discipline of linguistics. This distinction is mapped onto the journals marked for each pole. The discourse of "applied linguistics as education" essentially appears in *TESOL Quarterly*, *Applied Linguistics*, and *ELT Journal*, whereas the discourse of linguistic theory occurs more frequently in *IRAL* and *Language Learning* (Figure 18.2). An example of education-centered discourse is Porto's (2010) "Culturally responsive L2 education: An awareness-raising proposal," whereas an example of structural linguistic theory discourse is Kleinjans's (1958) "A comparison of Japanese and English object structures."

Sample of Dimension 2 (education)

ELT is seen as including much more than purely linguistic aspects, as it focuses also on broad **literacy issues** which acknowledge the **importance** of global **economic**, **social**, historical, and **cultural** factors in language learning and teaching. In other words, ELT in the twenty-first century means **culturally** responsive **literacy education**.

Sample of Dimension 2 (linguistic theory)

These **constructions** make the distribution of the Japanese indirect **object structure** cover less area than that of the indirect **object construction** in English. Although this should not cause the Japanese to make mistakes in English, it does indicate that they will have to concentrate harder on the **word order patterns** …

The third factor shows a further distinction between the two lexical sets: one that refers to studies largely looking at talk, classroom interaction, and discourse (positive pole) and

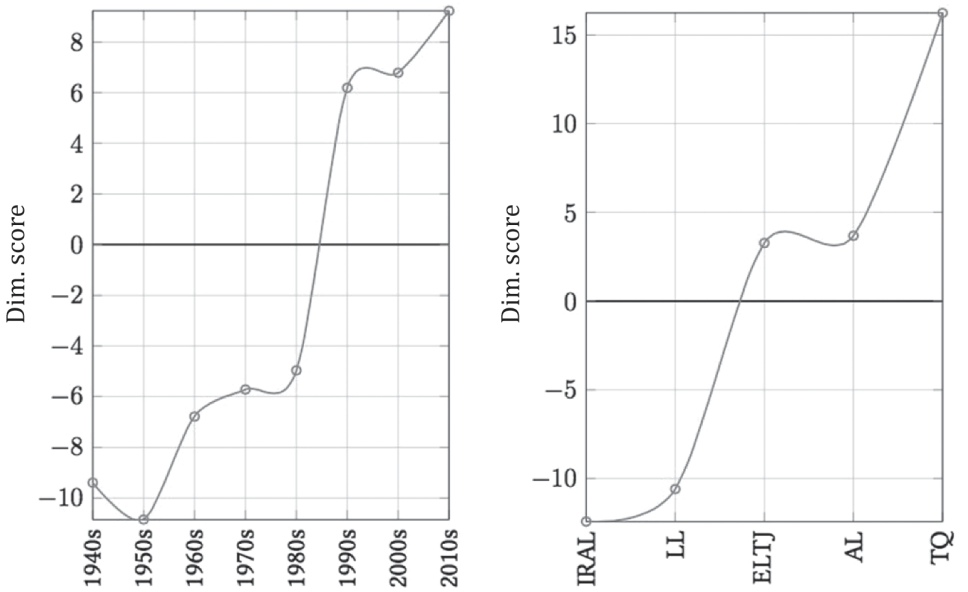

Figure 18.2 Mean Dimension 2 scores for decade and journal

one that reflects an interest in the description and theory of the sound system (phonology, phonetics, acoustics, etc.; negative pole). These sets seem to indicate two different discourses around speech. The first one sees applied linguistics as related to people using speech to communicate and interact ("speech as interaction"), whereas the other one views applied linguistics as related to describing and teaching speech at the level of pronunciation ("speech as pronunciation"). These two discourses occupy different time periods (Figure 18.3). The discourse related to pronunciation is typical of the early period (between the 1940s through the 1980s); the discourse on communication and interaction is found in the later period (from the 1990s on). Again, a marked difference exists between the journals, with *TESOL Quarterly*, *Applied Linguistics*, and *ELT Journal* being more frequently associated with interaction while *IRAL* and *Language* being more associated with pronunciation. Examples of typical texts for both poles (below) come from Al-Gahtani and Roever's (2012) "Proficiency and sequential organization of L2 requests" for interaction and Tibbits's (1947) "Pronunciation difficulties" for pronunciation.

Sample of Dimension 3, positive pole (interaction)

Bardovi-Harlig and Salsbury (2004) stimulated **conversations** through emotion cards (Rintell, 1989) or preset **topics**, recorded the ensuing **interaction**, and then identified and classified cases of oppositional **talk**. Elicited **discourse allows** more **researcher** control than **authentic** data but is limited to eliciting casual **conversation**.

