RiCL
Research in Corpus Linguistics

Corpus analysis of engagement discourse strategies in academic presentations

Carolina Viera[a] – Serena AP Williams[b]
Boise State University[a] / United States
Language and Heritage Institute[b] / United States

Abstract – Text analysis informed by Genre Theory (Hyon 1996) and methods in Corpus Linguistics provide the opportunity to describe language patterns that exist not only at the individual level but also in discourse communities. In this study, we investigate the discourse strategies used by novice and expert members of the academic United States (US) Spanish-speaking community to engage their audience, construct interpersonal meaning, and position themselves as expert speakers. We analyze two corpora: a specialized corpus of 32 conference presentations delivered by professors and doctoral students of Hispanic Studies, and a learner corpus of 24 in-class presentations to describe discourse patterning of social engagement expressed in text organization during presentation openings. Results indicate variation in engagement strategies between novice and expert presenters, with professors being the ones who make more use of interpersonal and interactive features to engage their audience. Our findings inform genre-based pedagogies by describing the language functions used to construct the different stages in which openings are organized. As oral presentations have been insufficiently studied (Robles Garrote 2016), this study contributes to the growing knowledge of academic oral Spanish in the United States.

Keywords – academic Spanish oral presentations; genre analysis; engagement; academic literacy; Spanish language teaching

1. INTRODUCTION[1]

Oral presentations are an important academic genre set comprising in-class student presentations, conference presentations, class discussions, lectures, and dissertation defenses, among others (Swales 2004; Biber 2006; Zareva 2012). Despite their importance, academic presentations had not been sufficiently studied until recently

---

(Ventola 2002; Hood and Forey 2005; Seloni 2012; Robles Garrote 2016), partly because the number of available online corpora has increased (Morell and Pastor Cesteros 2018: 126). Corpora of oral language are difficult to construct and analyze in comparison to written language corpora, and this is especially evident when referring to a corpus in Spanish. Regarding learner corpora, Alonso-Ramos (2016: 7) affirms that "[t]here is no Spanish academic learner corpus such as *CALE*", The *Corpus of Academic Learner English* for written texts.

Existing research on academic oral presentations suggests that while academic oral texts overlap in some ways with their written counterparts, distinctive features of these text types are that they showcase "research at various levels of completion, from work in progress to post-publication dissemination" (Hood and Forey 2005: 291–292), and possess a greater spontaneity than academic written texts, especially research articles or essays. Hood and Forey (2005: 292) emphasize that while "the oral performance is strongly associated with the development of a parallel written text," the presenters must interact with an audience in the present time and place, resulting in a more interactive text (Wulff *et al.* 2009; Hyland and Jiang 2017). This highlights "the importance of interpersonal management and politeness features" (Ventola 2002: 10) in oral academic texts.

While the interactive and interpersonal character of written texts has also been studied (Hyland 2005, 2009), oral presentations require a distinct way of establishing rapport with the audience. Perhaps one of the most salient examples of this establishment of rapport is the inclusion of an interpersonal stage known as the 'opening' (Thompson 1994; Rowley-Jolivet and Carter-Thomas 2005; Villar 2011), a kind of preamble to the presentation content which has as its function to establish initial contact, stimulate interest, and create a dialogical setting of solidarity (Hood and Forey 2005: 292). During this stage, presenters introduce themselves, greet and acknowledge the audience, and sometimes make known the limitations of their study. In so doing, presenters utilize different discourse strategies to pique listeners' engagement with the presentation. Openings are vital to facilitating initial understanding, which is crucial when processing information presented in real time. As this opening is not present in written texts, it constitutes a singular distinguishing element of the oral text. Presenters who include openings in their presentations show understanding of the social complexity of academic oral presentations in addition to an understanding of the

differences between oral and written texts. However, oral introductions can pose a problem, especially to novice presenters, because they are "the locus of complex pragmatic choices" (Rowley-Jolivet and Carter-Thomas 2005: 42).

In this study, we analyze the 'opening' in two corpora: 32 conference presentations (CPs) given by professors and doctoral students of Hispanic Studies and a learner corpus of 24 in-class presentations of learners of Spanish in the USA. We describe the language strategies used by both novices and experts to construct interaction or engagement with the audience in two different academic presentation modes. The following research questions guide our study: 1) What interactive and interpersonal discourse features are expressed in the text organization (stages) of the opening, and 2) What discourse elements are associated with expertise in academic public speaking in this context? The study is informed by Genre Theory (Flowerdew 2005; Martin and Rose 2008; Biber and Conrad 2009) and uses Corpus Linguistic methods for the data collection and analysis (Parodi 2008; Gries 2009; McEnery and Hardie 2011; Casas-Pedrosa *et al.* 2013). The study contributes to the growing field of academic oral corpus research through reporting the methodological decisions regarding annotation and tagset creation at the discourse level. The prevailing annotation of corpora is that of parts of speech while discourse-pragmatic annotation is rarer (Alonso-Ramos 2016: 14–15; Gries and Berez 2017). Consequently, the methodological decisions described in this study will be of interest to those pursuing analysis of oral language in academic settings.

Lastly, this study discusses how corpus analysis can contribute to our understanding of the Spanish academic discourse produced in academic presentations in the United States. The context of Spanish in the United States presents additional challenges to speakers in the academic community due to its multilingualism and multidialectalism. Even though research of oral academic Spanish exists in other contexts, it would be erroneous to assume that this discourse community follows the same conventions as other academic discourse communities that use Spanish. Academic oral texts in Spanish in the USA have been rarely studied. Though researchers have begun to address this sociolinguistic context (Achugar 2003, 2009; Viera 2017, 2019), there still remains a gap in knowledge with respect to the conventions of this academic community. The field of Contrastive Rhetoric has made clear that descriptions of texts within one cultural context do not always apply to those of another (Soler-Monreal *et al.*

2011). The creation of specialized corpora such as the ones discussed here allows the identification of distinctive features and discursive strategies of interaction that can later be compared with their use in other academic contexts in which Spanish is used.

2. GENRE-BASED AND CORPUS APPROACHES TO THE STUDY OF ACADEMIC LANGUAGE

Text analysis informed by a corpus approach provides what Flowerdew (2017) considers as an opportunity to describe language patterns that exist not only at the individual level but also in discourse communities: groups of individuals who share common goals, use and generate a set of distinctive text types (genres), develop some specific lexis and have participatory communication methods.[2] In our study, the academic 'Sociorhetorical Discourse Community' (Swales 1990) of focus consists of members who use Spanish in public places with an academic goal. Corpus analysis offers the possibility to analyze their use of the language at a larger scale than individual analyses. As Dressen-Hammouda (2012: 194) points out, these approaches at both the individual and discourse levels seek the "analysis of data toward a 'snapshot' view of language use, by providing a measurement of either the most frequent use or of its average use." Studies analyzing such linguistic snapshots within a framework of Genre Theory have shown that academic discourse communities develop linguistic and discourse conventions that characterize each discipline (Burns 2001; Ciapuscio 2005; Biber and Conrad 2009). Discourse communities favor a set of textual genres or "exemplars that share similarities in structure, style, content and intended audience" (Swales 1990: 58), and that are "staged, goal-oriented social process[es]" (Martin and Rose 2007: 8). That is, production of academic texts occurs in specific contexts that determine linguistic options. With this bottom-up perspective, the analysis of texts created in these specific contexts precedes more general description of academic language patterns, thus calling for studies that add such an analysis to the more general body of knowledge.

