UTTERANCE- AND PHRASE-INITIAL PARTS OF SPEECH IN GERMAN
INTERACTIONS AND TEXTBOOKS

by

CHRISTOPHER FABIAN SEIDLER

B.S., Kansas State University, 2013

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Modern Languages
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Dr. Janice McGregor

Copyright

CHRISTOPHER FABIAN SEIDLER

2015

Abstract

The current study investigates phrase-initial parts of speech as found in intermediate German textbooks and compares these findings to utterance-initial parts of speech as found in spontaneous speech in German-language interactions. This is important, because learning and using German word order appears to be a struggle for German learners whose first language is English. Research has shown that possible word order realizations in a language are partly restricted by the parts of speech system of that language (Hengeveld, Rijkhoff, & Siewierska, 2004; Vulanovic & Köhler, 2009). This is important because English and German have different parts of speech systems (Hengeveld et. al., 2004; Hengeveld & van Lier, 2010). Doherty (2005) analyzed English to German translations of an international science magazine and found that almost every second sentence begins differently. Instead, this study looks at talk in contexts of use and compares these findings with textbook language because, in recent years, communicative approaches to language teaching have been adopted by a large number of US German language programs. One would thus expect that textbooks used in these classrooms would contain at least some input with constructions that are typical to contexts of use.

The results of the study indicate that construction-initial parts of speech in textbooks and in contexts of use are quite different. These differences imply that if it is a communicative approach that is being promoted, textbook authors and German educators would do well to expose students to actual talk from contexts of use so that they might learn to make meaning based on considerations of context.

Table of Contents

List of Figures

Acknowledgements

1 Introduction

1.1 General Introduction

Learning and using German word order appears to be a struggle for German learners whose first language (L1) is English. This is because word order primarily determines the grammatical roles of elements within a sentence in English, whereas in German, case markings demonstrate who does what to whom (Jackson, 2007). In her work, Doherty (2005) compares the beginnings of sentences from English into German translations in the international science magazine, *New Scientist*. She finds that almost every second sentence begins differently. She attributes this difference to the process of topicalization, which occurs almost two times more frequently in German sentences as compared to English sentences. Topicalization, as defined by Reinhart (1981), is what the utterance is about. Doherty notes that "in German, topicalization refers to a position before the verb" (p 182). Jacobs (2001) reports that natural languages topicalize differently. In addition, he notes that the initial word of a sentence or question may not be what is considered the topic. Below are two examples of topicalization that have the same meaning (i.e, 'The lion ate the bird.'). Example 1 below is considered to be about the lion (the subject of the sentence), while example 2 is about the bird (the object of the sentence). Therefore, in example 2, 'the bird' is topicalized.

> 1) Der Löwe aß den Vogel.
> *the lion-NOM ate the bird-ACC*
> The lion ate the bird.

> 2) Den Vogel aß der Löwe.
> *the bird-ACC ate the lion-NOM*
> It was the bird that the lion ate.

In German, the finite verb must appear in the second position for main clauses (Saville-Troike, 2012; Mihalicek & Wilson, 2011; Schumacher & Hung, 2012). Also in German, the

clause-initial constituent is typically the subject, though adverbial phrases can replace the clause-initial constituent (Schumacher & Hung, 2012). This can be seen below in example 3; an adverbial clause appears as the first constituent and the subject must now appear in third position.

3)   Gestern aß der Löwe den Vogel.
        Yesterday ate the lion-NOM the bird-ACC
        Yesterday, the lion ate the bird.

This flexibility in German word order is permitted largely due to case markings. These case markings exist in particular parts of speech (POS), such as personal pronouns, definite (e.g. *the*) and indefinite articles (e.g. *a* or *an*), adjectives and some 'weak nouns' (e.g., *Held* or 'hero' becomes *Helden* or 'heros' in any other case excluding nominative) (Walter & Compernolle, forthcoming). These case markings are exemplified in example 4, a sentence composed of three masculine nouns. Each determiner and adjective are declined differently indicating to German speakers the noun's case or relationship to the verb in the sentence.

4)   Der freundliche Manager verkauft dem netten Mann den neuen Laptop.
        The friendly manager-NOM sells the nice man-DAT a new laptop-ACC.
        The friendly manager sells the nice man the new laptop.

For this reason, a sentence containing multiple phrasal components (e.g. subject, direct object, indirect object, prepositional phrase, adverbial phrase, etc.) could relationally have just as many grammatically correct realizations, since each noun phrase contains morphological information for the speaker to determine the role of the noun within the sentence. However, in actual use, questions of pragmatics and context arise. Doherty (2005) notes that selecting word order in German is constrained by the fact that one must select a structure that best fits the particular context. Doherty's data do not consist of spontaneous speech, however she examined scientific

2

English texts with German translations over the course of two years for topicalization of adjuncts and arguments in the Berlin corpus.

Thus, in this paper, I analyze word order realizations in contexts of use, spontaneous speech, and compare these with word order in textbooks based on POS. Given that many language programs in the United States have adopted communicative approaches to language teaching, meaningful language use in context is frequently encouraged. A communicative approach does not consist of a single method but involves developing communicative competence and emphasizes meaningful and purposeful interactions through language (Richards, 2006). Chuang (1993) notes that teaching materials may be the biggest factor in student's acquisition of a language. Therefore, German textbooks used in communicative classrooms should contain input with varied word orders, specifically those represented in contexts of use in order to align with many universities' approach to language teaching. It is for these reasons that in this study, I focus on POS that follow a full stop in two contexts: utterances in German conversational speech and sentences and phrases in intermediate German textbooks. A full stop is the indicator of a completed utterance and a start of a new utterance.

This paper is structured as follows. I will give an overview of past studies regarding the placement of certain constituents of German in section 1.2. In section 2, I discuss the data and methods for analyzing events of spontaneous speech in casual conversation from a German-language corpus followed by the data and methods for investigating the language in intermediate German textbooks. In section 3, I present the results and point out differences and similarities between beginnings of utterances in spontaneous speech and phrases and sentences in intermediate textbooks. Lastly, I conclude in section 4 with implications the data has for teaching and research.

1.2 Literature Review

Research on POS systems and their possible connection to word order is limited to a few researchers. Hengeveld (1992) evaluates nineteen languages for POS systems based on four categories of predicates which are verbal, nominal, adjectival, and adverbial. He categorizes languages into having either a flexible or rigid POS system based on the existence or nonexistence of these predicates in a language and finds that a language with all four predicates is considered to have a flexible POS system. On the contrary, a language lacking a class or classes of predicates is said to have a rigid POS system.

Hengeveld, Rijkhoff, & Siewierska (2004) expand on Hengeveld's (1992) work by examining fifty genetically, geographically, and typologically diverse languages. Similar to Hengeveld's (1992) methodology, Hengeveld et al. analyze the fifty languages based on four POS. The four POS that a language does or does not have follow a particular hierarchy. The hierarchy is as follows: the head of predicate phrase, head of referential phrase, modifier of referential phrase, and modifier of predicate phrase. These terms can be simplified into four syntactic lexemes: verb, noun, adjective, and manner adverb, respectively. The languages are assigned a numerical value, one through seven, depending on the existence or nonexistence of these lexemes. Languages assigned a one through three are considered to have a flexible POS system. A language assigned a four has a differentiated POS system. Languages assigned values between five and seven are categorized as having a rigid POS system.

