Project Narrative for “Language Contact in Argentina: An Opportunity for Hands-On Student Research in Sociolinguistics” [SP 404: Senior Capstone in Spanish Sociolinguistics]

1. Executive Summary. This project proposes to document the language use of German-Portuguese-Spanish speakers in the Misiones province of Argentina. Documentation will take the form of sociolinguistic interviews, which are structured yet naturalistic conversations with speakers about their everyday lives and language use. Such documentation will provide a record of the rich linguistic heritage of the region. It will also serve to provide a robust language data set for researchers exploring hypotheses related to language contact and change over time. I will complete these interviews on-site in Argentina during May/June of 2018, in collaboration with Dr. John Lipski, a researcher from Penn State University.

The students enrolled in SP 404, the Senior Capstone in Spanish Sociolinguistics, will be the driving force behind the transcription, annotation, and analysis of the interviews. In September 2018, students will travel to Penn State University’s Center for Language Science to receive training and mentoring as they begin to develop their initial hypotheses about the Misionero speech communities. They will meet with my collaborator, Dr. Lipski, as well as other language researchers doing work in Misiones and other bilingual communities. Throughout the Fall 2018 semester, they will work in small research teams to develop and test research hypotheses about language use in Misiones. Our class meetings will serve both as a space for students to develop their understanding of language research methods and to expand their awareness of scholarship on language contact and change. At least once a week, each team will share their research progress with their peers in the class. The ultimate goal of the course will be to produce a conference-ready abstract and presentation, attend two conferences, and collaborate on a publication.

2. Background Information/Literature Review. The Misiones province of northeastern Argentina is home to a unique community of trilingual German-Portuguese-Spanish speakers, the majority of whom are small farmers of German-Brazilian descent. These speakers use German as their first home language, Portuguese as a second home or local community language, and Spanish as an extra-community language acquired in late childhood or adolescence. Despite their spoken fluency, the speakers have little to no literacy in any of their three tongues. They are not exposed to standard varieties of the languages they speak at home (Portuguese and German); they are not formally educated in their three languages; and in
the case of Portuguese (from Brazil) and German (from Germany), the speakers have no contact with their linguistic countries of origin. This unique speaker profile distinguishes the Misionero communities from the majority of multilingual communities previously studied, providing a singular opportunity to explore hypotheses on linguistic structures and language contact in the near absence of external social pressures from schools, peer groups, or mass media. Specifically, the students of SP 404 will explore if and to what extent prolonged contact between languages can result in change or similarity over time.

The extant language documentation and analysis of these speech communities is limited, with initial contact having been carried out only recently by Dr. John Lipski, with whom I am collaborating on further data collection. Such documentation is urgent, as both Portuguese and German are rapidly fading in the youngest generation of the Misiones province – a consequence of comprehensive standardized schooling, greater access to Spanish-language mass media, and migration to urban areas of Argentina. The transcription and linguistic analysis of Misionero speech carried out by the students of SP 404 will thus be crucial for documenting the historically rich linguistic landscape of the region.

3. Goals, Objectives, and Significance. The first goal of this research project is to develop a transcribed and annotated corpus of sociolinguistic interviews with the German-Portuguese-Spanish trilingual communities of the Misiones province in Argentina. This corpus will provide a window into how speakers naturally use their three languages in everyday life, preserving a record of the linguistic richness of these communities, even as changing socioeconomic dynamics make the use of all three languages less and less common amongst younger speakers in the region. The corpus will also offer a linguistic data set that other researchers can use to test hypotheses related to language contact and language change.

Though this language data will be collected on-site prior to the start of SP 404 during Summer 2018, the SP 404 students will all be involved in transcribing the content of the sociolinguistic interviews that make up the corpus. They will be responsible for annotating the interview recordings in such a way that future researchers will be able to quickly and easily find subsections of the recordings that will be most relevant for their research. The students of SP 404 will also be among the first to work with data from the Misiones region to explore the potential consequences of prolonged contact between languages at the phonological (pronunciation), morpho-syntactic (grammatical), and lexical (word) levels.
Scholarship on language convergence, whereby two or more languages gradually become more similar over time as a result of language contact, is hypothesized to occur at all three levels of a speaker’s language systems (phonological, morph-syntactic, and lexical), but recent research suggests disparities in the extent to which these different parts of language may be affected. Morpho-syntactic structures like verb conjugation, case marking, and pronoun usage, for instance, may be more resistant to contact-induced change. On the other hand, speakers’ phonological systems show more malleability in contact situations, especially when it comes to the precise articulation of individual sounds. For this reason, the students of SP 404 will work in small teams of two to four members to test separate hypotheses, allowing us as a class to get a more complete understanding of the language data we are analyzing.

