Afterword

Edward Finegan

University of Southern California

Edward Finegan, Professor Emeritus of Linguistics and Law at the University of Southern California, provides this afterword to synthesize and provide commentary on the six articles in this issue. He has been involved with research on register for more than 30 years, publishing a large number of empirical studies on register and the book *Sociolinguistic Perspectives on Register* (Biber & Finegan 1994, Oxford University Press). He is also co-author on the *Longman Grammar of Spoken and Written English* (Biber, Johansson, Leech, Conrad, & Finegan 1999, Longman), the first comprehensive reference grammar to systematically account for register. He is currently involved in research on the relationship between register variation and social dialect variation. In addition to his research on register, Finegan has made tremendous contributions in the areas of general linguistics, language variation in the U.S., and language attitudes toward correctness, publishing widely-used textbooks in all three areas. He has also been an influential figure in the application of linguistics in legal proceedings, acting as expert witness in many legal cases, particularly those related to defamation and trademark. He currently serves as the Editor of the journal *Dictionaries: The Journal of the Dictionary Society of North America*.

1.   Introduction

The announcement of a new journal called *Register Studies* was cause for celebration months ago, and its arrival now is most welcome and, if the contents of the inaugural issue are any indication, full of promise. For the many researchers already engaged in a range of methodological and theoretical approaches to register – and convinced of its importance in understanding language variation and change – and also for others who may benefit from a focus on register, the journal offers a significant boost. For the inaugural issue, the journal's editors commissioned six scholars to describe the state of register studies in their respective disciplines and subdisciplines – from the vantage point of Systemic Functional Linguistics, text linguistics, variationist linguistics, historical linguistics, applied

linguistics, and computational language research – and to endeavor to anticipate the future.

Leaping from every page of each article in the inaugural issue is proof that studies of register are abundant and thriving. While register analyses are spreading to other languages and language types, to date they are thriving principally in English-language contexts. Another note apparent in several of the articles is that the notion of register and descriptions of particular registers have shown themselves adaptable, even essential, to applications in domains outside linguistics proper – from language teaching and natural language processing to text translation and text generation. Across the wide range of registers that have been analyzed, some of the underlying patterns in the distribution of linguistic features and their contextual correlates in social, communicative, and other situational functions appear widespread, perhaps universal in some ways.

The editors have asked that, in an afterword, I try to synthesize the articles here and add some reflections of my own. Any synthesis across so rich a range of perspectives is not realistic; more significantly, though, it would shortchange the depth of analysis offered in each of the six articles. Instead, I reflect on select observations made in each article. After discussion of the articles, I offer some more general observations.

2.   Register in Systemic Functional Linguistics

"Language sciences" rather than "linguistics" is the rubric Christian Matthiessen inclines toward in characterizing the significant broadening of investigation into language structure and use that has occurred since the narrow view that prevailed in some theoretical circles starting in the late 1950s and the 1960s. Matthiessen identifies what is not universally recognized as a central property of language systems – in fact, of complex semiotic systems more generally – and not accommodated in theoretical frameworks among linguists, namely, the inherent variability and adaptability of languages. Charged by the editors with the task of relating register to Systemic Functional Linguistics (SFL), he lays out SFL's relevant history, tracing it to the mid-20th century and identifying preceding and subsequent functionalist schools. "Academically," he notes, "SFL has always been outward looking and 'permeable'…, engaging in dialogues with anthropology …, sociology …, educational concerns, computational linguistics and AI, neuroscience, film studies, and a range of other disciplines" (Matthiessen, this volume, p. 12). In SFL, register is viewed as the instantiation of language realization across the full range of functional needs – in a nutshell, "variation in language according to context of use" (Matthiessen, this volume, p. 15).