Knowledge of genre conventions is vital to becoming an expert member in a discourse community (Swales 2004; Biber and Conrad 2009; Dressen-Hammouda 2012), and represents a challenge to the novice member of the community who has not

[2] See an extended definition, critical review and update of the concept in Swales (2016).

yet fully experienced the process of language socialization. This process implies acquiring a different style or type of discourse through participation in a new social context (Bolívar 2005; Moyano 2009; Seloni 2012). Tailoring a text for a specific, live, academic community, making necessary adjustments to the text while presenting, and interacting with a present audience are important aspects of presentations to be learned by novice members. In-class student oral presentations, which we will call 'academic oral presentations' (AOPs), have a pedagogical objective of adding the skill of public speaking to the student's oral repertoire. In the academic world, conference presentations (CPs) are generally the venue in which public speaking also occurs.

Following the situational framework proposed by Biber (1994) to compare different registers, CPs and AOPs have in common the public place of communication, the planned text, and the common purpose to transfer academic knowledge. In each of these public speaking genres, presenters make their expository texts accessible to their immediate audience with whom they engage and interact. Additionally, presenters are being evaluated by their audiences, which can create language anxiety and interfere in speech production, especially at the initial part of the presentation. In a similar study, Csomay (2015: 4) compares teacher lectures and student in-class presentations and concludes that they differ in: "a) participant characteristics; b) relations among participants and c) production circumstances." More precisely, she points out expertise and communicative purposes as the main differences between these two genres.

As part of the addressor's epistemological stance towards the text, Parodi (2010) indicates that academic discourse should be marked by credibility and prestige. Achugar (2003) states that academic presenters should position themselves in the role of an expert. This positioning is expected in the case of the CPs where the addressor is engaged with the topic of the presentation and usually has the goal of argument in favor of an original idea. In contrast, in the AOPs, the addressor is presenting a topic that has been selected by the instructor and might not be engaging or familiar to the presenter. Expertise is achieved through participation and practice. As such, oral texts produced in AOPs and CPs represent two instances at a continuum of expertise in academic public speaking. An exploratory corpus study permits an initial approach to describing variation in expertise instantiated in the text within this particular discourse community.

3. METHODOLOGY

The data presented in this exploratory study come from a learner corpus and a specialized corpus. We follow a corpus-driven, "inductive approach, which progressively generalizes from the observation of data to build up the theory or rule" (Granger 2011: 13), in this case, focusing on a necessary stage of corpus research: description. The analysis of this study focuses on the opening sub-stages of the presentation (henceforth simply 'stages') in both AOPs and CPs to determine interpersonal and interactive (engagement) discourse features.

3.1. Participants and data

We analyzed two corpora: a corpus of 32 CPs of professors and doctoral students of Spanish language and literature and a corpus of 24 groups of Spanish learners giving academic oral presentations in class. We describe the generic structure of the presentation openings of each. Table 1 describes the general characteristics of each corpus.

		AOP Corpus	**CP Corpus**
General features	Number of texts	24 groups: 91 students	32 (28 tagged)
	Number of words	43,729	74,571
	Total recorded hours	7h39	9h33
	Stage analyzed	Opening	Opening
Participant features	L1	English; English/Spanish bilingual	Spanish (28) English (4)
	Language level	Advanced; superior	Advanced; superior
	Education level	Upper division university	Graduate students (15) Professors (17)
Linguistic features		Spoken Academic Planned Monologic	Spoken Academic Planned Monologic
	Genre	In-class student presentations	Conference Presentations
Textual features	Topic	Sociolinguistics	Literature (18) Linguistics (14)

Table 1: Description of AOP and CP corpora

Twenty-eight CP presenters were native speakers of Spanish, and four were near-native bilingual English and Spanish speakers. Native Spanish-speaking participants originate from various Spanish-speaking countries but completed undergraduate studies at US institutions (18). Three completed bachelor's degrees in Mexico and four in Spain while

the remaining seven participants completed their degrees in other Latin American countries. In the US, university professors are understood to have experience and membership in the academic discourse community, as they are expected to disseminate the findings of their research in public venues; therefore, we assume that professors have gained exposure to the presentational genre and are active participants of this discourse community.

The learner corpus corresponds to what Granger (2011: 11) classifies as a 'local learner' corpus: a smaller corpus "collected by teachers as part of their normal teaching activities and directly used as a basis for classroom materials." Data for AOPs were collected in an upper-division Spanish class in a large Hispanic Serving Institution (HSI) in the West of the United States. The Hispanic-Serving designation is obtained when 25% of degree-seeking domestic students are classified as Hispanic. Spanish classes in the US are characterized by a mixed student population of learners of Spanish as a second language and students who learned Spanish by interaction with their family (Burgo 2017). As a result, learners of Spanish in upper division classes have varying degrees of proficiency in Spanish. In this sense, our AOP corpus is representative of the sociolinguistic learning context in the US. We acknowledge this fact proves problematic for its replicability in other Spanish-speaking communities; however, homogeneity is not a feature of naturally-occurring speech samples, especially in territories or contexts where languages are in contact. Nevertheless, we consider that the methods of this study can be replicated in other contexts in which engagement function is the focus.

3.2. Corpus design and task description

Both corpora were collected between 2011–2012. The CP corpus was collected by one of the authors following all research-with-humans protocol for the protection of rights. The CP corpus data were collected in 8 different professional academic venues in different US regions.[3] Each presentation was part of a panel presentation of between 15 and 20 minutes in duration. Four literature CPs, which correspond to graduate students, did not include an opening stage; therefore, the total number tagged was 28.

The learner corpus was created by Cecilia Colombi, (University of California, Davis). No sociocultural or proficiency-level data accompany this AOP corpus. Since

---

[3] This CP is also described and used with different research purposes in Viera (2017, 2019).

the proficiency level of the presenters was not determined at the time that both the corpora had been created, the researchers assessed proficiency by listening to the presentation video recordings, applying oral proficiency interview assessment standards of the American Council of Teachers of Foreign Language (2012).[4]

Presentations were video-recorded, then transcribed orthographically, manually tagged, and analyzed using a corpus-based approach (McEnery and Hardy 2011). An analysis of the data followed, informed by Genre Theory (Martin and Rose 2007, 2008; Biber and Conrad 2009). Paralinguistic and visual elements were not included in the analysis because we were interested in the textual mode of communication. Both researchers checked the accuracy of all the transcriptions.