The most significant finding from Hengeveld et al. through the comparison of fifty languages is that possible word orders in a language are partly restricted by the POS system of that language. In addition, they show that languages with unpredicted syntactic properties resolve

ambiguity with morphological features. For instance, German resolves such ambiguity by using case markings.

German is classified as a three and thus has a flexible POS system (Hengeveld & van Lier, 2010). German is considered a three because the same modifiers are used for verbs and nouns. Conversely, English belongs to the differentiated part of speech system and is a four on the continuum (Hengeveld et al., 2004). English is considered a four because different modifiers can be employed for nouns and verbs. Therefore, it seems plausible that English and German have various word orders due to different POS systems.

In response to Hengeveld et al., Vulanovic and Köhler (2009) propose a two-dimensional classification of POS system types that claims it is simpler and more accurate. Vulanovic and Köhler evaluate the same set of languages that Hengeveld et al. analyzed for POS system types and other aspects of each language. They examine whether a rigid word order is used or grammatical markers are employed for clarifying between the subject and the predicate, the head and the modifier of the referential noun phrase, and the head and the modifier of the predicate phrase (p. 290). Vulanovic and Köhler agree with Hengeveld et al and find that a language's POS system is related to word order. More specifically, languages with a rigid POS system are found to have a less fixed word order and use fewer grammatical markers to disambiguate between the four propositional functions than languages with a flexible POS system.

Though some research exists on POS systems and word order, I was unable to find research pertaining specifically to word order patterns and POS in German. Due to Hengeveld et al. and Vulanovic and Köhlers' claims that morphological markings, such as the German language's case marking system, directly relate to POS systems and influence possible word

orders of a language, I will discuss research comparing German to English, and word order preferences of German with a few references to English.

As has been shown, English word order primarily determines the grammatical roles within a sentence. On the contrary, German has case markings that demonstrate who does what to whom (Jackson, 2007) and is more flexible than English with regard to the positioning of constituents within a sentence. Though German grammatically allows for a number of arrangements of constituents within a sentence, distinct preferences for how these constituents are organized exist, often due to particular thematic roles and discourse functions (Pechmann, Uszkoreit, Engelkamp, and Zerbst, 1996). Exhaustive research exists on the reasons for varied word orders in German and I will now discuss these preferences. Important to note is that most of the research does not analyze spontaneous speech but rather deals with data in isolated contexts.

Rosengren (1993) found that topicalization in German occurs to connect elements with a previous utterance, perhaps allowing discussion to flow smoother. Psycholinguists report similarly that preference exists to keep the topic first instead of making a topic shift. This topic first paradigm often occurs with highly salient constituents. Further complicating this paradigm is the common practice observed in spoken German called pronoun dropping. Dropped pronouns qualify as a highly salient constituent (Arnold, 2010; Bosch, Katz & Umbach 2007; Cowles, Walenski, & Kluender, 2007; Kaiser 2006). In German, pronoun dropping is often explainable by linguistic contextual factors. For instance, in actual speech prior to a dropped pronoun, the speaker may have already articulated the pronoun (e.g *ich* 'I'). For this reason, this shared understanding, the pronoun, may not need to be explicitly said again. Below examples 5 and 6 are similar, however in example 6, the pronoun is dropped.

5) Ich habe mir eine neue Jacke bei Walmart gekauft.
   I bought myself a new jacket at Walmart.

6) Habe mir eine neue Jacke bei Walmart gekauft.
   Bought myself a new jacket at Walmart.

Roder, Schicke, Stock, Heberer, Neville & Rösler (2000) find in an empirical study that native German speakers process sentences slower when they deviate from the canonical word order, which is Subject (S) – Verb (V) – Indirect Object (IO) – Direct Object (DO). For instance, the sentences DO – V – S – IO and DO – V – IO – S were processed the slowest. This proves that although German allows for many possible word order realizations, those deviating from the canonical order are cognitively more difficult to process.

Preferences for various word orders are proven to exist and have been discussed in the contexts of written language. However, spoken German differs in many ways from written language (Schwitalla, 2003). Deppermann (2006) notes that what is still missing is a comprehensive theory of grammar in spoken language, and Fox (2007) addresses the concept of emergent grammar in the fields of Conversation Analysis, Interactional Linguistics, and usage-based approaches to grammar. She reports that due to context grammar in conversation is heavily influenced and in many ways co-constructed in interaction.

With this in mind, Günthner (2007) analyzes word order in *die Sache ist* or 'the thing is' clauses. To this end, she analyzes 91 everyday interactions collected in 1986-2006 in various regions of Germany. She finds that the prescriptive notion of moving the finite verb to the end of an utterance containing the subordinate conjunction, *dass* 'that' is often not followed. Her reasoning is backed by discovering that the independent clause in *die Sache ist* clauses actually contains the most important information and thus loses independent clause word order. The data set comprises of events in which prescriptive word order patterns of subordinate clauses are

followed. Furthermore, the articulation of the single word, *dass* 'that' is sometimes left out and sometimes present. All in all, her findings demonstrate the inconsistency and flexibility of syntax in actual spoken German usage.

A study similar to my design was performed by Chuang (1993). He analyzed word frequency and POS in English textbooks used in high schools in Taiwan. Due to the different POS systems of the learners' first language, Mandarin, to the target language, English, one of the research questions he hoped to answer was, do students react to and perform differently on the four parts of speech: adjective, adverb, noun, and verb (p. 98) His methodology was simple; he first tabulated the frequencies of different POS in the textbooks. Then, the students were tested. The test results revealed that students performed lower on verbs than nouns, adjectives, and adverbs.

As far as I know, no research exists that compares utterance- and phrase-initial POS in German interaction and textbooks. Though Doherty (2005) finds that English and German sentences begin differently fifty percent of the time in written text, her data is not spontaneous speech nor is there an overt concentration on POS. The following research questions I pose here are thus in need of investigation:

- What parts of speech follow a stop in intermediate German textbooks and utterances in spontaneous conversational speech?

- Are the results similar or different between textbooks? If so, how are they different?

- Are the results similar or different between textbook and spontaneous conversational speech data? If so, how are they different?

2 Data and Methods

In order to compare phrase-initial parts of speech in textbooks and utterance-initial parts of speech in spontaneous spoken speech, I use a two-part methodology. The two parts include an analysis of German textbooks and an analysis of spoken German corpus data. The reasoning for a two-part methodology is due to the analysis of two separate data sets. In the first part of my methodology, I analyzed sentences and phrases in textbooks from three intermediate German textbooks. I searched the department and bookstore websites of twenty-seven universities in United States with a German PhD program and found that *Stationen: Ein Kursbuch für die Mittelstufe (2nd Ed.)* is the most frequently required intermediate textbook. The other two textbooks I analyze are *Denk Mal!* and *Anders gedacht* which are also required textbooks at many universities with large German programs.

Cengage Learning is the publisher of both *Stationen* and *Anders gedacht*. According to the authors of *Anders gedacht*, the textbook was designed to prepare students to communicate effectively through intellectually engaging content while learning about German cultures. Furthermore, the authors claim that *Anders gedacht* transforms German into a content course and allows students to build connections to other fields. This textbook makes no explicit mention of adhering to a communicative approach, though developing language skills to communicate effectively is a goal.