Sample of Dimension 3, negative pole (pronunciation)

The following list of words is divided into groups within which the two **vowels** occur in a similar **phonetic** context, that is, surrounded by similar **sounds**, and, since the words are

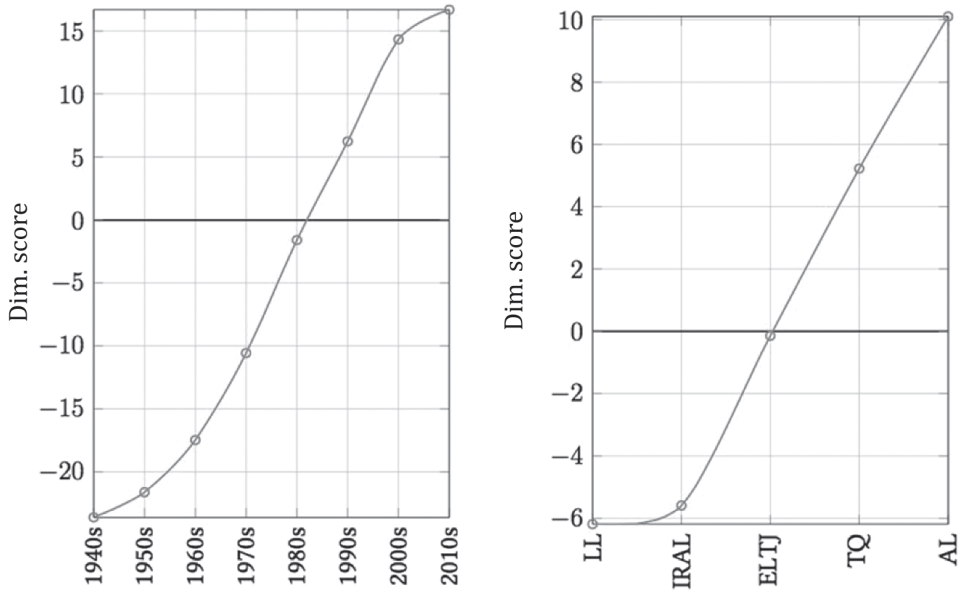

Figure 18.3 Mean Dimension 3 scores for decade and journal

obviously intended to be pronounced in isolation, with similar **stress**. All the relevant **vowels** in group a (followed by a **final** voiced **consonant**) are relatively somewhat **longer** than those in group b (followed by a weakly **stressed syllable**).

The amount of variation occurring among the different journals and time frames was measured using ANOVAs. As these ANOVAs returned significant results, the differences among the journals on the one hand and among the decades on the other are significant. In addition, a procedure known as the coefficient of determination (indicated by R^2) tested the dimensions for their predictive power by showing the percentage of variation that can be predicted by knowing the levels of the variables (either the individual journals or time periods, or both). The individual results are shown in Table 18.3. Overall, the interaction between journal and time frame is a better predictor of the variation than either journal or time alone, suggesting that journal and time of publication act together to influence how the discourses change over time. For instance, more than 40% of the variation for Dimension 1 can be predicted by knowing both the journal and the time period, as opposed to 17% for the journal alone and 15% for time alone. For Dimensions 1 and 2, the journal in which the discourse took place is a more important predictor than the time period, but for Dimension 3 it is the opposite—that is, the time period is more important. This suggests that both the "empirical science" (Dim. 1) and the "education versus linguistics" (Dim. 2) discourses are both products of the agendas of particular journals, whereas the "speech as interaction versus speech as pronunciation" dimension (Dim. 3) is a more consensual discourse.

The results show clear historical shifts for all three dimensions, yet each dimension portrays an individual variation pattern, so it is hard to see how the dimensions interact with each other. To display this interaction among the discourses, it is necessary to combine the individual variation patterns, which was achieved using a cluster analysis, with each cluster comprising together texts with similar dimensional profiles. By aligning these

Table 18.3 Results of ANOVAs for the first three dimensions

Variable	*Dim.*	*DF*	*F*	*p*	R^2	*%*
Journal	1	4	267.98	<.0001	.173143	17.31
	2	4	304.97	<.0001	.192442	19.24
	3	4	95.56	<.0001	.069482	6.95
Decade	1	7	128.29	<.0001	.149321	14.93
	2	7	69.54	<.0001	.086878	8.69
	3	7	385.82	<.0001	.345509	34.55
Both journal and time period	1	31	112.48	<.0001	.406444	40.64
	2	31	74.47	<.0001	.311957	31.20
	3	31	106.59	<.0001	.393538	39.35

clusters in a time series, based on the time period in which the majority of the texts were published, a timeline of the discursive eras of the discipline was derived.

A hierarchical cluster analysis was conducted using the factor scores for the six dimensions as the data. The number of clusters in the data was determined by means of procedures such as Cubic Clustering Procedure, Pseudo-F, and Pseudo-t^2, which suggested both a two- and a six-cluster solution as possible. Only the former is discussed here, for reasons of space. This two-cluster solution yields a split between two distinct time periods: from 1946 through the late 1980s, and from the late 1980s to 2015. The first period, comprising about 40 years, is based on negative scores of dimensions 1, 2, and 3, indicating a confluence of the discourses of structural and phonetic linguistic theory in addition to an incipience of the discourse of quantitative research. In other words, it is a period marked by the discourse of applied linguistics as an offshoot of linguistics. In contrast, the time period that followed, covering the last 25 to 30 years, is marked by the confluence of the discourses of scientific rigor, education, and interaction. As can be seen, these are sharply distinct eras, each with a very different view of applied linguistics.