The course from which the AOP corpus was created dealt with topics related to Spanish in the United States. Students attended conference presentations, participated in pair and group discussion on each topic, and completed written exams. In addition to serving as a model, the conference sessions offered students the opportunity to learn theoretical concepts. Finally, students produced an oral presentation on one of a selection of linguistic articles related to course content. Although AOPs were group presentations, the opening is mostly delivered by one student in the group. We consider the final text a product of negotiation that reflects the linguistic options of the group. While half the groups (12) were face-to-face presentations, the other half (12) were completed by voice recording on a PowerPoint slide deck. Recorded presentations were listened to and evaluated by the members of the class. Table 2 displays the instructions excerpted from the handout provided to students for this summary task.

1. Read the assigned article.
2. Create a summary of the article.
3. Explain the most important ideas.
4. Use a formal register and academic vocabulary.
5. Use a PowerPoint or other visual materials.
6. Follow this structure:

a. **Introduction:** Introduce yourself. Specify the topic and objective and greet the audience. Announce your topic and goal. Make connections with the class topics.
b. **Development:** Cover the most important points.
c. **Closing:** the closing is as important as the introduction. The function is to remind the audience of the main concepts so they remain in the minds of the listeners.

Table 2: AOP task instructions

---

[4] Please note, however, that these standards were created to rate a conversational mode of communication.

3.3. Corpus annotation

We created a taxonomy and tagset to identify engagement function stages in the opening. The following taxonomies served as a basis for developing a coding scheme: Rowley-Jolivet and Carter-Thomas (2005); the MICASE tagset described in Maynard and Leicher (2007) and Alsop and Nesi (2014). The corpus was manually annotated by both researchers who were familiar with both modes of presentation. Previous studies have noted the difficulty in deciding the boundaries of the tag units in the process of the corpus annotation (Alharbi and Hain 2016; Navarro and Simões 2019). To establish the cut-off points between stages, we followed Rowley-Jolivet and Carter-Thomas (2005) in using textual clues in the transcript and visual clues in the PowerPoints. Textual clues included discourse markers (i.e., *well*, *so*), and visual clues were given by the different slides of the PowerPoint and their corresponding title that acted as text organizers. We also considered pauses and gestures. Table 3 presents an example of establishing these cut-off points in a single AOP opening, along with the codes used for its annotation.

Tag	Transcript
GR (greeting)	*Buenas tardes,* 'Good afternoon'
[pause]	
SI (speaker presentation)	*mi nombre es X* 'my name is X'
[pause]	
TA (topic announcement)	*y a continuación, mis compañeras y yo tendremos a cargo el siguiente capítulo número cuatro titulado: [título del capítulo].* *Por lo cual, pido de su amable atención.* 'following, my classmates and I have been tasked with the following chapter number four titled: [title of chapter]. For this reason, I ask you for your kind attention'
[change of slide]	
PL (plan)	*Para iniciar con el primer tema de esta presentación, [nombre] nos expondrá* 'To start with the first topic of this presentation, [name] will present'
DEF (definition)	*sobre la lingüística sistémica funcional y género.* 'about Systemic Functional Linguistics and genre.'

Table 3. Example of discourse cues used to establish cut-off points between generic stages in one AOP opening (recordedcap4a)

Manual validation was carried out in the totality of the analyzed openings. We used a one-pass re-annotation; that is, the tagging of the corpus was repeated independently by the two different researchers. As different factors might affect intra-coder reliability (Révész 2011: 217), especially for holistic data, the researchers re-coded the data three times. We calculated inter-coder reliability following Miles and Huberman (1984) by dividing the number of agreements by the total number of decisions made. Inter-coder reliability was high (0.95), likely due to the fact that our categories are low-inference categories that "require little judgment" (Révész 2011: 212). Disagreements were resolved through discussion and where disagreement continued, elimination of the annotation from the corpus. Table 4 shows the taxonomy created for the annotation of the opening.

Functional Stage	Description and Function
Greeting	Speakers greet the audience
Contextualizing the topic	Speakers provide background information for the presentation itself or connect the presentation to a major context
Topic announcement	The speaker announces the topic; text functions as a text organizer (like a written title)
Relevance	Speakers claim the importance of the topic (centrality or need)
Personal narratives	Speakers present from the 1st person perspective, usually in the form of an anecdote that explains their interest in the topic of the presentation.
Speaker introduction	Speakers introduce themselves
House-keeping	Speakers pay attention to technical or organizational issues
Defining the topic	Speakers provide a brief explanation of the topic such as explanation, elaboration, clarification, delimiting the scope, exemplifying, reviewing, or stating the focus
Thanks	Speakers thank the audience or moderator
Goodwill	Speakers use any rhetorical strategies to achieve audience solidarity or benevolence such as self-deprecation or asking for forgiveness
Humor	Speakers make use of humor
Presentation plan	Speakers provide an outline of the organization of the presentation

Table 4: Opening structure tagset

3.4. Corpus analysis

A genre perspective usually entails both sequential and distributional analysis. Determining the sequential formula of the different stages is out of the scope of this study which focuses on the distribution of functional stages. Researchers working with genre analysis have proposed that a percentage of occurrence lower than 25% be considered an unstable stage of the generic structure, and values above 75% be considered prototypical, or obligatory, stages of the genre (see Navarro and Simões

2019) for a review. We classified the frequency of sub-stages as a) 25%–45%, occasional; b) 46%–74%, frequent; and c) 75%–100%, prototypical stages.

After the identification and tagging of the stages, we used the concordancer software *AntConc* 3.2.4 (Anthony 2013) to identify and quantify frequent stages. The *AntConc* Concordance Tool and Concordance Plot Tool were used to find the examples of the tags in context and the number of occurrences in the corpora. Absolute frequencies were normalized per 1,000 words. The four sub-corpora in our study where compared to determine differences in engagement discourse features: professors, graduate students, face-to-face, and recorded presentations.

3.5. Corpus size and representativeness

The size of the analyzed corpora is similar or larger than those discussed in the existing literature for academic oral language (see Wulff *et al.* 2009 or Robles Garrote 2016), a size that is smaller than typical written corpora because spoken data are more difficult to collect than written corpora and entail a time-consuming transcription stage. Because of the size and representativity of our corpus, our analysis applies only to our corpus: a pilot corpus that can inform a future larger corpus study. Despite its limitations, to our knowledge, no other similar corpus has been compiled in regard to spoken academic US Spanish. Therefore, the description and analysis hereby presented constitute a contribution to the field of Language for Specific Purposes as well as to genre-based approaches to teaching and learning Spanish.