Similar to *Anders gedacht*, in *Stationen,* there is a preface to the instructor. There is little explicit attention given to developing students' conversational skills. Rather, it can be gleaned from the preface that the creators designed the textbook to bring students along on an adventure through different cities, or *Stationen*, in which authentic texts are presented and used to improve language skills and promote critical thinking skills. However, *Stationen* and *Anders gedacht* both

deliberately introduce *Redemittel*, or useful phrases for discussion, that demonstrate the authors'

desire for students to improve conversational skills.

Denk Mal! has a preface similar to that of *Stationen* and *Anders gedacht*, which reads,

"*Denk Mal!* takes an interactive, communicative approach. It focuses on real communication in

meaningful contexts to develop and consolidate students' speaking, listening, reading, and

writing skills." Therefore, we can draw the conclusion that developing students' communicative

competence is one of the authors' main goals.

In the textbook analysis, I first tabulated the first two words' POS in phrases and

sentences found in the first and third chapters of the above mentioned textbooks. Words were

categorized into various POS as outlined in Schiller, Teufel, Stöckert, & Thielen (1995). I

selected this work because words are classified in the same manner in the analysis of the

spontaneous speech corpus data. The different parts of speech in this study are: adverb (e.g. *dann*

'dann'), adjective (e.g. *blau* 'blue'), article (e.g. *die* 'the'), conjunction (e.g. *aber* 'but'*)*,

interjection (e.g. *ähm* 'um'), name (e.g. Richard), noun (Fußball 'soccer'), number (e.g. 1980),

particle (e.g. *mit* 'along'*,* i.e. *Ich komme mit* 'I am coming along'), preposition (e.g. *mit* 'with'*,*

i.e *Ich fahre mit dem Bus* 'I am driving with the bus.'*)*, pronoun (e.g. *sie* 'she'), and verb (e.g.

schwimmen 'to swim'). I went to great lengths to assure that the analysis of data from the

textbooks was consistent with the analysis of the corpus data.

In the second part of my methodology, I analyze spontaneous speech events from a

corpus called the *Forschungs- und Lehrkorpus Gesprochenes Deutsch* (FOLK) 'Research and

Teaching Corpus of Spoken German'. The data I analyzed taken from the FOLK corpus

comprise 10 events with a total audio duration of two hours fifty-seven minutes and fifty-six

seconds. Each event is categorized as *Alltagsgespräch* 'everyday speech/small talk'. To process

the various two-word POS combinations, koRpus was used so that TreeTagger was able to get POS tags on words. The R package koRpus was developed by Michalke (2015) to calculate similarities and differences among texts. Over time, koRpus has developed into software that individuals can use to conduct research on multiple features of texts such as readability and lexical diversity. TreeTagger is software developed by Schmid (1994). Words were determined for POS according to the same guidelines in the textbook analysis, in order to compare the two data sets appropriately. In the current study, only the first two words following a full stop were analyzed to find the various two-word POS combinations.

2.1 Textbook Analysis

Textbook materials often form the backbone of a curriculum and may be quite central to the planning and development of introductory and intermediate foreign language courses. In order to verify whether phrase- and sentence-initial POS in intermediate German textbooks are similar to utterance-initial POS in spontaneous speech, I examined three intermediate German textbooks.

I selected intermediate German textbooks because intermediate learners have been exposed to many POS, especially in comparison to beginning learners of a language. Also influencing my selection of intermediate textbooks is the fact that the main approach to the teaching of foreign languages at an overwhelming majority of higher education institutions across the United States is the communicative language teaching approach. The communicative approach does not consist of a single method. Rather, developing communicative competence is the main goal and emphasizes meaningful and purposeful interactions through language (Richards, 2006). It is for these reasons that one may assume that intermediate textbooks contain input that approaches the kind of language one might encounter in contexts of use.

In the analysis of the textbooks, words were categorized into the following parts of speech: adverb, adjective, article, conjunction, interjection, name, noun, number, particle, preposition, pronoun, or verb. Example 7 from *Stationen* below contains two sentences.

(7)   Damals beobachtete der Berliner Kurt Heuwer, wie die Amerikaner Steak mit Ketchup aßen. Er erzählte seiner Frau Herta.
The Berliner Kurt Heuwer observed back then how Americans ate steak with ketchup. He explained to his wife Herta.

The first sentence would be analyzed as *adverb + verb*. The next sentence would be analyzed as *pronoun + verb*.

To determine the input to be analyzed, I decided to include text following a period, question mark, hyphen or colon. In written language, a period, question mark, hyphen, and colon are often interpreted as a stop. Thus, the words following those orthographic punctuation marks start the beginning of a new sentence or phrase. In addition, I analyzed the first two words' POS in titles of activities or exercises (e.g. *Eine berühmte Berlinerin* 'a famous female person from Berlin', *richtig oder falsch* 'true or false', *Fragen zum Nachdenken* 'questions to contemplate') and other phrases found such as directions for an activity in the textbooks (e.g. *sich erschrecke*n 'to be frightened'). My rationale for including titles in the textbook data is due to two factors. First, I want the textbook data to represent all input the reader encounters. Second, I want my analysis of the textbook data to be as consistent as possible with my analysis of the FOLK corpus data. I did not omit one word utterances from the FOLK corpus data as that belongs to that type of speech. Similarly, I did not omit titles from the textbook data as these belong to that genre of text. Figure 3 below shows how the title would be tabulated as *noun + preposition*. In addition, Figure 3 shows that the first and second sentences in the directions under the title would be tabulated as *verb + pronoun*.

Figure 1. Phase-Initial POS in *Stationen*
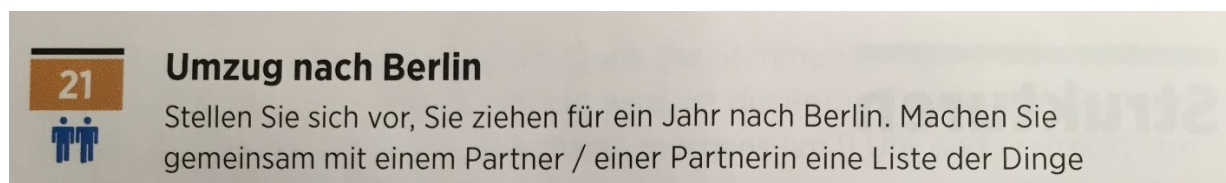
I did choose to omit text written in English, which often only appeared to explain complex grammatical features, titles located in corners of pages, which were often highly repetitive, and vocabulary lists with one-to-one translations.

2.2 FOLK Corpus Analysis

As mentioned, the corpus consulted in the study is called FOLK. The FOLK corpus is one of 21 corpora made available by the *Institut für die Deutsche Sprache* 'Institute for the German Language'. The FOLK corpus contains naturally-occurring conversational data that can be used to analyze natural spoken speech (Deppermann & Schmidt, 2014).