Analysis of the TESOL Quarterly *article corpus*

As mentioned above, the *TESOL Quarterly* corpus comprises articles published in the journal from its inception in 1967 until 2016. Five factors were extracted through a Promax-rotated factor analysis. The resulting factorial pattern appears below (for reasons of space, only a subset of the words in each factor are shown).

Factor 1

Positive pole: power (.55), identity (.45), perspective (.45), social (.44), world (.41), understanding (.41), practice (.40), people (.40), cultural (.40), issue (.39), pedagogy (.39), society (.38), critical (.38), notion (.37), life (.36), argue (.36), community (.35), discourse (.35), challenge (.34), context (.33).

Negative pole: score (.55), proficiency (.48), test (.48), measure (.46), correlation (.45), high (.45), total (.45), level (.44), significant (.43), ability (.40), sample, (.39), item (.37), determine (.35), variable (.35), grade (.34), design (.32), performance (.32), comparison (.31), procedure (.31).

Factor 2

Positive pole: form (.51), acquisition (.45), occur (.42), grammatical (.40), produce (.40), structure (.40), grammar (.35), meaning (.35), production (.33), attention, (.33), explicit (.32), utterance (.32), sequence (.32), rule (.31), present, (.31), noun (.30), simple (.30), function (.29).

Negative pole: school (.47), education (.47), educational (.43), English (.39), program (.37), country (.35), university (.34), survey (.34), academic (.30), international, (.30), policy (.27), home (.26), graduate (.26), bilingual (.25), experience, (.24), project (.23), member (.23), educator (.22).

Factor 3

Positive pole: study (.62), finding (.52), research (.45), effect (.44), examine (.44), researcher (.43), participant (.42), learner (.39), multiple (.37), datum, (.35), focus (.34), task (.34), address (.33), construct (.31), influence, (.29), interaction (.28), knowledge (.25), positive (.23).

Negative pole: material (.42), teach (.40), sentence (.40), exercise (.39), basic (.36), situation (.35), fact (.34), deal (.32), certain (.32), point (.31), learn, (.31), paper (.31), problem (.28), statement (.27), hear (.26), drill (.26), technique (.25), step (.25), necessary (.24).

Factor 4

Positive pole: teaching (.51), teacher (.48), classroom (.43), need (.41), curriculum (.39), activity (.36), communicative (.34), goal (.34), instruction (.32), course, (.32), instructional (.31), learning (.31), lesson (.30), observation (.29), area (.28), skill (.28), class (.26), method (.26), concern (.25).

Negative pole: word (.29), native (.24), frequency (.20).

Factor 5

Positive pole: text (.58), writing (.58), write (.55), reader (.53), writer (.48), idea (.42), written (.41), essay (.39), read (.38), reading (.38), topic (.31), process, (.31), passage (.31), rhetorical (.30), composition (.30), comprehension (.29), paragraph (.29), main (.27), organization (.27).

Negative pole: speaker (.34), speech (.32), sound (.27), speak (.27), linguistic (.26), environment (.24), dialect (.23), pronunciation (.23), perception (.21).

The interpretation of these factors gave rise to the dimensions discussed here. The positive pole of Dimension 1 reflects a discourse of language teaching and learning as being literacy-driven, critical, (socio)cultural, and discursive, whereas the discourse embedded in the negative pole highlights the view of language teaching and learning as geared toward attaining proficiency. These provide two sharply distinct views of the goals of language teaching and learning: language teaching and learning as essentially social (positive pole) and language teaching as essentially individual (negative pole). These two poles represent two distinct time periods: an early time period from the 1960s to the 1980s (negative

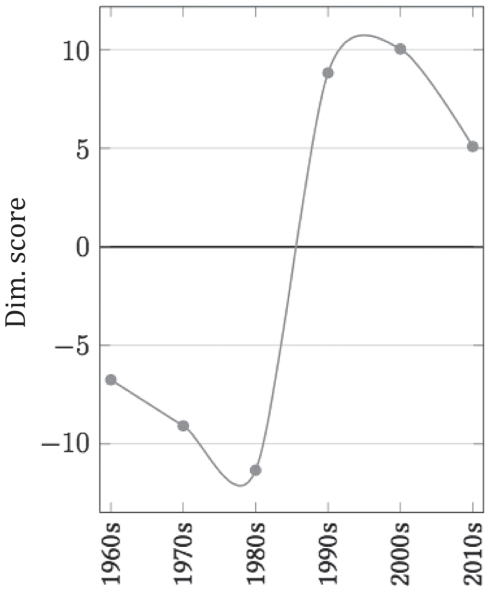

Figure 18.4 Mean dimension scores for *TESOL Quarterly* Dimension 1

pole) and a more recent time period that was initiated around the 1990s (positive pole) (Figure 18.4). A typical example of the negative pole is Alderson's (1979) "The cloze procedure and proficiency in English as a foreign language"; a typical example of the positive pole is Kubota's (1999) "Japanese culture constructed by discourses: Implications for applied linguistics research and ELT."