4. RESULTS

Below, we discuss the engagement discourse strategies instantiated in the text structure of the student presentation openings and expert conference presentation openings, including the frequency of such strategies in both corpora in order to provide a description of the sub-stages used in each genre. We also compare the engagement discourse strategies of professors and graduate students within the conference presentation openings.

Table 5 reflects the number of participants that incorporated each generic stage into their presentation as well as the percentage of total participants using that stage.

The table also shows the number of individual occurrences of the feature in the corpus, indicating the frequency in each stage. The normalized frequency is indicated per 1,000 words (N=2,492).

Stage Used in Opening (Word count = 2,492)	Participants N=28	Participant use (%)	Raw frequency	Normalized frequency
Announcing the topic	19	67.86	22	8.83
Contextualizing the topic	18	64.29	28	11.24
Defining the topic	14	50.00	28	11.24
Giving thanks	14	50.00	20	8.03
Personal narratives or personal asides	11	39.29	25	10.03
Explaining relevance of topic	10	35.71	4	1.61
Greeting audience	9	32.14	9	3.61
Goodwill	12	42.86	17	6.82
Housekeeping	9	32.14	10	4.01
Humor	6	21.43	12	4.82
Presentation plan	3	10.71	2	0.80

Table 5: CPs opening structure (frequencies per 1,000 words)

Text structure analysis shows that the most frequent CP opening stages are "content-oriented and listener-oriented" in terms of Rowley-Jolivet and Carter-Thomas (2005), while the frequency of other interpersonal strategies in the CPs is occasional (less than 36%). The high-frequency stages that orient toward content include 'topic announcement', 'contextualization', and 'defining the topic'. After announcing the topic, which is the equivalent of a title in the written mode, speakers provide background information for the presentation itself or connect the presentation with other related topics that construct shared knowledge, and succinctly define the topic by elaborating, clarifying, delimiting the scope, exemplifying, reviewing, or stating the focus of the presentation. The following examples in Table 6 illustrate the functions of the high-frequency content-oriented stages.

Topic announcement	The speaker announces the topic; text functions as a text organizer.	1. *OK el título es (3LitP)* 'OK the title is' 2. *ah mi presentación ah tiene que ver con lo que es…(8LitE)* 'uh my presentation ah has to do with what is…' 3. *Bueno, yo titulé mi presentación ah (10LinP)* 'Well, I titled my presentation uh'
Contextualization	Speakers provide background information for the presentation itself or connect the presentation to a major context	4. *trabajé con estudiantes en México así que lo que voy a presentar (10LinP)* 'I worked with students in Mexico so what I'm going to present' 5. *eeh un trabajo que consta de tres partes (19LitP)* 'eeh a study that consists of three parts'
Defining the topic	Speakers provide a brief explanation of the topic such as explanation, elaboration, clarification, delimiting the scope, exemplifying, reviewing, or stating the focus	6. *en otras palabras, lo que se conoce como…(3LitP)* 'in other words, what is known as…' 7. *Entonces, un poco, este es justamente el entrecruce de esos dos capítulos. (10LitEH)* 'So, in a way, this is the point at which these two chapters intertwine' 8. *más concretamente, es una puesta en común de…(10LitE)*'

Table 6: High-frequency content-oriented opening stages

The high-frequency stages that orient toward listeners include giving thanks and personal narratives or personal asides. Table 7 displays functions of the high-frequency listener-oriented stages among the CPs.

Personal narratives	Speakers present from the 1st person perspective, usually in the form of an anecdote that explains their interest in the topic of the presentation.	9. *Cuando empecé a hacer esta investigación mi mii idea era encontrarme con estudiantes recién llegados, ¿verdad? (11LinP)* 'When I started this study, my, my idea was to meet with recently-arrived students, right?'
Giving thanks	Expressing appreciation to organizers, audience-members, or other relevant individuals	10. *Gracias por venir (1LitEM)* 'Thank you for coming' 11. *Gracias a los organizadores (1Lit EM)* 'Thank you to the organizers' 12. *Gracias Fernando (6LitPH)* 'Thanks, Fernando'

Table 7: High-frequency listener-oriented opening stages

Table 8 displays examples of the less frequent engagement opening stages.

Relevance of the topic	Speakers claim the importance of the topic (centrality or need)	*13. ehhh sobre todo lo que quiero llamar la atención de ustedes que trabajan con el.. la alguno…el grupo latino, que en muchos de ellos pueden llegar a ser indígenas, ¿verdad? (4LinP)* 'ehhh above all what I want to call to your attention is'
Goodwill	Speakers use any rhetorical strategies to achieve audience solidarity or benevolence such as self-deprecation or asking for forgiveness	*14.…es un poquito más complicada cuando se trata de aplicar al Caribe, ¿no?; y no es que sea imposible, ¿no? pero para mí, en este momento ha sido un poquito difícil, ¿no? Entonces, este trabajo muestra esa dificultad (3LitP)* 'it's a bit more complicated when applied to the Caribbean, right?; and it's not that it's impossible, right? but for me, at this time it's been a bit difficult, right? Thus, this study demonstrates that difficulty.'
Humor	Speakers make use of humor	*15. Entonces<FM>…no puedo leer con las gafas (risas) (6LinE)* 'So… I can't read with my glasses (laughter)'
Presentation plan	Speakers provide an outline of the organization of the presentation	*16. la estructura de mi presentación es esta empiezo con la pregunta central luego voy a hablar un poco brevemente<PL> (3APNLin)* 'the structure of my presentation is this: I start with the central question and then I will speak briefly'