I limited the data I analyzed to 10 events from FOLK, because Golato (2003) also limited her data. She investigated compliment responses in natural spoken speech in German and did not mix types of talk, for example face-to-face conversations vs. oral presentations. Her data came from face-to-face or phone conversations during activities that friends would normally engage in (e.g. meals, social gatherings, etc). Levinson (1992) suggests that interactions differ between institutional and ordinary conversation settings and it is for these reasons that I limit the data and exclude other events from the FOLK corpus which are classified as *Institutionelle Kommunication* 'institutional communication', *Medien- bzw. Öffentliche Kommunkation* 'media and public communication', *Experimentsituation bzw.Kommunicationsspiel* 'experimental situation and communication game', *Sprachbiografisches Interview* 'spoken biographical interview'. The 10 events that I analyzed are classified as *Alltagsgespräch* 'everyday speech/small talk', which belong to spontaneous speech (Deppermann & Schmidt, 2014). The events that I analyzed were recorded in different regions in Germany, primarily the Ripuarian and north-lower German linguistic regions. A description of the FOLK corpus events analyzed in the current study can be found in Appendix A.

In order to make claims about conversational speech, it is important to have subjects representing multiple generations. The 10 events comprise 24 speakers (12 male and 12 female). The oldest speaker was born in 1959 and the youngest in 1996. The occupations of the speakers

are also very diverse, for example, the participants' occupations include: radio host, journalist, full time student, engineer, author, and teacher. The number and gender of speakers between events also changes. For example, event 1 consists of four male friends in a band, whereas event 8 is between three friends (one male and two female). Finally, event 10 is between a husband and wife.

As a part of the analysis of the beginning two words of utterances, the 10 events from the FOLK corpus were tagged for POS information. I used TreeTagger as it has achieved an accuracy of 97.53% for POS information (Schmid, 1995). However, while investigating the accuracy of three events from FOLK using TreeTagger, Westpfahl and Schmidt (2013) found the error rates to be 18.79, 18.62, and 19.5 percent. To systematically check the tagging software for accuracy; I thus spot checked event 4, every tenth word. The error rate was similar to Westpfahl and Schmidt´s findings at 19.75%. Although these error rates are higher than desired, I am still able to discuss trends.

After every full stop, the various two-word POS combinations were processed. Example 8 on the next page shows the output from event 4 generated by TreeTagger which has been simplified to include information pertinent to understanding the process. In example 8, the first column titled "token" shows the utterance the way it was articulated in the event. A period is used to indicate a fullstop. The column titled "lemma" shows the lemma of the word from the event. A lemma is an abstract concept that subsumes all possible lexical variations according to that word's abstract representation (Crystal, 2008). For instance, in English the lemma *want* subsumes *wants*, *wanted*, and *wanting*. The next column provides a gloss of the lemma. Lastly, the column titled "word class" shows the POS assigned to the word. Example 8 below shows the output of koRpus and thus the word order of the utterance, *verb + pronoun*.

15

(8) Utterance-initial POS in FOLK Data

token	lemma	Gloss	word class
.		.	*fullstop*
willst	wollen	to want	*verb*
du	du	You	*pronoun*
was	was	something	pronoun
trinken	trinken	to drink	verb
.		.	fullstop

3 Results

3.1 Textbook Analysis

3.1.1 *Stationen*

In the first sentence of the "preface to the instructor" in *Stationen*, the authors emphasize authentic readings in the textbook; nonetheless, there is no overt mention of improving communicative competence. This is important to keep in mind, because the texts found in *Stationen* do not seem to promote conversational skills and prioritize teaching cultural information over developing communicative competency. In fact, they often feature a city and famous inhabitant of that city which are referred to in exercises for the remainder of the chapter. This can be exemplified by the first reading in Chapter 1. Although the text is an authentic reading and aligns with the authors' aims, the text is not designed for improving conversational skills. It may be for these reasons 11.0% of sentences and phrases begin with a name (e.g. *Chris*) in *Stationen*. All in all, sentences beginning with a proper name coincide with the authors' intent of exposing students to locations and people in the German-speaking world.

The authors of *Stationen* do however expose students to unique linguistic features. In the first chapter of *Stationen*, POS are explicitly covered. The reader is given an overview, purpose, and direct connection of POS to German and English. Then, the authors introduce the different POS, which are clearly defined. In addition, every POS is accompanied with contextualized examples.

3.1.2 *Anders Gedacht*

The analysis of *Anders gedacht* revealed that 18.9% of sentences and phrases begin with a noun. For instance, *Modalverben brauchen einen Infinitiv ohne zu* 'Modal verbs need an infinitive without to' (p. 108). Though sentences and phrases beginning with a noun do not make

up the largest POS frequency, this is certainly a uniquely high figure. Article-initial sentences and phrases are similar to noun-initial sentences and phrases. These POS initial sentences and phrases are related due to the nature of articles, as a noun almost always follows an article. The analysis too reveals that 13.4% of all sentences and phrases begin with an article. Therefore, it can be gleaned from the data that a large majority of sentences and phrases begin with a noun phrase.

3.1.3 *Denk Mal!*

Pronoun-initial sentences and phrases are the most common in *Denk Mal!*. The second most common is verb-initial utterances which account for 18.7% of utterances. The remaining POS initial-utterances each account for less than 15% of the data. Another interesting finding is that in the first chapter of *Denk Mal!*, word order in statements is explicitly addressed. In fact, the concept of topicalization is mentioned. The term itself is not included, but the concept of emphasizing information in the first position is explicitly addressed with examples.

3.1.4 Comparison across Textbooks

The analysis revealed stark differences between the first words' POS in sentences and phrases among textbooks. For instance, sentences beginning with a name (e.g. *Chris*) demonstrate major differences in the data among the three textbooks. In *Stationen*, 11.0% of sentences and phrases are name-initial, whereas in *Denk Mal!,* sentences and phrases are name-initial 6.9%, and the frequency is even lower in *Anders gedacht* with 3.4%. Nonetheless, the analysis of the three textbooks reveals additional differences and some similarities of the first two words' POS in sentences and phrases. To begin, I discuss frequency of sentences and phrases based on word initial POS. Figure 2 below shows the percentage of sentences beginning with the eight POS most frequently found in each of the textbooks and then from the textbooks

as one entity. Sentences and phrases beginning with a number, particle, or interjection have been omitted from Figure 2 because each one make up less than 1.4 percent.