Sample of Dimension 1, positive pole (literacy, critical, social, cultural concerns)

In this **perspective** of **critical literacy**, pluralism is sought not by neglecting to teach the dominant language and **culture** but by adding a **discourse** of **power** to the repertoire that students bring to the mainstream **society**. Overall, respecting and preserving **cultural** and linguistic codes that are different from one's own are necessary to create equality in **society**.

Sample of Dimension 1, negative pole (proficiency)

There is clear interaction between deletion **rate** and text which makes it impossible to **generalize**. Nevertheless, it is clear that different texts, using the same deletion **rate**, result in different **correlations** with the **criterion**, which suggests that different texts may well **measure** different aspects of EFL **proficiency**, or the same aspect more efficiently or less efficiently.

Dimension 2 also includes an opposition between two very different discourses. On the positive pole, language teaching and learning relate to applying linguistic theory, which is typical of the 1970s and 1980s (Figure 18.5). On the negative pole, the view of language teaching and learning is part of a wider framework that includes multilingualism

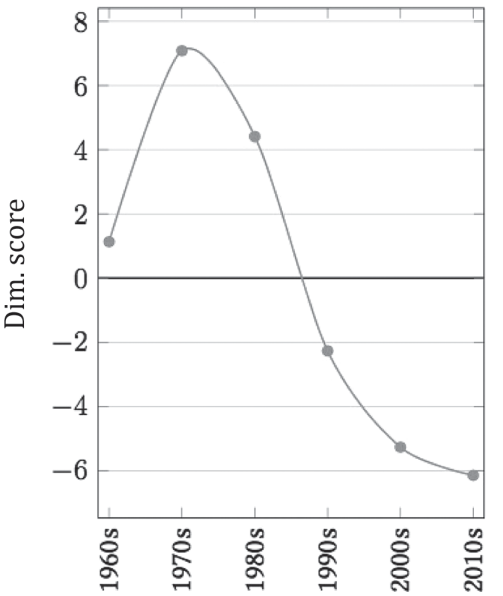

Figure 18.5 Mean dimension scores for *TESOL Quarterly* Dimension 2

and large-scale policies in the educational system, which is often found from the 1990s on. Examples of these are DeCarrico's (1986) "Tense, aspect, and time in the English modality system" (positive pole) and Hornberger and Johnson's (2007) "Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice" (negative pole).

Sample of Dimension 2, positive pole (linguistic theory)

With the modal **simple past** in the main **clause**, the **past** perfect **occurs** in the other **clause**, as usual. It should be noted that in certain contexts, **present** perfect **meaning** can be pragmatically forced, as in "He should have been here by now."

Sample of Dimension 2, negative pole (education)

We show that this institution in turn opens up additional ideological and implementational spaces for indigenous rights and indigenous **education**, surpassing those initially envisioned even in multilingual national **policies**. The PROEIB Maestría is ethnically and linguistically diverse, enrolling indigenous **educators** through a selection process in each **country** involving their respective ministries of **education**, sponsoring **universities**, and indigenous organizations.

Dimension 3 incorporates a distinction between the discourse of research on language teaching and learning as empirical science (on the positive pole) and the discourse of language teaching and learning as materials and techniques (on the negative pole). These occupy different time periods. The discourse of materials and techniques is typical of the early phase of the journal, especially the 1960s and 1970s, whereas the discourse of

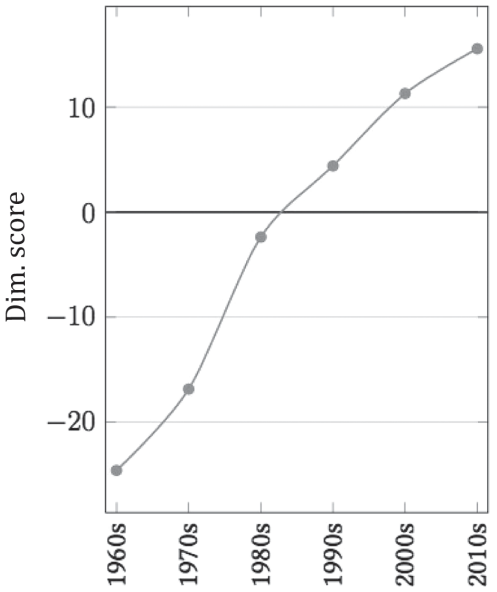

Figure 18.6 Mean dimension scores for *TESOL Quarterly* Dimension 3

research rigor is typical of the later phase (after the 1980s) (Figure 18.6). Articles that illustrate these discourses include Reigel's (2008) "Positive feedback in pairwork and its association with ESL course level promotion" (positive pole) and Paulston's (1971) "The sequencing of structural pattern drills" (negative pole).