Registers of a language are thus its functional varieties. Whether and to what degree registral variation and a language's named registers are distinct from one another and enumerable or, by contrast, form a continuum isn't clear in SFL, at least in Matthiessen's characterization of a language as "nothing but the total aggregate of registers at any given period during its evolution" (Matthiessen, this volume, p. 16). Probably more so than other functional approaches to language analysis, SFL relies on independent, increasingly familiar, terminology, invoked also in the articles by Conrad and Argamon in this issue. Besides explaining SFL's terms of art, Matthiessen captures the essence of its view of register variation in less technical terms: it is "semantic variation resonating with contextual varia-tion" (Matthiessen, this volume, p. 19). The subtlety of *resonate* in that formulation creates an expansive space for exploration across contexts and linguistic features alike, as well as exploration of an intimate dance as context and linguistic feature partner with one another in still mysterious ways. Matthiessen's exposition distin-guishes "appliable" linguistics (Halliday's term) from the more familiar "applied" linguistics, the former aiming for a theory with the potential for application, the latter presuming (though sometimes ignoring) existing linguistic theory and dis-cussed further in the articles by Conrad and Argamon. In Matthiessen's words, "register variation is semantic variation in the first instance – the meanings at risk in a recurrent context of use" (Matthiessen, this volume, p. 18); in the words of SFL founder Michael Halliday, a register is "recognizable as a particular selection of words and structures [but] it is defined in terms of meanings … the selection of meanings that constitutes the variety to which a text belongs" (Matthiessen, this volume, p. 19).

3.   Text-linguistics approaches to register

As evidenced by the range of his register-related work and the frequency of ref-erences to that work throughout this issue of *Register Studies*, Douglas Biber is a towering figure in the field. Starting with his seminal 1988 book, he has profoundly influenced the field and the work of those working in it. In his contribution here, he focuses on differences between variationist studies of register (as described in Szmrecsanyi's contribution to the special issue) and those Biber has in recent years been calling text-linguistic register analyses.

Biber sees the potential value of register studies not only in their powerful descriptive reach, but beyond that – and perhaps more optimistic than some oth-ers – his is a visionary perspective, seeing register studies as having the poten-tial to reconfigure linguistic theory to a degree, a reconfiguring that would shift emphasis from sentence to text, without in any way discounting the importance

of sentences and their constituents in any text-linguistic analysis. In his framework – and in register studies generally – the text is fundamental, even when, in some register studies, multiple texts are aggregated into a single observation (a convenience recognized as something of a compromise). Not Biber's nor any of the other five articles in this special issue defines *text* or *a text*, even though its pivotal role in register studies more than warrants definition. A text, after all, is the fundamental unit of analysis in register studies. The object of analysis in register studies is the register, and that too needs definition, though the many names attaching to what some researchers call 'registers' convey an idea of the range of objects clustered under that umbrella. Some are broad and high level – spoken and written texts or formal and informal ones; others are relatively narrow – Reddit comments, tweets, and newspaper editorials.

Biber concludes his article with a challenge to bring variationist studies and text-linguistic studies into a unified or 'triangulated' approach that, he says, should make it possible "to learn more about the patterns of linguistic register variation than would be possible through either approach on its own" (Biber, this volume, p. 69). But, he concedes, "[t]ext-linguistic register analysis … differs from many other sociolinguistic approaches in its foundational claim that linguistic variation is functional rather than indexical or purely conventional" (Biber, this volume, p. 45), and he contrasts the variationist approach with "the text-linguistic approach [which] emphasizes the functional basis of linguistic variation, claiming that linguistic features are frequent and pervasive in a register because they perform communicative functions required by the situational context" (Biber, this volume, p. 46).

4.   Register in variationist linguistics

The origins of variationist linguistics are associated chiefly with William Labov, and it is characterized by Benedikt Szmrecsanyi in this issue partly in Labov's words as "a discipline in the field of variation studies that investigates variation between 'alternate ways of saying "the same" thing'" (Szmrecsanyi, this volume, p. 76). Variationists "draw on quantitative methodologies to model the conditioning factors that regulate the way language users choose between semantically and functionally equivalent variants" (Szmrecsanyi, this volume, p. 76). At first blush, readers may be surprised to see variationist perspectives on register in an inaugural issue of a journal called *Register Studies*. After all, as Szmrecsanyi acknowledges, while "analyzing registers means … investigating the *functional relationship(s)* between a set of linguistic features and the situational context," variationist linguistics is "about variation between different ways of accomplishing

the same function" (Szmrecsanyi, this volume, p. 77). Consequently, by definition, "functional variation … does not come under the remit of variationist linguistics," and Szmrecsanyi concedes that "variationist linguistics is largely agnostic about functional relationships" (Szmrecsanyi, this volume, p. 77).