Figure 2. Frequency of Phrase-Initial Parts of Speech
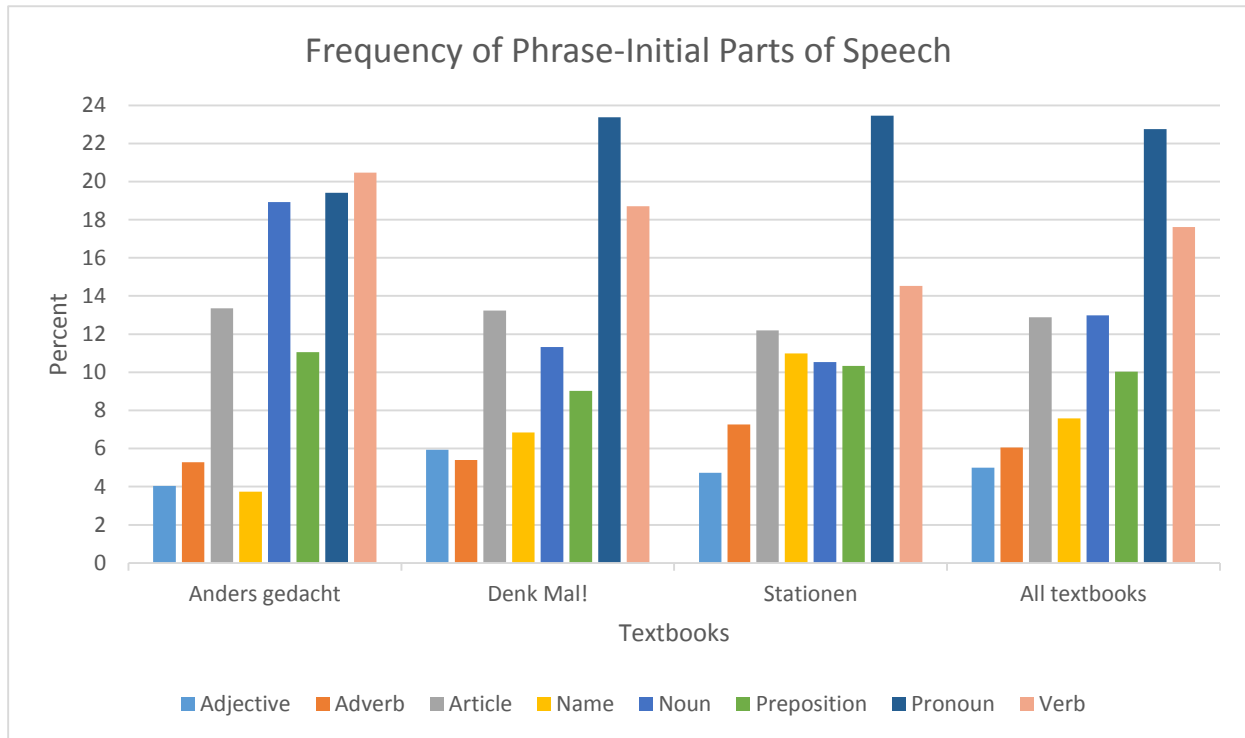
Interestingly the analysis revealed some similarities in sentence and phrase initial POS. The similarities across the textbooks concern two parts of speech, articles and prepositions. In *Denk Mal!*, 9.0% of sentences and phrases are preposition-initial, whereas in *Stationen* and *Anders gedacht* the frequency is slightly higher at 10.3% and 11.0%, respectively. This finding is rather consistent between the textbooks. The next similarity across textbooks is sentences and phrases that begin with an article. In fact, the margin of difference is extremely low. In *Stationen*, 12.2% of sentences and phrases begin with an article. Even more closely related are sentences and phrases beginning with an article in *Anders gedacht* and in *Denk Mal!* which is 13.4% and 13.2% respectively.

Across all of the textbooks, sentences and phrases begin more frequently with a pronoun than any other POS. In fact, 22.8% of sentences and phrases across all textbooks begin with a pronoun. It is important to remember the many different kinds of words which are classified as pronouns. Examples 9-13 below demonstrate that interrogative pronouns, reflexive pronouns, and personal pronouns are tagged in the textbook data and FOLK corpus.

9) *Was ist der Mittelpunkt des Bildes*? (*Anders gedacht* p.46)
'What is the focus of the picture?'

10) *sich erschrecken*
'to scare oneself'

11) *Sie wartete nur darauf, dass kleine Kinder zu ihrem Kuchenhäuschen kamen.* (*Anders gedacht* p. 56)
'She merely waited for small children to come to her small cake house.'

12) *Wen möchte sie mit diesem Lied ansprechen*? (*Anders gedacht* p. 112)
'Whom would she like to appeal to with this song?'

13) *Mir gefällt …* (*Anders gedacht* p. 31)
'I like …'

Given that examples 11-13 also demonstrate different case markings (e.g. nominative, accusative, and dative, respectively), it is clear that flexibility of German word order is due to case as Walter and van Compernolle (forthcoming) show, and since the tagging software was not set up to tag for case, this study cannot make claims about the frequency of varied word orders in terms of case.

3.2 FOLK Analysis

3.2.1 FOLK

The first general finding from the events in the FOLK corpus reveal that single word utterances are the most common. In fact, 28.8% of all utterances from the FOLK data are made up of a single word. The first most frequent single-word utterance is a noun; the second most frequent is an adverb. Additionally, other single-word utterances appear quite often, including adjectives, interjections, and verbs. These findings demonstrate the high frequency of single-word utterances in spontaneous speech.

Single-word utterances are very frequent, and I will now expand on these findings. In the FOLK data, 12.1% of utterances are single-word noun-utterances, which is an extremely high frequency considering the various two POS combinations that could appear. Before analyzing a single-word noun-utterance found in the FOLK data, it is important to recognize four key factors that make every utterance unique and greatly impact the meaning of an utterance. These four factors are time, place, volume, and speaker of the utterance. In event 2, a single-word noun-utterance appears and is seen below in example 14.

(14) Example from FOLK Corpus

Speaker	Deutsch	English
OM	welche war _s	which was 'it
TV	Herbststurm	Herbststurm
OM	ach so	I see
OM	wollt ihr vielleicht aufhören oder	Do you all want to maybe quit that

In the same way that utterances have unique factors, contextual factors also contribute to an utterance's meaning. Contextual factors can be categorized into three factors. Linguistic contextual factors include that which comes prior to the utterance. For instance, in example 14,

OM asks TV a question regarding a band. The next contextual factor is situational which allows individuals communicating to make references to objects found in the vicinity without explicitly articulating them. In this case, the speakers are at a venue setting up for a band performance. Lastly, social contextual factors which deals with the relationships between speakers. Speakers OM and TV in example 14 are male friends in the band.

This example illustrates an interesting finding in the data. It appears that in the context of spontaneous speech, a single-word utterance is an appropriate response. This is due to the major role contextual factors have in discourse which are also unique to conversation. It is considered an appropriate response because OM acknowledges TV's single-word answer and takes another turn by addressing the other members in the band to quit making a racket. This is interesting because it sheds light on a common prescriptively taught notion that students should answer questions with complete sentences.

Next, I will discuss two interesting aspects of spoken German found in the FOLK corpus data. It is important to discuss these aspects because these two POS frequencies appear quite often in spontaneous speech. First, the pronoun may sometimes be dropped as seen in example 15 on the next page from event 9. In JN's fourth turn in example 15, there is no pronoun. Rather, the speaker, JN, only articulates the conjugated verb, *hat* 'has' which is the third person singular conjugated form of *haben* 'to have'. This form, *hat*, agrees to the dropped pronoun, *sie* 'she'. Once again, it is important to consider the three contextual factors when analyzing conversational data. One reason the pronoun is dropped may be due to linguistic contextual factors. In JN's previous utterance, she uses a demonstrative pronoun, *die* 'she.' Thus, though the pronoun may not be clearly articulated in her fourth turn, a shared understanding probably exists between the two speakers which doesn't require JN to articulate the pronoun again.

In addition to pronoun dropping, example 15 has an example of an article which is used to refer to a person. *Die* 'the', is the subject. In this case the girl or she with magnificent head of red curls. This can be demonstrated with the italicized word in JN's third turn. The article, *die* 'the'*,* can also be used to replace *sie* or 'they' in 3rd person plural.