Sample of Dimension 3, positive pole (empirical science)

Ferguson and Houghton (1992) found that praise **positively affected** student classroom behavior, but as the frequency of praise decreased so did the desired on-task behaviors. Students are alert to rote **positive** feedback. During the **data** collection, the **researcher** observed students giving each other mock praise.

Sample of Dimension 3, negative pole (materials and techniques)

The **learning** process varies according to the structural **pattern drilled**. At this **point**, however, there is still no real communication taking place. Students have a tendency to **learn** what they are **taught** rather than what we think we are **teaching**. If we want fluency in expressing their own opinions, then we have to **teach** that.

Dimension 4 corresponds to the discourse of language teaching and learning that could be broadly described as centering on the role of teachers, which encompasses different issues such as teacher training, planning and sequencing, method and curriculum development, and course design (the negative pole is not interpreted because it includes very few words). This discourse was particularly prevalent in the 1980s and 1990s (Figure 18.7), an example of which is Grosse's (1991) "The TESOL methods course."

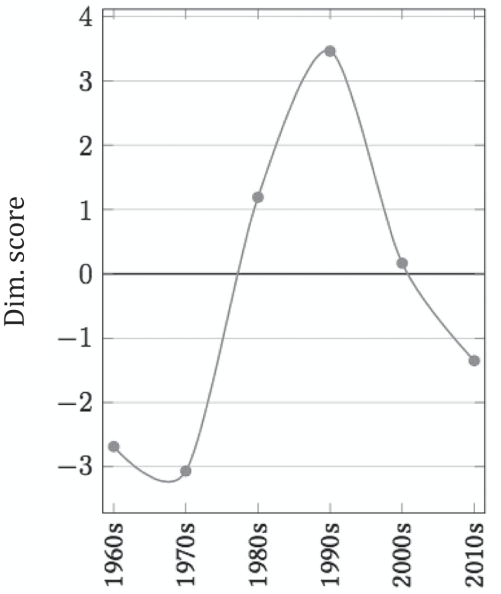

Figure 18.7 Mean dimension scores for *TESOL Quarterly* Dimension 4

Sample of Dimension 4, positive pole (teachers, planning, and sequencing)

A third key issue in the **methods course** of the 1990s will be the **development** of metacognitive awareness in effective **teaching**. Freeman (1989) clearly describes the "trigger" effect of awareness on the three bases of **teaching**, namely knowledge, **skills**, and attitudes, calling it a vital aspect the **development** of **teachers** and their "internal monitoring systems" (p. 40).

Finally, Dimension 5 is based on the discourse of language teaching and learning as skills development, which defines two sets of skills: reading and writing on the positive pole and listening and speaking on the negative pole. The former is most active in the 1980s and 1990s, whereas the latter is more frequent in the 1960s and 1970s as well as later in the 2000s and 2010s (Figure 18.8). Typical examples are Weissberg's (1984) "Given and new: Paragraph development models from scientific English" (positive pole) and Jacobson's (1970) "The teaching of English to speakers of other languages and/or dialects: An oversimplification" (negative pole).

Sample of Dimension 5, positive pole (skills: reading and writing)

The given/new contract, with its associated patterns of **topic** development, offers the ESL **composition** teacher a rational alternative to the traditional **rhetorical** models used in teaching **paragraph** construction. It has been shown here to be especially valid for use with students in scientific and technical **writing** courses.

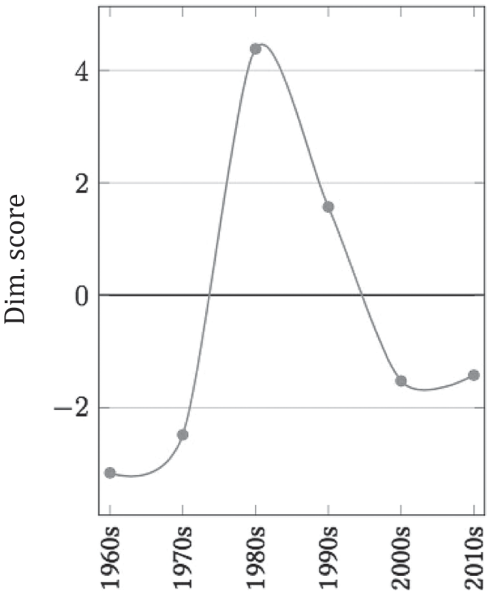

Figure 18.8 Mean dimension scores for *TESOL Quarterly* Dimension 5

Sample of Dimension 5, negative pole (skills: listening and speaking)

The performance in Standard English by the **speaker** of another language will result from the learner's acquisition of deep and surface structure rules as well as of an appropriate lexicon, whereas the performance in Standard English by the **speaker** of a non-standard **dialect** requires only the adjustment of a few deep structure and of a larger number of surface structure roles.