Table 8: Less frequent interpersonal stages in CP openings

While the presenters seem to vary in terms of the selection and frequency in types of these other less frequent stages, it is important to note that when considered together, we conclude that there is an overall consistent attempt by all speakers to include interpersonal stages in the opening; on average, presenters include 5 distinct interpersonal stages in an average opening (see Figure 1 below).
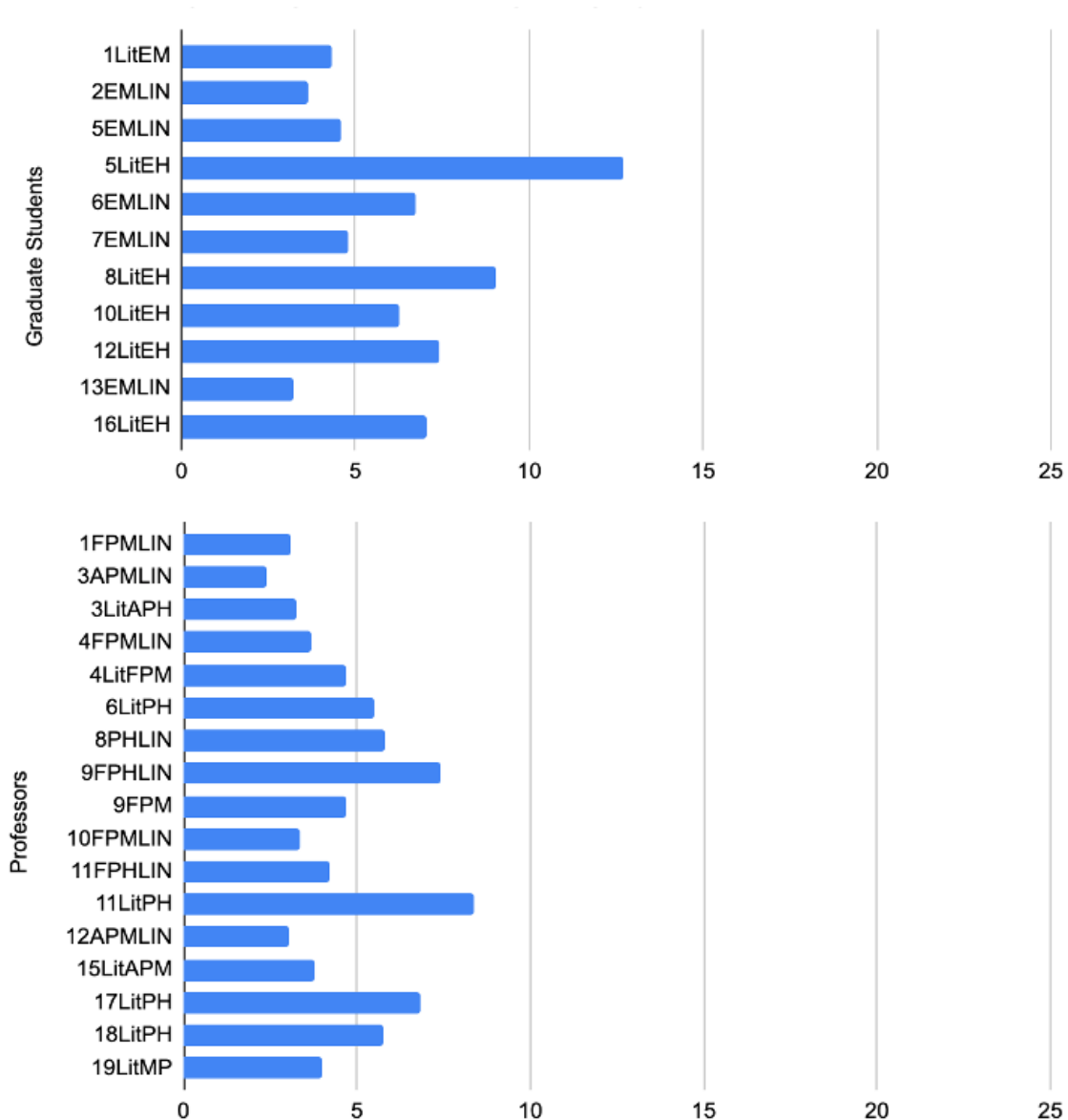
Figure 1: Number of interpersonal generic stages in CP opening (per 89 words, average CP opening length)

As graduate students may be considered peripheral members of the discourse community of academic conference presenters, we describe the frequencies of the opening stages for each group of both professors and graduate students in order to see if any distinctions in stage use exist. Table 9 below shows the stages used by the 17 professors in the CP corpus. We classified the frequency of stages as a) 25%–45%, occasional; b) 46%–70%, frequent; and c) 71%–100%, recurrent.

Stage Used	Professors N=17	Participant use %	Raw frequency	Normalized frequency
Announcing the topic	13	76.47	16	6.42
Contextualizing the topic	13	76.47	20	8.03
Personal narratives or personal asides	10	58.82	22	8.83
Defining the topic	9	52.94	12	4.82
Goodwill	9	52.94	13	5.22
Housekeeping	6	35.29	6	2.41
Giving thanks	8	47.06	13	5.22
Humor	5	29.41	9	3.61
Explaining the relevance of the topic	3	17.65	3	1.20
Greeting audience	2	11.76	2	0.80

Table 9: Professors CPs opening structure (frequencies per 1,000 words)

The most frequent content-oriented opening stages for professors include topic announcement (77%) and contextualization (77%). Graduate students also frequently utilized content-oriented opening stages (topic announcement 55%, topic definition 55%, and topic contextualization 46%); however, these frequencies do not arrive at the recurrent classification as they do for professors. For the listener-oriented opening stages, professors frequently used personal narratives or asides (59%) to engage with their audience while graduate students rarely made use of this stage (9%), preferring instead to utilize an audience greeting (64%).

In the next section, we present the findings of the text structure analysis of students' academic presentations. Table 10 reflects frequency categories for the AOP openings, with each stage within the opening showing the individual occurrences of the feature in the corpus and the number of groups that incorporated this stage into their presentation. The percentage is calculated to reflect participant usage, and the subsequent frequency categorization is indicated for each stage as well.

Stage Used	Participants (N=24)	Participant use (%)	Occurrence (hits)	Normalized Frequency
Speaker introduction	22	91.67	31	28.86
Announcing the topic	22	91.67	24	22.35
Greeting audience	20	83.33	25	23.28
Defining the topic	20	83.33	53	49.35
Contextualizing the topic	19	79.17	35	32.59
Explaining relevance of topic	4	16.67	5	4.66
Presentation plan	2	8.33	2	1.86

Table 10: Schematic structure of AOP openings (frequencies per 1,000 words)

Similar to the expert group results, students frequently include stages that are content-oriented: announcing, defining, and contextualizing the topic (see Table 11 below for examples). Indeed, when looking at the distribution of the stages, the percentage of inclusion of these stages is higher in the AOP corpus, which is an expected outcome considering that all AOPs in this corpus were collected in a similar context and students followed assignment guidelines provided by the instructor.