(15) Event 9 from FOLK Corpus

Speaker	Deutsch	English
JN	[die ro]te lockenpracht	the red magnificent head of curls
SG	j[a]	yeah
JN	[a]ber ähm die zum beispiel die ham wer gesehen beim beim theaterfestival	but uh it for example we saw it at the at the theater festival
JN	da hat *die* so ne inprovisierte völlig überdrehte ähm (.) pippi langstrumpf gespielt	she played such an improvised completely overexcited uh (.) pippi longstocking
MD	ach so ja h°	oh yea h°
JN	un hat aber ganz viel von dem was sie jetz hier in ihre rolle geworfen hat auch bei der (.) pippi langstrumpf [gehabt]	and she however added here a real lot to the role from that which she did with (.) pippi longstockings
MD	[ha ha] ha [ha]	[ha ha] ha [ha]
JN	[un] des is so witzig des dann wiederzusehn	[and] that is so funny to then see that again

The two findings, the use of a demonstrative pronoun and evidence of a possible pronoun dropping, are worth mentioning because utterances with this construction are likely to not appear in textbooks.

3.2.2   Comparison of Textbook Results to FOLK Corpus Results

The data from the three textbooks revealed drastic differences compared to data from the FOLK corpus. Figure 3 compares the frequencies of utterances based on the initial word's POS between data from the FOLK corpus and the textbooks. The initial-words' POS is measured along the x-axis across eight different variables: adjective, adverb, article, name, noun,

preposition, pronoun and verb. The bar graph presents the frequencies of these variables in percentage, along the y-axis.

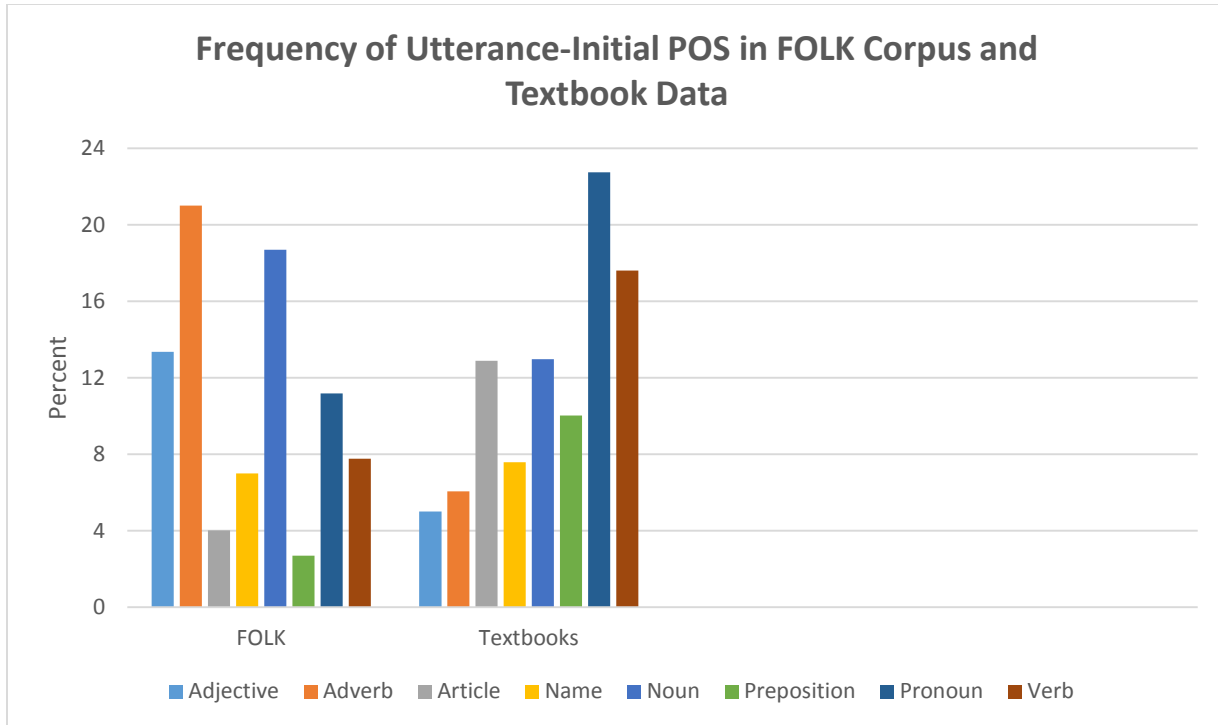**Figure 3. Frequency of Utterance-Initial POS**

Figure 3 demonstrates that 21% of utterances begin with an adverb in the FOLK data, whereas only 6.1% of phrases and sentences in the textbooks are adverb-initial. It is evident from the data that more utterances begin with adverbs than any other POS in spontaneous speech. Figure 3 also demonstrates that 10.0% of phrases in textbooks are preposition-initial, whereas only 2.7% of utterances in FOLK are preposition-initial. Also interesting to note are verb-initial phrases. In the textbooks, 17.6% of phrases are verb-initial, whereas only 7.8% of utterances are verb initial in FOLK.

Example 16 on the following page is a segment from FOLK containing an adverb-initial utterance. Once again it is crucial to recognize the contextual factors. In example 16, two friends

are discussing Spotify, an application that can be used with a cell phone, tablet, computer, or similar electronic device that allows its users to listen to millions of songs.

(16) Adverb-initial Utterance from FOLK Corpus

Speaker	Deutsch	English
GB	die neue band bei (.) von emi	the new band by (.) from emi
XM	h°	h°
GB	Hier	Here
GB	weißte (.) leg drauf	you know (.) put it on
GB	dann machst du das °h aber wenn (.) irgendwo (.) bei spotify wo auch im mer du dir die songs danndownloade n kannst für (.) kohle	then you do that °h but whenever (.) wherever (.) with Spotify or wherever you can download songs for (.) cash

Adverb-initial sentences and phrases are far less frequent in the textbook data and comprise only 6.1% of phrases and sentences. Interestingly, the 14.9% margin is the largest difference between the FOLK and textbook data on the basis of constructions' first word's POS. Therefore, textbooks are not presenting L1 English learners of German with phrases and sentences that begin with adverbs as often as they appear in spontaneous speech. This may be due to the different modes of communication. In conversation, the placement of adverbs in German may be different than in written communication. This is especially concerning because the placement of adverbs in English is subject to different patterns. López (1995) exemplifies that comparing the adverbs *well* and *carefully* as seen in example 17 below:

17) a. *Carefully*, he put away the dishes.
       He *carefully* put away the dishes.
       He put away the dishes *carefully*.
    b. * *Well*, the professor speaks.
        * The professor *well* speaks.
         The professor speaks *well*.

For these reasons, textbook authors may want to select input that considers how adverb placement in German occurs in spontaneous speech.

Among the six chapters of the textbook data, preposition-initial sentences and phrases make up 10.0% of the data, whereas preposition-initial utterances make up only 2.7%. Example 18 below is a sentence from *Anders gedacht* (p. 40).

18) Auf Bild 1 sieht man Johann Wolfgang von Goethe.
    On page 1, one sees Johann Wolfgang von Goethe.

Preposition-initial phrases may appear more common in the textbook data due to the genre of textbooks. Textbooks often prompt the reader with some sort of authentic artifact. This is the case with example 18; the artifact is a picture of the well-known German author Johann Wolfgang von Goethe. Following an artifact often appears a question or statement in relation to the artifact as seen above in example 18.