Though I expect each student team to develop their own hypothesis, based on key readings and their own observations of the data, I do plan to present them with the following research questions, among others, as a way to orient them during the initial stages of their work:

- Does the use of Portuguese lead to increased vowel nasalization in Spanish and German?
- Is the articulation of voiceless stop consonants like /ptk/ in each language stable or changing?
- Does the variable production of /s/ to mark plural nouns differ from non-contact varieties?
- Is the flexible verb order of Spanish and Portuguese evident in German V2 constructions?

4. Research Methods & Timeline. May-June 2018: Data Collection. Data collection will take place over a period of approximately one month during the summer of 2018, prior to the start of the SP 404 Capstone Seminar in Spanish Sociolinguistics. I will collect the sociolinguistic data myself, along with Dr. John Lipski, a researcher who has significant experience working with the Misionero communities and whose fluency in Portuguese and German complements my fluency in Argentine Spanish. Dr. Lipski is one of only a handful of language researchers to have close ties and personal contacts in this region, thus his collaboration is key for the successful completion of the data collection stage.

In accordance with the protocol detailed in IRB #1212183-1, which has been approved for Exempt Status by the Ball State University IRB, participants will be recruited through Dr. Lipski’s extensive personal contacts in Misiones Province. By means of these contacts, we will request the participation of
individuals capable of speaking Portuguese and German, in addition to Spanish. Since these communities are very small, we will personally visit prospective participants to request their participation.

After giving informed consent, participants will be asked about their knowledge of Portuguese and German, including their use of these languages during childhood and adulthood. They will then be engaged in conversations about their daily activities and interests for approximately 45 minutes to 1 hour. At the end of the interview, participants will be asked to produce given structures in Portuguese and German, and will also be presented with possible constructions in Portuguese and German and asked to comment on the likelihood that such combinations could be used in their communities. Participants will be asked to verbally consent to the audio recording of their responses. No written consent form will be signed given the widespread lack of literacy in the community. Participants who consent to be recorded will also be asked if they will allow the recordings to be archived indefinitely for future researchers.

**July 2018: Data organization.** After data collection is completed, all interview files will be transferred to a password-secured cloud storage account, as well as to the hard drive of a computer housed in my office in North Quad building on Ball State University. I will complete an initial review of the files to ensure that no personally identifying information is present on the recordings, and if I identify such information, I will remove it from the interview files using software to delete the relevant audio data.

**August 2018: CITI Training, Introduction to Transcription, Extant Research.** During the first week of the semester, all students in SP 404 will complete CITI training, and an amendment to the IRB submission will be added to allow for their participation in this research project. Students will be asked to download free open-source transcription software during the first week, and an in-class tutorial will be provided to them during the second week to show them the basics of the software. During the first week of classes, I will also submit a purchase order for headphones. This will ensure that I order only the number of pairs necessary, since I will at that point know the final enrollment for SP 404. I will complete a student travel request for the September trip to Penn State, as well.

During this time, students will also complete a series of readings on sociolinguistic variation in the Spanish-speaking world, and will be introduced to established research methods in the field, including data collection, data organization, and data analysis. They will have several small assignments that will
require them to listen to the interview recordings and begin partial transcription. The aim of these assignments, combined with targeted readings and class discussions, will be to expose students to the language data while also encouraging them to begin generating potential research hypotheses that they would like to explore. Starting from the beginning of the semester, one of our three weekly classes will be organized as a lab meeting during which students share ideas, ask questions, and discuss progress in their review of the literature, as well as their transcription and annotation progress.

**September 2018: Penn State Visit, Hypothesis Generation, Data Annotation.** On either September 5-8 or September 19-22, the SP 404 class will visit Penn State University’s Center for Language Science. There, students will receive training and guidance from the foremost researchers in Misionero language contact: Dr. John Lipski and Dr. Michael Putnam. They will discuss their research ideas, attend a lecture on transcription and data analysis methodologies, and visit the research and data collection facilities. I have already confirmed with Dr. Lipski his willingness to meet with the students. He was enthusiastic in his response and eager to meet with students from Ball State to discuss their research.

During regular class time in September, our energies will be devoted to continuing our review of the relevant literature, beginning transcription and annotation of the interview files, and discussing how the Misiones interviews in particular can be used to explore the influence of language contact.

**October-December 2018. Data Annotation, Class Presentations, and Abstracts.** Throughout October, students will continue to annotate data from the Misiones interview. This is a time-consuming process, so each team will have weekly assignments to find and annotate a specific number of speech segments relevant to their project to keep them on schedule. By the beginning of November, students should have completed the data annotation and will begin data analysis. I will guide them in the use of mixed effects modeling in R to analyze their findings, and we will spend several class meetings completing this work. By the end of the Fall 2018 semester, students will produce a conference-ready abstract and presentation of their projects. During Spring 2019, interested students will apply to and present at two undergraduate research venues, and in early Summer 2019, they will submit abstracts to peer-reviewed linguistics conferences. The ultimate goal, following these presentations, is to collaborate with students on a set of publications to disseminate their findings within the field of language research.