Contextualizing the topic	Speakers provide background information for the presentation itself or connect the presentation to a major context	17. *...en español, en los Estados Unidos (Rec10c)* *[in Spanish, in the United States]* *la populación de los hispanos en los Estados Unidos está creciendo cada día (live85)* 'the population of Hispanic in the United States is growing every day' 18. *..una educación formal con el español puede extender el conocimiento de la lengua (Rec 9b)* 'a formal education with Spanish can extend knowledge of the language'
Defining the topic	Speakers provide a brief explanation of the topic such as explanation, elaboration, clarification, delimiting the scope, exemplifying, reviewing, or stating the focus	19. *El propósito de este estudio es encontrar las cuestiones relativas a la adquisición del español (live 74)* 'The purpose of this study is to find questions relative to the acquisition of Spanish.' 20. *Y el propósito del estudio es para analizar la comparación entre el nivel um del español recibido en el aula y el porcentaje de formas, consideradas... consideradas no estándares en la producción oral de los hablantes mexicanos americanos (live80)* 'And the purpose of the study is for analyzing the comparison between a level um of Spanish received in the classroom and the percentage of forms, considered… considered nonstandard in the oral production of Mexican-American speakers.'
Topic announcement	The speaker announces the topic; text functions as a text organizer	21. *vamos a hablar sobre capítulo cinco, la enseñanza del español en Nuevo México (live54)* 'we will talk about Chapter Five, the teaching of Spanish in New Mexico]' 22. *y vamos a discutir el capítulo de este libro que se llama XX (Live74)* '[and we will discuss the chapter from this book that is called XX'

Table 11: Content-oriented opening stages for AOPs

Of interest for our research question on generic structure, other than greetings, students do not include interactive or interpersonal stages that were present in the expert corpus of reference, such as 'personal narratives', 'humor', 'goodwill' or 'housekeeping.' Students do include an interpersonal stage of speaker introduction which is not present in the CP corpus, and as explained in the discussion below, likely motivated by the assignment instructions in which presenters are asked explicitly to introduce themselves to their audience.

Additionally, in this student corpus, 12 of the presentations were conducted using PowerPoint narration while the other 12 were presented in a face-to-face context. Table 12 below shows that all content-oriented stages were frequent, but that there were fewer topic announcements and contextualizing stages with the face-to-face mode than with the PowerPoint narration mode.

	Topic announcement	Contextualizing	Defining
Face-to-face (12)	83 %	67%	83%
PowerPoint with recorded narration (12)	100 %	92%	83%

Table 12: Content-oriented stages of face-to-face and PowerPoint-narrated presentations

In sum, a text structure analysis of the openings of conference presentations and student presentations shows high-frequency content-oriented stages, but a difference in structural component categories and their frequency. This difference is mostly at the level of listener-oriented stages. The stages used by both experts in the reference corpus and students are summarized in Figure 2 below.
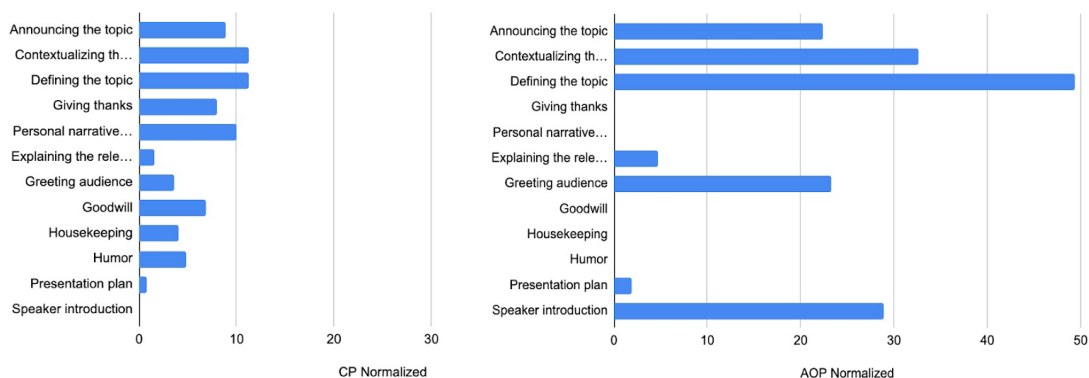

Figure 2: CP and AOP opening structure

Conference presentations use a wider variety of stages than student presentations, though both announce, contextualize, and define the topic, and to some extent, explain the role of the presentation, orienting to content. The listener-oriented stages used by students are fewer, mainly greetings and speaker introductions, while conference presenters give thanks, tell narratives, use goodwill and humor, and manage housekeeping issues.

5. DISCUSSION

As discussed in previous sections, the differences that we have found in the structure of CP and AOP openings reflect the variation in situational contexts of both speech events. For instance, AOP openings include a stage which is not present in the CP corpus where the speakers introduce themselves. In a conference presentation, moderators introduce the speaker, which makes this stage unnecessary. However, our analysis shows that even though these two events have different functional goals, they have structural similarities that are the result of both being public, academic speech events where speakers present cognitively-demanding information and must persuade the audience of their capabilities as valid academic communicators. Openings are crucial to achieve this interpersonal communication. Consistently, most presentations (28) analyzed in this study include an opening. With the audience in mind, they acknowledge the audience through greetings and giving thanks for their presence and make an effort to facilitate the understanding of the content of the presentation. However, at a closer look, we notice differences that show how language socialization has an impact on the academic oral texts produced by the members of a discourse community.

First, we notice that openings were absent in four graduate student conference presentations, whereas all professors included this stage. In the AOPs, students were instructed to include an opening in their presentations; therefore, the presence of openings in this corpus reflects task instructions. However, students create this stage in a very basic way, usually keeping language at the sentence level (see Table 11). For instance, in most AOPs, the topic announcement, an obligatory stage in the professor sub-corpus, is realized by stating the number of the chapter being presented or reading aloud the title of the chapter, without further defining its scope or connecting the topic of the presentation with other topics or theories discussed in class. It is important to note that the task instructions mentioned to state the goal and make connections with class

topics as part of the AOP introduction. Thus, while students include obligatory content-oriented stages in their openings— topic announcement, contextualization and defining the topic—linguistically, they construct these stages in a simpler way than professors tend to do. In doing so, they communicate less investment in engaging or facilitating the comprehension of the information they will present. One possible explanation is that students lack the language proficiency to accomplish this function, but this upper division class consisted of advanced speakers of Spanish who were able to present a complex sociolinguistic chapter in an appropriate way. Additionally, the fact that the graduate students, who were mostly native speakers of Spanish, made use of these functions, but less frequently than professors, suggests that pragmatic awareness rather than proficiency may explain the less frequent use of engaging listener-oriented stages in the opening.

Another difference between the professor corpus and the student corpus is that expert openings are divided into more stages. These additional stages consist of personal asides, housekeeping, and humor that the speaker creates in response to a specific circumstance. In the professor sub-corpus, we notice that 'personal narratives' and 'personal asides' are included in 21% of the presentations. More experienced presenters in our corpus (professors) construct a scholar identity with an active agency in the process of investigation, one in which their motivations and personal stories related to the topic or the research are equally important to the information presented. This approach is consistent with Hyland (2005: 173), who affirms that academics position themselves not as "simply producing texts that plausibly represent an external reality, but also as using language to acknowledge, construct and negotiate social relations." In contrast, in their introductions, both graduate and undergraduate students focus on presenting information without further interaction with the audience or attempts to make the information personal or relevant. The focus on the information expressed in content-oriented stages in AOPs and lack of interaction create a text in which speakers do not position themselves in dialogue with the audience, and establish a more formal text register (Poynton 1989). In doing so, they distance themselves from the information they are presenting. This is evident when comparing face-to-face to recorded presentations in our corpus. One would expect that the first ones evidenced the presence of the audience by the inclusion of more interpersonal discourse strategies; however, the analysis shows few differences between them. It is also important to note

that making the information presented relevant at a personal level (making connections) was part of the task assignment. To the contrary, professors favor interaction and solidarity with the audience in their openings.