There are two other significant points of difference between variationist linguistics and register studies, Szmrecsanyi notes. First, unlike register studies, "variationist analysis eschews aggregation and proceeds in … a 'single-feature-based' mode … (typically, one variable per research paper)" (Szmrecsanyi, this volume, p. 77). Second, unlike register studies, whose unit of observation is a text, "variationist analysis is focused … on individual linguistic choices, [and] [t]he task … is therefore to link register differences not to text frequencies but to linguistic choice-making" (Szmrecsanyi, this volume, p. 78). Szmrecsanyi captures the variationist perspective on register in a nutshell: "typically conceptualized as stylistic variation in aesthetic preferences … thought of as one of the language-external factors … that regulates variation of individual linguistic variables …. Register is specifically analyzed in terms of how it influences linguistic choices between functionally equivalent variants" (Szmrecsanyi, this volume, p. 78). The vantage point from which variationists view register variation seems so fundamentally different from, if not plainly at odds with, the aims of register studies that Szmrecsanyi's and Biber's expressed hope for a fruitful triangulation seems as much aspirational as realistic. At the same time, as other research in some of the citations here and in Biber's article has shown, what each approach brings to the table can produce competing, seemingly contradictory, findings, and that certainly opens the door to attempts to reconcile what may be seen less as contradictory and more as complementary results.

5.   Register in computational language research

The research Shlomo Argamon describes in his contribution displays a range of applications of multi-dimensional and other methods of analyzing registers. More than that, and not addressed in other contributions here, he examines the important role of register *synthesis* and the role that register *analysis* plays on the flip side of the coin. Most work in register analysis relies on linguistic features whose communicative and social functions are apparent or reasonably well-understood. Work in computational linguistics relies naturally on those same features and explanations. In addition, however, researchers in computational linguistics venture outside more traditional linguistic features, exploring patterns with little and sometimes no readily apparent lexical or structural linguistic basis but that nevertheless prove useful in certain applications. Argamon points, for example, to

punctuation marks and paragraph length as useful in select systems for paraphrasing texts, translating them from one register (or "style") to another, and generating register-appropriate texts. Similarly, alongside transparently linguistic features like passive voice constructions and personal pronouns, such linguistically opaque features as character n-grams have proven useful in setting parameters that constrain "style-related choices" in generating summary texts of medical information and other synthesis tasks (as well as in forensic applications such as authorship attribution).

Argamon points to work by Teich and Frankhauser (2009) that shows, among other things, that analysis of a curated set of registers can provide "insight into how scientists in different disciplines understand the key activities of their disciplines, and how they construct themselves as acting, in their research reports" (Argamon, this volume, p.120). For example, as shown by verb choice in their research reports, the "most significant" activities for computational linguists are experimental (*collect*, *examine*); for computer scientists, formal (*prove*, *define*); and for linguists, communicative (*argue*, *read*) and cognitive (*see*, *feel*).

Argamon concludes with a paragraph worth quoting because it is applicable to so much else of what is offered in the special issue of the journal:

> The other grand challenge is the development of the methodological and computational tools necessary for empirical verification of theories of register. This would seem to require the ability to specify clear articulated models of the causal relationships between situational parameters (social roles and relationships, communication medium, etc.), linguistic features, and intermediate representational levels, and then to test these models empirically, either by analysis of large corpora, or by generating texts according to specification and rigorously measuring human responses to the texts.   (Argamon, this volume, p.126)

6.   Register in historical linguistics

The reliance on notions of register and register-sensitive corpora in examining language variation and language change, especially in English, has been abundant, to judge by the nearly 100 items in Merja Kytö's list of references. While some of the cited studies predate electronic corpora, Kytö makes clear their value in tracing the development of individual linguistic features in a single genre or register (both terms are intentionally used in her article without careful distinction) or even a single author. She also traces the most significant recent studies of variation to the lengthy contribution by Uriel Weinreich, William Labov, and Marvin Hertzog in 1968 and gives a significant nod as well to Suzanne Romaine's 1982 book in initiating socio-historical scholarship. The abundance of scholarly interest over

the past six decades highlights the important role that registers are thought to play in the study of language change. The literature on the interplay between register and feature distribution (chiefly in terms of frequencies of occurrence and correlations with external factors, notably register but also gender, region, time) has been abundant.