Comparison of dimensions

This part of the analysis compares the dimensions found in the description of the larger applied linguistics corpus and the more specialized *TESOL Quarterly* articles corpus. The comparison shows that, although differences exist between the two sets of dimensions, all applied linguistics dimensions match at least one *TESOL Quarterly* dimension (Table 18.4). However, the opposite is not true, as materials and techniques and reading and writing did not surface directly as dimensions in their own right in the analysis of applied linguistics.

Conclusion

This chapter has focused on showing how corpora can be explored for the identification of the discourses of academic/scientific fields. By using case studies of diachronic corpora, it was possible to show how one particular applied human science (applied linguistics) has been shaped historically by a range of discourses. The results show that the history of the field is marked by a series of discourses that unfold over time, forming complex patterns of variation. The analyses also suggest that journals are important agents

Table 18.4 Comparison of discourses in applied linguistics and *TESOL Quarterly* discourses

Applied linguistics discourses		*TESOL Quarterly discourses*	
Dim 1 pos.	Empirical science	Dim 1 neg.	Proficiency
		Dim 3 pos.	Empirical science
Dim 2 pos.	Education	Dim 1 pos.	Literacy, critical, social, cultural
		Dim 2 neg.	Education
Dim 2 neg.	Linguistic theory	Dim 2 pos.	Linguistic theory
Dim 3 pos.	Interaction	Dim 4 pos.	Planning and sequencing
		Dim 1 pos.	Literacy, critical, social, cultural
Dim 3 neg.	Speech	Dim 5 neg.	Skills: Listening and speaking

of change as particular journals introduce, support, and disrupt particular discourses at particular times.

The existence of a range of different discourses in applied linguistics is evidence that it is a diversified, multidisciplinary field in which different knowledge bases and multiple disciplinary concerns coexist under a single umbrella. The historical shifts revealed in the analyses indicate that the field has changed quite considerably from its beginnings more than 70 years ago, moving from a discipline that drew mostly on language teaching and linguistics to a discipline that incorporates knowledge from multiple sources to tackle a wide range of research issues.

From a methodological standpoint, the chapter has highlighted the importance of using multivariate statistical methods for spotting and tracking discourses in a diachronic corpus. Corpus methods that rely on the analysis of single variables, or even on many different variables one at a time, are not appropriate for detecting complex patterns of discourse variation. Because MD Analysis was originally designed to account for variation patterns across registers, it is well suited for spotting discourse variation patterns. A diachronic discourse-based MD Analysis can show the analyst when discourses are present and absent, which discourses are mutually exclusive, and how these discourses take turns through time, which are powerful tools for the discourse analysis of science.

Further reading

Atkinson, D. (1996). The Philosophical Transactions of the Royal Society of London, 1675–1975: A sociohistorical discourse analysis. *Language in Society, 25*(3), 333–371.

Atkinson presents a cross-time study of a corpus of texts published from 1675 to 1975 in the world's first scientific journal in English, *The Philosophical Transactions of the Royal Society of London*. The study provides findings from both a rhetorical and an MD Analysis of the corpus. The MD Analysis confirms the results of the analysis of the *Edinburgh Medical Journal* discussed above (Atkinson, 1992). The rhetorical analysis identifies the major textual developments, such as (1) the visibility of the author being greatly reduced, serving an "object-centered orientation," which resulted in the avoidance of first-person statements; (2) the letter losing preference as the register of excellence for scientific writing, being replaced by the experimental article, whose structure underwent several changes, leading to the introduction–methods–results–discussion (IMRD) structure becoming conventionalized; and (3) a community of research scientists being gradually built and recognized through, for example, detailed literature reviews. The author discusses how these findings fit into the broader context of British society in the seventeenth century, especially the "genteel form of life," a social formation that helps explain the origins of the discourse of *Philosophical Transactions* in particular and academic writing in the English-speaking world in general.

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story.* Amsterdam: John Benjamins.

Gray provides a detailed MD Analysis of current-day academic articles of three kinds (theoretical, qualitative, and quantitative) in six disciplines: philosophy, history, political science, applied linguistics, biology, and physics. The MD Analysis revealed four dimensions: academic involvement and elaboration versus information density; contextualized narration versus procedural description; human focus versus non-human focus; and academese. Broadly speaking, each dimension reveals a contrast among groups of disciplines. The first dimension distinguished philosophy, which came up as "extremely involved," from the other disciplines, especially biology, which was heavily marked for information density. The second dimension provided a contrast between biology and physics, marked for "procedural description," and all the other disciplines, which were marked to some degree for "contextualized narrative description." The third dimension clustered applied linguistics and philosophy as having a "human focus" and the remaining disciplines as having a "non-human" focus. Finally, the fourth dimension found that only political science and applied linguistics had some degree of "academese." This analysis provides a nuanced view of contemporary scientific writing that contrasts with the trends found for medical writing by Atkinson (1992, 1996), such as involved discourse still being accepted in the human and social sciences but shunned in the natural and hard sciences in favor of information-based discourse. Furthermore, it shows that narrative discourse continues to be produced in the human and social sciences.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Amsterdam/Philadelphia, PA: John Benjamins.