The professor sub-corpus analyzed here shows a discourse patterning of social engagement expressed in text organization. Interestingly, in our corpus we see a progression with respect to the importance of such strategies in relation to expertise. Regarding the differences, we observe that professors are the ones who make more use of interpersonal and interactive features of the language. Professors' openings are the site for the inclusion of personal narratives that connect the topic to the personal interests of the presenter, humor, house-keeping, and request for the benevolence or understanding of the audience if the work presented is inconclusive or a technical problem arises at the moment of the presentation. An incipient use of these strategies is seen in the graduate student sub-corpus and absent in AOPs.

By developing the taxonomy used in this research, we found that some of the categories described in previous research on academic presentations do not apply to our contexts, which to our knowledge, is a novel contribution to the field. For instance, 'presentation plan', 'explaining the relevance of the topic', and 'greeting the audience' are not prototypical stages in our expert (professor) corpus.

6. CONCLUSIONS

The first step in becoming aware of the main characteristics of a genre is through description because genres vary according to contexts. This study highlights that even when looking at a discrete stretch of discourse, such as the opening of a presentation, it is evident that academic discourse is complex. Our results demonstrate the importance of language description via discourse analysis and corpus research. Considering that, to our knowledge, there are no larger corpora available of oral academic Spanish in the context of the United States, this exploratory study offers a preliminary view and tools to develop more representative, larger studies.

We identified the most frequent stages of the openings in two different speech events that are part of what are collectively considered to be 'academic presentations'. We found that the number of engagement discourse strategies, which are a distinctive feature of this genre, progress with expertise. Professors create openings that facilitate

understanding of the information. They also situate themselves as active and engaged producers of the knowledge they present. In our findings, even graduate students, who are more experienced and engaged than undergraduate students, showed incipient use of such strategies. Exploring the reasons why students focus on the informational aspect of the communication escape the scope of this study but represent an interesting avenue of investigation. Our findings assert the need to investigate whether explicit instruction of academic discourse, with a genre-based approach (Schleppegrell 2004; Martin 2009) would impact language development of these markers in advanced oral proficiency in our corpora. Genres are learned through exposure, practice, and explicit teaching (Swales 2004; Fang *et al.* 2006; Antilla-Garza and Cook-Gumperz 2015) and identifying novice and expert discourse strategies provides instructors valuable information about what might be explicitly taught.

As it is the case with exploratory corpus studies, we believe that one of our main contributions can be found in the methodological decisions taken during the research process. Since we describe different genres, the analysis yielded a tagset that can be used in different contexts in future studies to analyze engagement, a crucial discourse skill for public presenters. The findings of this study, with respect to the preferred stages of an opening can also inform teaching activities designed to promote advanced literacy. Corpus informed, educational research (even a small-scale one) may contribute to our understanding of the patterning of Spanish academic discourse in specific contexts.

REFERENCES

Achugar, Mariana. 2003. Academic registers in Spanish in the U.S.: A study of oral texts produced by bilingual speakers in a university graduate program. In Ana Roca and Maria Cecilia Colombi eds. *Mi Lengua: Spanish as a Heritage Language in the United States, Research and Practice*. Washington: Georgetown University Press, 213–234.

Achugar, Mariana. 2009. Constructing a bilingual professional identity in a graduate Classroom. *Journal of Language, Identity and Education* 8/3: 65–87.

Alharbi, Ghada and Thomas Hain. 2016. *The OpenCourseWare Metadiscourse (OCWMD) Corpus*. LREC. http://www.lrecconf.org/proceedings/lrec2016/pdf/1085_Paper.pdf (April 11, 2020.)

American Council of Teachers of Foreign Language (2012). https://www.actfl.org/ (April 11, 2020.)

Alsop, Siân and Hilary Nesi. 2014. The pragmatic annotation of a corpus of academic lectures. In Calzolari Nicoletta, Kalid Choukri, Thierry Declerck, Hrafn Loftsson and Bente Maegaard eds. *Proceedings of LREC 2014, Ninth International Conference on Language Resources and Evaluation*, 1560–1563.

Alonso-Ramos, Margarita. 2016. *Spanish Learner Corpus Research: Current Trends and Future Perspectives*. Amsterdam: John Benjamins.

Anthony, Lawrence. 2013. *AntConc* (Version 3.2.4). http://www.antlab.sci.waseda.ac.jp. (April 11, 2020.)

Antilla-Garza, Julie and Jenny Cook-Gumperz. 2015. Debating the world – choosing the word: High school debates as academic discourse preparation for bilingual students. *Linguistics and Education* 31: 276–285.

Biber, Douglas. 1994. An analytical framework for register studies. In Biber Douglas and Edward Finegan eds. *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press, 31–58.

Biber, Douglas. 2006. *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.

Biber, Douglas and Susan Conrad. 2009. *Register, Genre and Style*. Cambridge: Cambridge University Press.

Bolívar, A. 2005. Tradiciones discursivas y construcción del conocimiento en las humanidades. *Signo y Seña* 14: 67–91.

Burgo, Clara. 2017. Meeting student needs: Integrating Spanish heritage language learners into the second language classroom. *Hispania* 100/5: 45–50.

Burns, Anne. 2001. Analysing spoken discourse: Implications for TESOL. In Anne Burns and Caroline Coffin eds. *Analysing English in a Global Context: A Reader*. London: Routledge, 123–148.

Casas-Pedrosa, Antonio, Jesús Fernández-Domínguez and Alejandro Alcaraz-Sintes. 2013. Introduction: The use of corpora for language teaching and learning. *Research in Corpus Linguistics* 1: 1–5.

Ciapuscio, Guiomar. 2005. La noción de género en la Lingüística Sistémico Funcional y en la Lingüística Textual. *Revista Signos* 38: 31–48.

Csomay, Eniko. 2015. A corpus-based analysis of linguistic variation in teacher and student presentations in university settings. In Viviana Cortés and Eniko Csomay eds. *Corpus-based Research in Applied Linguistics. In Honor of Doug Biber*. Amsterdam: John Benjamins, 1–23.

Dressen-Hammouda, Dacia. 2012. Measuring the construction of discoursal expertise through corpus-based genre analysis. In Alex Boulton, Shirley Carter-Thomas and Elizabeth Rowley-Jolivet eds. *Corpus-informed Research and Learning in ESP: Issues and Applications*. Amsterdam: John Benjamins, 193–216.