One of the striking notes in Kytö's article – one that falls on the challenging side of the ledger – is the relative balance between patterns of distribution and explanations for those patterns in the cited literature. To judge by the analyses described, there is yet too little theorizing as to the role registers play in language change. The literature Kytö cites is chockablock with observations about correlations between frequencies and particular registers and their situational characteristics, but the correlatives are often seen by the studies' authors as self-explanatory of cause. Correlations with gender or register may provide the basis for hypotheses very much worth testing, but they cannot in the end be viewed themselves as explanatory. As Biber points out in his contribution here, exemplification is an essential ingredient in the study of registers and register variation, a point underscored in Conrad's article, but exemplification is illustrative: it is not explanatory. In discussing Paula Rodríguez-Puente's book manuscript on phrasal verbs as a typical register study, Kytö notes it "seeks to pin down the role played by the processes of grammaticalisation, lexicalisation and idiomatisation" (Kytö, this volume, p. 149). Such attempts to relate distributional patterns to the processes that undergird language change are potentially explanatory and should be expected in studies of language change in addition to any correlations that are uncovered.

7.   Register in EAP/ESP

Taking English for Specific Purposes (ESP) as a cover term that includes English for Academic Purposes (EAP), Susan Conrad laments that even though register and register variation are crucial concepts in all ESP work, "most ESP professionals have only a vague concept of register" (Conrad, this volume, p. 172). She includes "as register analysis any studies, or parts of studies, that analyze linguistic features and tie them to their functions in their situational contexts," but she excludes "the more global organizational features included in genre analysis, such as rhetorical moves, except as they are connected to studies of linguistic features" (Conrad, this volume, p. 173).

In her contribution, she reminds us, despite strikingly different terminologies, of the similarities between approaches to register characteristic of ESP and those of Systemic Functional Linguistics (discussed in Matthiessen's contribution). Citing Susan Hunston, Conrad sees "the approaches as parallel but compatible, with

the corpus-based conceptualization having a 'more commonsense notion of the "situation", and moving more from observation to theory, while the SFL perspective theorizes the situational context and moves from theory to observation" (Conrad, this volume, p. 172). When it comes to illustrating the role of register studies in ESP, Conrad finds it "difficult to pick a 'typical' study" and chooses her own examination of passive voice as a feature of impersonal style in engineering practice. In the qualitative analysis in her study, she addressed a matter often overlooked in other studies but crucial in examining passive voice use – namely, the effects of information structure essential to an adequate understanding of passive voice use (as in pursuit of a focused topic chain within a paragraph). While studies of register variation that incorporate passive voice occurrences often, even typically, relate them to impersonal functions, the passive voice has a more basic and naturally functional task in information structure, and Conrad's general acknowledgement of that is useful. She underscores the significant but regrettable fact that corpus-based register studies too often content themselves with a recitation of surprising, even fascinating, statistical patterns but neglect sufficient explanation of the patterns, taking explanation for granted. Quoting an article she co-wrote with Biber and Viviana Cortes, she repeats here this important note: "we do not regard frequency data as explanatory. In fact, we would argue for the opposite: frequency data identifies patterns that must be explained" (Conrad, this volume, p. 178).

8.   Some concluding observations

This first issue of *Register Studies* achieves two important ends. It admirably furnishes readers with theoretical, methodological, and applicational perspectives on register from several research traditions. Even among those who engage in register studies directly or follow them closely, few could fail to be impressed with the range of perspectives and traditions that take register as a central construct in their endeavor to understand language – especially language variation and language change. It also raises questions about methodologically sound research and the larger questions that should motivate linguistic research – likewise valuable contributions of scholarship. In taking up the larger issues that register analysis aims to address, individual elements of the answers will have to be found. It is important at every step for register studies to keep the larger questions in mind rather than diving opportunistically into the magnificent corpus resources now available, utilizing the sophisticated tools at the ready for probing those corpora.