Biber reports a comprehensive MD Analysis of a corpus of both spoken and written academic English that comprises six areas: business, education, engineering, humanities, natural science, and social science. Four dimensions were identified: oral versus literate discourse; procedural versus content-focused discourse; reconstructed account of events; and teacher-centered stance. Two of these were associated with significant differences among the disciplinary spoken and written registers. The second dimension (i.e., procedural versus content-focused discourse) distinguished between spoken classroom teaching in engineering, business, and education and all the other disciplines and registers (especially natural science textbooks), with the former employing procedural discourse and the latter, content-focused discourse. This contrasts with Gray's (2015) finding that procedural discourse was a feature of written biology and physics. The third dimension (i.e., reconstructed account of events) also pointed out disciplinary differences: Classroom teaching in education, humanities, social sciences, and business, in addition to textbooks in education and humanities, all activate narrative discourse, unlike the other registers and disciplines—a finding confirmed by Gray (2015) and Atkinson (1992, 1996).

Bibliography

Alderson, J.C. (1979). The cloze procedure and proficiency in English as a foreign language. TESOL Quarterly, 13(2), 219–227. doi:10.2307/3586211

Al-Gahtani, S., & Carsten Roever, C. (2012). Proficiency and sequential organization of L2 requests. *Applied Linguistics*, 33(1), 42–65. https://doi.org/10.1093/applin/amr031

Atkinson, D. (1992). The evolution of medical research writing from 1735 to 1985: The case of the "Edinburgh Medical Journal". *Applied Linguistics, 13*, 337–374.

Atkinson, D. (1996). The Philosophical Transactions of the Royal Society of London, 1675–1975: A sociohistorical discourse analysis. *Language in Society, 25*(3), 333–371.

Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.

Baker, P., & Ellege, S. (2011). *Key terms in discourse analysis*. London: Continuum.

Baker, P., & McEnery, T. (2015). Introduction. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating discourse and corpora* (pp. 1–20). Basingstoke: Palgrave Macmillan.

Bardovi-Harlig, K., & Salsbury, T. (2004). The organization of turns in the disagreement of L2 learners: A longitudinal perspective. in D. Boxer & A. Cohen (Eds.), *Studying speaking to inform second language acquisition*. *Multilingual Matters*, 199–227.

Berber Sardinha, T. (2016). *Corpus linguistics and history: Lexical dimensions in TESOL Quarterly*. Paper presented at the American Association for Corpus Linguistics, Ames, IA, USA.

Berber Sardinha, T. (2017). Lexical priming and register variation. In M. Pace-Sigge & K. Patterson (Eds.), *Lexical priming: Applications and advances* (pp. 190–230). Amsterdam: John Benjamins.

Berber Sardinha, T. (2019). Using multi-dimensional analysis to detect representations of national culture. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 231–258). London/New York: Bloomsbury/Continuum.

Berber Sardinha, T., & Reppen, R. (2019). *A corpus-based history of applied linguistics*. Unpublished manuscript.

Berber Sardinha, T., & Veirano Pinto, M. (Eds.). (2014). *Multi-dimensional analysis, 25 years on: A tribute to Douglas Biber*. Amsterdam/Philadelphia, PA: John Benjamins.

Berber Sardinha, T., & Veirano Pinto, M. (2019). *Multi-dimensional analysis: Research methods and current issues*. London: Bloomsbury Academic.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam/Philadelphia, PA: John Benjamins.

Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory, 8*(1), 9–37.

Biber, D. (2019). Text-linguistic approaches to register variation. *Register Studies, 1*(1), 42–75.

Biber, D., & Jones, J. K. (2007). Vocabulary-based discourse units in biology research articles. In D. Biber, U. Connor, & T.A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 175–212). Amsterdam/Philadelphia, PA: John Benjamins.

Biber, D., Connor, U., & Upton, T.A. (2007). Discourse analysis and corpus linguistics. In D. Biber, U. Connor, & T.A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 1–21). Amsterdam/Philadelphia, PA: John Benjamins.

Burr, V. (1995). *An introduction to social constructionism*. London: Routledge.

Cantos Gómez, P. (2019). Multivariate statistics commonly used in multi-dimensional analysis. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 97–124). London/New York: Bloomsbury/Continuum.

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora, 3*(1), 43–57.

Crossley, S., & Louwerse, M.M. (2007). Multi-dimensional register classification using bi-grams. *International Journal of Corpus Linguistics, 12*(4), 453–478.