Fang, Zhihui, Mary J. Schleppegrell and Beverly E. Cox. 2006. Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research* 38/3: 247–73.

Flowerdew, Lynne. 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: 321–332.

Flowerdew, John. 2017. Corpus-based approaches to language description for specialized academic writing. *Teach* 50/1: 90–106.

Granger, Sylviane. 2011. How to use foreign and second language learner corpora. In Alison Mackey and Susan M. Gass eds. *Research Methods in Second Language Acquisition: A Practical Guide*. Malden: Blackwell, 7–29.

Gries, Stefan. Th. 2009. What is Corpus Linguistics? *Language and Linguistics Compass* 3: 1–17.

Gries, Stefan Th. and Andrea L. Berez. 2017. Linguistic annotation in/for corpus linguistics. In Nancy Ide and James Pustejovsky eds. *Handbook of Linguistic Annotation*. Dordrecht: Springer Science, 379–408.

Hood, Susan and Gail Forey. 2005. Introducing a conference paper: Getting interpersonal with your audience. *Journal of English for Academic Purposes* 4/4: 291–306.

Hyland, Ken. 2005. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies* 7/2: 173–192.

Hyland, Ken. 2009. Corpus informed discourse analysis: The case of academic engagement. In Maggie Charles and Susan Hunston eds. *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum, 110–128.

Hyland, Ken and Feng Jiang. 2017. Is academic writing becoming more informal? *English for Specific Purposes* 45: 40–51.

Hyon, Sunny. 1996. Genre in three traditions: Implications for ESL. *TESOL Quarterly* 30/4: 693–722.

Martin, James. 2009. Genre and language learning: A social semiotic perspective. *Linguistics and Education* 20/1: 10–21.

Martin, James R. and David Rose. 2007. *Working with Discourse: Meaning beyond the Clause*. London: Continuum.

Martin, James R. and David. Rose 2008. *Genre Relations: Mapping Culture*. London: Equinox Publishing.

Maynard, Carson and Sheryl Leicher. 2007. Pragmatic annotation of an academic spoken corpus for pedagogical purposes. In Eileen Fitzpatrick ed. *Corpus Linguistics beyond the Word: Corpus Research from Phrase to Discourse*. Amsterdam: Rodopi, 107–115.

McEnery, Tony and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Miles, Matthew B. and Michael Huberman. 1984. *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills: Sage.

Morell, Teresa and Susana Pastor Cesteros. 2018. Multimodal communication in academic oral presentations by L2 Spanish students. *Journal of Spanish Language Teaching* 5/2: 125–138.

Moyano, Estela Inés. 2009. Negotiating genre: Lecturer's awareness in genre across the curriculum Project at the university level. In Charles Bazerman, Adair Bonini and Débora Figueiredo eds. *Genre in a Changing World*. Indiana: Parlor Press and WAC Clearinghouse, 442–464.

Navarro, Federico and Alex Simões. 2019. Potencial de estructura genérica en tesis de ingeniería eléctrica: Contrastes entre lenguas y niveles educativos. *Revista Signos* 52: 306–329.

Parodi, Giovanni. 2008. Lingüística de corpus: Una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada* 46/1: 93–119.

Parodi, Giovanni. 2010. *Academic and Professional Discourse Genres in Spanish*. Amsterdam: John Benjamins.

Poynton, Cate. 1989. *Language and Gender: Making the Difference*. Oxford: Oxford University Press.

Révész, Andrea. 2011. Coding second language data validly and reliably. In Alison Mackey and Susan M. Gass eds. *Research Methods in Second Language Acquisition: A Practical Guide*. Malden: Blackwell, 201–221.

Robles Garrote, Pilar. 2016. Aportaciones de la lingüística de corpus al estudio de la conferencia como género académico de divulgación científica. *Chimera: Romance Corpora and Linguistic Studies* 3: 1–21.

Rowley-Jolivet, Elizabeth and Shirley Carter-Thomas. 2005. The rhetoric of conference presentation introductions: Context, argument and interaction. *International Journal of Applied Linguistics* 15/1: 45–70.

Seloni, Lisya. 2012. Academic literacy socialization of first year doctoral students in the USA: A micro-ethnographic perspective. *English for Specific Purposes* 31/1: 47–59.

Schleppegrell, Mary. J. 2004. *The Language of Schooling: A Functional Linguistics Perspective*. Mahwah: Erlbaum.

Soler-Monreal, Carmen, María Carbonell-Olivares and Luz Gil-Salom. 2011. A contrastive study of the rhetorical organisation of English and Spanish PhD thesis introductions. *English for Specific Purposes* 30/1: 4–17.

Swales, John. M. 1990. The concept of discourse communities. In John M. Swales ed. *Genre Analysis: English in Academic and Research Settings.* Cambridge: Cambridge University Press, 21–32.

Swales, John. M. 2004. *Research Genres: Explorations and Applications*. Cambridge: Cambridge University Press.

Swales, John. M. 2016. Reflections on the concept of discourse community. *ASp* 69: 7–19.

Thompson, Susan. 1994. Frameworks and contexts: A genre-based approach to analysing lecture introductions. *English for Specific Purposes* 13/2: 171–186.

Ventola, Eija. 2002. Why and what kind of focus on conference presentations? In Eija Ventola, Celia Shalom and Susan Thompson eds. *The Language of Conferencing*. Bern: Peter Lang, 15–50.

Viera, Carolina. 2017. Genre and register variation: Academic conference presentations in Spanish in the United States. In Juan Colomina-Almiñana ed. *Contemporary Advances in Theoretical and Applied Spanish Linguistic Variation*. Columbus: Ohio State University Press, 148–161.

Viera, Carolina. 2019. La informalidad como recurso en el español académico de los Estados Unidos. In Gregory L. Thompson and Scott Alvord eds. *Contact, Community, and Connections: Current Approaches to Spanish in Multilingual Populations*. Wilmington: Vernon Press, 59–82.

Villar, Claudia. 2011. Las presentaciones académicas orales de los estudiantes alemanes de E/LE. Del discurso monológico al dialógico. *Revista Nebrija de Lingüística Aplicada* 10/5: 130–172.

Wulff, Stephanie, John. M. Swales and Kristen Keller. 2009. 'We have about seven minutes for questions': The discussion sessions from a specialized conference. *English for Specific Purposes* 28/2: 79–92.

Zareva, Alla. 2012. Self-mention and the projection of multiple identity roles in TESOL graduate student presentations: The influence of the written academic genres. *English for Specific Purposes* 32/2: 72–83.

Corresponding author
Carolina Viera
Boise State University
World Languages Departament
1910 University Drive
83725 Boise, Idaho
United States
e-mail: carolinaviera@boisestate.edu