   In this inaugural issue of *Register Studies*, two words repeatedly compel attention. One of course is *register*. Understanding register and its role in language variation and change from different perspectives is at the heart of this journal issue.

The other word is *text*. It is so central that a single article witnesses these collocates, among others: *text type*, *written text*, *natural language texts*, *bodies of text*, *text generation*, *more useable texts*, *source text*, *equivalent text*, *text classification*, *categories of texts*, *natural text*, *training texts*, *input text*, *text analysis*, *text topic*, *text chunkers*, *registers or genres of texts*, *academic biology texts*, *popular nonfiction texts*, *scientific text*, *general text*, *text construction*, *text summarization*, *text length*, and *actual texts*. In other contributions, reference is made to *a complete text*, *learner texts*, *the psychology text*, *the linguistics text*, *the chemistry text*, *the full texts*, *text messages*, *text samples*, *professional texts*, *non-academic texts*. There's even mention of "the intertextual context of the text" and to "text, however defined"! For a quarter of a century a journal now called *Text & Talk* was published simply as *Text*. So readers of this issue of *Register Studies* must wonder just what's meant by a text, a complete text, a general text, a natural text – and just what "text" is and isn't. Whereas other definitional matters are raised in the articles, for example about *register* and *genre* in Conrad's contribution, it is perplexing that *text* and *a text* are taken for granted, as pre-theoretical notions, in all the contributions. It is also remarkable that so much understanding can be derived from the reports published here despite the absence of a definition for the central unit of analysis.

On another note altogether, researchers must take care that feature sets originating for a particular purpose in a particular language apply appropriately to other languages and research questions before employing them elsewhere and for other questions and that features in other languages that could impact the seeming universality of any dimension are not overlooked. The dimensions Biber's research suggest as universal – the oral vs. literate style and the narrative dimension – are intuitively satisfying and the research findings reassuring, and these putative universal dimensions may well prove universal, even if the features defining them across languages turn out to differ to some degree, depending on which features serve which purposes in particular languages.

With respect to the exemplification that is essential in expositions of register variation and change, post-hoc explanation of feature distributions across registers is endemic – to some degree it is necessary – in register studies. Analysis of features and their uses, when done superficially, can be problematical. An occasional register study risks appearing to begin not with a hypothesis but with a probe of convenience: easy because the available corpora and statistical tools to investigate them seem so inviting. Once the statistical results are extracted from the data, an extract from a text in the corpus that gave rise to the statistics is highlighted with typographically adapted fonts to signal the presence of features characterizing the register. Such exemplification is useful and rhetorically necessary, but trouble arises if the correlative situational factor – say, "impersonal" for passive voice structures – applies to only some of the highlighted examples in

the extract, while others seem more clearly or strongly related to a different factor – say, information structure. Readers may legitimately wonder whether such researchers haven't focused on the quantitative patterns to the neglect of deeper explanations that are more difficult to divine but fundamentally more important in understanding studies of register. A continuing challenge for scholars engaged in register studies is that superficial or merely plausible explanations are allowed to suffice as actual explanations instead of what they actually are – plausible hypotheses in need of further investigation.

The scores and scores – indeed hundreds – of published books and articles cited in this first issue of *Register Studies* testify to the vitality of the field. To the extent it remains catholic in its representation of approaches and theoretical perspectives but demanding in its expectation of explanation beyond correlation, as in this special issue, the journal – and register studies themselves – will prosper and prove influential. At the same time, it is crucial that *Register Studies* serve as a resource for those in other fields not (or not yet) closely allied with register studies or aware of their potential. An understanding of the potential import of register and its possible contributions to a growing range of language and semiotic interests will spread the insights that register studies are increasingly uncovering.

Address for correspondence

Edward Finegan
Department of Linguistics
University of Southern California
GFS 301
Los Angeles, CA 90089-1693
USA

finegan@usc.edu