DeCarrico, J.S. (1986). Tense, aspect, and time in the English modality system. *TESOL Quarterly, 20*(4), 665–682. doi:10.2307/3586517

Egbert, J., & Staples, S. (2019). Doing multi-dimensional analysis in SPSS, SAS, and R. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 125–144). London/New York: Bloomsbury/Continuum.

Ferguson, E., & Houghton, S. (1992). The effects of contingent teacher praise, as specified by Canter's Assertive Discipline Programme, on children's on-task behaviour. *Educational Studies,* 18(1), 83–93. https://doi.org/10.1080/0305569920180108

Fitzsimmons-Doolan, S. (2014). Using lexical variables to identify language ideologies in a policy corpus. *Corpora, 9*(1), 57–82.

Freeman, D. (1989). Teacher training, development, and decision-making: A model of teaching and related strategies for language teacher education, *TESOL Quarterly*, *23*(1), 27–45. doi: 10.2307/3587506

Friginal, E., & Hardy, J. A. (2014). Conducting multi-dimensional analysis using SPSS. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis, 25 years on: A tribute to Douglas Biber* (pp. 298–316). Amsterdam/Philadelphia, PA: John Benjamins.

Friginal, E., & Hardy, J. (2019). From factors to dimensions: Interpreting linguistic co-occurrence patterns. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 145–164). London/New York: Bloomsbury/Continuum.

Gee, J.P., & Handford, M. (Eds.). (2012). *The Routledge handbook of discourse analysis*. London/New York: Routledge.

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. Amsterdam: John Benjamins.

Grosse, C.U. (1991). The TESOL methods course. *TESOL Quarterly*, 25(1), 29–49. doi:10.2307/3587027

Hajer, M. (1995). *The politics of environmental discourse, ecological modernization and the policy process*. Oxford: Clarendon.

Hall, S. (1992). *Formations of modernity*. Cambridge: Polity Press.

Hardy, J. (2015). Multi-dimensional analysis of academic discourse. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating discourse and corpora* (pp. 155–174). New York: Palgrave Macmillan.

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*(1207–1231).

Hearst, M. (1997). TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics, 23*(1), 33–64.

Hornberger, N.H., & Johnson, D.C. (2007). Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice. *TESOL Quarterly, 41*(3), 509–532. doi:10.1002/j.1545- 7249.2007.tb00083.x

Jacobson, R. (1970). The teaching of English to speakers of other languages and/or dialects: An oversimplification. *TESOL Quarterly*, *4*(3), 41–253. doi:10.2307/3585725

Kanoksilapatham, B. (2007). Rhetorical moves in biochemistry research articles. In D. Biber, U. Connor, & T.A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 73–120). Amsterdam/Philadelphia, PA: John Benjamins.

Kleinjans, E. (1958). A comparison of Japanese and English object structures. *Language Learning*, *8*(1–2), 47-52. https://doi.org/10.1111/j.1467-1770.1958.tb01216.x

Kubota, R. (1999). Japanese culture constructed by discourses: Implications for applied linguistics research and ELT. *TESOL Quarterly*, *33*(1), 9–35. doi:10.2307/3588189

Paulston, C.B. (1971). The sequencing of structural pattern drills. *TESOL Quarterly*, *5*(3), 197–208. doi:10.2307/3585692

Porto, M. (2019). Culturally responsive L2 education: An awareness-raising proposal. *ELT Journal*, 64(1), 45–53. https://doi.org/10.1093/elt/ccp021

Pülzl, H., Kleinschmit, D., & Arts, B. (2014). Bioeconomy—an emerging meta-discourse affecting forest discourses? *Scandinavian Journal of Forest Research, 29*(4), 386–393.

Reigel, D. (2008). Positive feedback in pairwork and its association with ESL course level promotion. *TESOL Quarterly*, *42*(1), 9–98. doi:10.1002/j.1545-7249.2008.tb00208.x

Rintell, E. (1989). That reminds me of a story: The use of language to express emotion by second-language learners and native speakers. In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp. 237–257). New York: Plenum Press.

Tibbitts, E.L. (1947). Pronunciation difficulties: Corrective treatment: II. The I.P.A. broad transcription. *ELT Journal, 1*(3), 78–79. https://doi.org/10.1093/elt/1.3.78

Vercellotti, M.L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics*, *38*(1), 90–111. https://doi.org/10.1093/applin/amv002

Weissberg, R.C. (1984). Given and new: Paragraph development models from scientific English. *TESOL Quarterly*, *18*(3), 485–500. doi:10.2307/3586716

Zuppardi, M.C., & Berber Sardinha, T. (in press). A multi-dimensional view of collocatios in academic writing. In U. Römer, V. Cortes, & E. Friginal (Eds.), *Advances in corpus-based research in academic writing*. Amsterdam/Philadelphia, PA: John Benjamins.