A relatively low percentage of utterances beginning with a verb were found in the FOLK data, 7.8% compared to the textbooks' data, 17.6%. In German, there are multiple ways to form questions. The two most common constructions involve the use of an interrogative pronoun (e.g. *wer, wen 'who, whom'*) or verb-subject inversion (e.g. *Spielen Sie Fußball? vs. Ich spiele Fußball.* 'Do you place soccer. vs. I play soccer'). However, in the FOLK corpus data, questions also seem to be formed in what might prescriptively be considered incomplete chunks. This is exemplified in example 19 below:

(19) Segment 2 Of Event 9 From FOLK Corpus

Speaker	Deutsch	English
SG	h° tessa zu hause (.) oder	h° Tessa at home (.) right
SG	[is sie au]ch [weg]	[is she al]so away
MD	[bidde]	excuse me
JN	[hoffe]	I hope
MD	tessa ist	tessa is
MD	zu hause	at home
MD	ja	yea

In example 19, the speaker SG tags her utterance with *oder* 'or' and uses rising intonation to indicate that her utterance is a question. Interestingly, SG takes another turn in the conversation and poses another question, this time a Y/N question. JN chimes in demonstrating her understanding by articulating *hoffe* 'I hope'. Thus it appears that both ways of question forming are appropriate in spontaneous speech, which is, once again due to contextual factors.

There is a possibility that the phrase might appear in a textbook as, *Ist Tessa zu Hause?* 'Is Tessa at home?' To compare, in the chapters of the textbooks I analyzed, only one tag question appeared. However, German speakers use such constructions a great deal in FOLK. This may be problematic for students who encounter German speakers who form questions in ways that they are not very familiar. Thus, I believe that textbooks including more of these types of utterances as well as activities in which students can participate in the meaningful use of these constructions will better prepare students to participate in contexts of use.

In section 3.2.1, the FOLK analysis, I introduced the high frequency of single word utterances. However, single word utterances are far less common in the textbook data, accounting for only 9.4% of all sentences and phrases, which is primarily due to titles of exercises (e.g. *Fragen* 'questions' from Denk Mal! p. 26). There is a 19.4% difference in the appearance of single-word constructions between the FOLK and textbook data. Therefore, textbooks are not presenting L1 English learners of German with single-word utterances as often as they actually appear in spontaneous speech. This is concerning as it appears to be an appropriate response in spontaneous speech due to the previously discussed three contextual factors: linguistic, situational, and social.

I found a similarity across the two data sets worth drawing attention to. The similarity concerns the construction *pronoun + verb*, which is by and large the most frequent two word

POS sequence in the textbook data and the third most frequent two word POS sequence in the

FOLK data. The three intermediate textbooks present L1 English learners of German with

phrases and sentences that begin with *pronoun + verb* in a way that is similar to their appearance

in spontaneous speech. This is interesting when one considers how subject and verb agreement is

quite different in German than English. Since this is an important aspect that appears in both

modes of language, it is worthy of attention in German classrooms.

In one particular group of utterances, a particular POS was omitted entirely. The

interrogative article, *welch* 'which/what', is used to form a question. Nonetheless, a noun

following the interrogative is absent in the FOLK data. This highlights another difference

between the data from the textbooks and spontaneous speech. In the textbooks, questions are

often posed with *welch-* 'which/what,' however, a noun almost always follows this interrogative

in the textbooks. For instance, in *Stationen* (p. 56), *Welche Farben verwendet der Künstler?*

'What colors does the artist use?' In example 14 on page 21, the speaker, OM, does not articulate

the noun. He simply says, *welche war _s* 'which was it'. Interestingly, the other speaker answers

OM's question. Thus suggesting that OM's utterance is appropriate. This is due to linguistic

contextual factors. The speakers were previously conversing about music. Thus, in addition to

the social contextual factors that the speakers are friends, it appears that articulation of the noun

following the interrogative, *welch-* 'which/what,' is not necessary in this context of spontaneous

speech.

4 Discussion

The textbook analysis revealed similarities and differences of the first two words' POS in sentences and phrases between the three textbooks. I then compared the textbook results, as one entity, to the results of the FOLK corpus analysis. The FOLK corpus analysis showed stark differences in the first two words' POS compared to the textbooks. In the rest of this section, I answer my research questions, note the limitations of the study and state implications for both pedagogy and future research.

The FOLK corpus analysis shows that in spontaneous conversational speech, 28.8% of utterances are made up of a single word, especially significant are nouns. This phenomenon suggests that single-word noun-utterances are appropriate in this context. This result may be due to the particular contextual factors in spontaneous speech. In conversation, linguistic, situational, and social factors greatly impact the meaning and formation of utterances. Given the social contextual factors, the speakers in all of the events are friends and furthermore, the events can be classified as everyday speech or small talk. Due to nature of conversation, speakers often make utterances that relate back to what was previously communicated. This particular linguistic contextual factor influences how utterances are formed which is noted in examples 14, 15, 16, and 19. Thus different POS combinations appear due to the different contexts.

Two additional interesting findings resulted from a closer examination as I analyzed events line for line. Evidence for possible pronoun dropping is present in the events. This result also suggests that pronoun dropping is acceptable, and perhaps preferred or expected due to the three previously mentioned contextual factors. Other possible reasons speakers drop pronouns may be because an individual would like to belong to a speech community. For example, if an individual wants to fit into a group he may begin communicating like other members of the group such as

dropping pronouns. The second interesting finding is related to the first and deals with articles which are used to refer to people or things. This may also be due to contextual factors. Perhaps it is common to use a demonstrative pronoun (e.g. *die* 'the' in *die ist ganz lieb* 'she is really kind') for subjects of objects in the vicinity, thus emphasizing the importance of situational context. Another reason may be due to social contexts. Perhaps the use of demonstrative pronouns are used more often among persons who are familiar with one another.

Another focus of the study was to analyze textbooks and spontaneous speech to see whether the results are similar or different. I compared the findings reported for both textbooks and spontaneous conversational speech. Figure 3 (p. 24) shows differences between the two data sets based on the constructions' initial POS. Interestingly, 21% of utterances in the FOLK events begin with an adverb. This elevated result may be due to the nature of topics discussed in the events. Seven events were recorded during an intermission of a theatrical performance; the speakers often conversed about the performance chronologically. For this reason, the use of linking adverbs (e.g. *danach* 'afterwards/ subsequently') may contribute to this finding. This finding also suggests that adverbs appear very frequently in interactive contexts of use, since emphasis may be placed on manner adverbs due to objects in the vicinity, i.e. there may be situational contextual factors that no longer need to be articulated because of a shared understanding.

A result that differs between the textbook and FOLK corpus analysis regards constructions beginning with a preposition. In the textbooks, every tenth sentence or phrase begins with a preposition, whereas every fiftieth utterance begins with a preposition in the FOLK data. Comparing these results suggests that there are different ways that people use language in spoken contexts of use, specifically everyday speech among friends, as compared to the way language is

presented in written contexts, specifically in the genre of textbooks. This may also be due to formal and informal contexts.

Another major difference between the textbooks and FOLK data is the percentage of constructions beginning with a verb. Under 8% of utterances begin with a verb in the FOLK data, whereas almost 18% of sentences or phrases in the textbook data begin with a verb. This result may be due to prosodic features such as intonation, tone, stress, and rhythm unique to spoken speech in German. These suprasegmental phonological processes signal to the interlocutor that the utterance is actually a question. This finding suggests that prosodic features are very important to question formation. Utterances, which are often small fragments, as in example 19 (p. 26), are often tagged with words such as *ne* 'right,' *oder* 'correct,' *nicht wahr* 'ya know' that transform utterances into questions. These words in German can be translated into English words *right* or *correct* with rising intonation at the end of an utterance. This finding suggests that these tagging words may significantly impact how one elicits information. An additional interpretation of this finding may be due to the type of questions posed in the two different data sets. This finding may also suggest that eliciting the same information looks different in various modes of communication.

In dispute is the finding that utterances in the FOLK data and sentences and phrases in the textbooks frequently begin with the two POS combination *pronoun + verb. Pronoun + verb* is the most frequent two POS combination present in the textbook data and third most frequent in the FOLK data. This finding may be due to a limitation in my analysis. Many classes of words were classified as a pronoun as seen in examples 9-13 (p. 20). This methodology therefore did not allow me to distinguish between different cases (e.g. nominative, accusative, dative, and genitive).

I have noted the high frequency of single word utterances, adverb-initial utterances, pronoun dropping, demonstrative pronouns, and tag question formation that appear in the FOLK data. These findings suggest that these differences belong to spontaneous speech contexts for various reasons which have lead me to identify areas of future research and implications for teaching.

The finding that 21% of utterances begin with an adverb is startling. It would be fascinating to examine adverb placement in utterances of different speech contexts. This is fascinating because the placement may have subtle effects on meaning. In addition, it would be interesting to examine how the data from my analysis of intermediate textbooks compares to beginner textbooks or even advanced textbooks.

The findings also lead me to identify the following pedagogical implications. Language educators may also find it conducive for some students to have explicit instruction about the concept of demonstrative pronouns, tag words, or pronoun dropping. I suggest that students be given the time to actually start using these forms in meaningful ways in consideration of contextual factors.

Language educators may also find it beneficial for students to be exposed to input containing meaningful authentic conversations with single-word utterances, dropped pronouns, and tag questions, since they are appropriate in spontaneous speech. However, bringing meaningful input into the classroom with the above phenomena is not enough. Equally important is developing students' awareness of the three contextual factors: linguistic, situational, and social. Students must understand the complexity of these relationships and how they impact speech in use. This will prepare students for appropriately communicating in German interactions. In general, the results of this study show just how important it is to bring students to

observe, analyze, and evaluate these contextual factors so that they might be better prepared to

use the language in a wider variety of contexts.

References

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass, 4*(4), 187-203.

Augustyn, P. & Euba., N. (2011). *Stationen: Ein Kursbuch für die Mittelstufe (2ⁿᵈ Ed.)*. Boston: Heinle Cengage Learning.

Barske, T., McKinstry, M., Schestokat., K., & Sokolosky, J. (2012). *Denk Mal! Deutsch ohne Grenzen (2ⁿᵈ Ed.).* Boston: Vista Higher Learning, Inc.

Bosch, P., Katz, G., & Umbach, C. (2007). The non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten, & M. Knees (Eds.), *Anaphora in text: Cognitive, formal and applied approaches to anaphoric reference*. Amersterdam: John Benjamins.

Chuang, Y. (1993). *A Quantitative Corpus Analysis of Word Frequency And Part of Speech in the English Textbooks Used in Senior High Schools in Taiwan* (Doctoral dissertation.) Retrieved from Linguistics and Language Behavior Abstracts. (Accession No. 9407084)

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*, 26, 3-18.

Crystal, D. (2008). *Dictionary of Linguistics and Phonetics* (6ᵗʰ Ed.). Oxford: Blackwell.

Deppermann, A. (2006). Construction Grammar – Eine Grammatik für die Interaktion?. In: Deppermann, A., Fiehler, R., Spranz-Fogasy, T., (Eds.), *Grammatik in Interaction. Untersuchungen zum Zusammenhang von grammatischen Strukturen und Gesprächsprozessen*. (pp. 43-65). Radolfzell: Verlag für Gesprächsforschung.

Deppermann, A., & Schmidt, T. (2014). Gesprächsdatenbanken als methodisches Instrument der Interaktionalen Linguistik - Eine exemplarische Untersuchung auf Basis des Korpus

FOLK in der Datenbank für Gesprochenes Deutsch (DGD2). In: Domke, Christine & Gansel, Christa (Eds.), *Korpora in der Linguistik - Perspektiven und Positionen zu Daten und Datenerhebung. Mitteilungen des Deutschen Germanistenverbandes* 1, 4-17.

Doherty, M. (2005). Topic-Worthiness in German and English. *Linguistics, 43*(1), 181-206.

Engelkamp, J., Pechmann, T., Uszkoreit, H., & Zerbst, D. (1996). Wortstellung im deutschen Mittelfeld. Linguistische Theorie und psycholinguistische Evidenz. In Habel., C., Kanngießer, S. & Rickheit, G. (Eds.), *Perspektiven der Kognitiven Linguistik, Modelle und Methoden.* Opladen: Westdeutscher Verlag.

Fox, B. A. (2007). Principles shaping grammatical practices: An exploration. *Discourse Studies, 9*(3), 299-318.

Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics, 24*(1), 90-121.

Günthner, S. (2008). Projektorkonstruktionen im Gespräch: Pseudoclefts, die Sache ist-Konstruktionen und Extrapositionen mit es. In *Gesprächsforschung- Online-Zeitung zur verbalen Interaktion*, 86-144.

Hengeveld, K. (1992). Parts of Speech. In Layered Structure And Reference In Fortescue, M, Harder, P & Kristoffersen, L (Eds.), *A Functional Perspective: Papers From The Functional Grammar Conference, Copenhagen, 1990* (pp. 29-55). Amsterdam, Netherlands: John Bejamins B.V.

Hengeveld, K., Rijkhoff, J., & Siewierska, A. (2004). Parts-of-speech systems and word order. *Journal of Linguistics, 40*(3), 527-570.

Hengeveld, K., & Lier, E. v. (2010). Connectivity in implicational maps. authors' reply to 'the
added value of the connectivity hypothesis for the map of parts of speech' by caterina
mauri (2010). *Linguistic Discovery, 8*(1), 160-161.

Jackson, C. N. (2007). The use and non-use of semantic information, word order, and case
markings during comprehension by L2 learners of German. *The Modern Language
Journal, 91*(3), 418-432.

Jacobs, J. (2001). The dimensions of topic-comment. *Linguistics, 39*(4), 641-681.

Levinson, S. C. (1992). Activity types and language In P. Drew and J. Heritage (Eds.). *Talk at
Work. Interactions in Institutional Settings*. Cambridge: Cambridge University Press (pp.
265-399.

López, A. I. O. (1995). The Distribution of Adverbial Phrases in English. *Atlantis*, *17*(1/2), 181–
206.

Kaiser, E. (2006). Effects of topic and focus on salience. In C. Ebert & C. Endriss (Eds.).
Proceedings of the Sinn und Bedeutung 10, ZAS working papers in linguistics (Vol. 44,
pp. 139-154). Berlin: ZAS.

Mihalicek, V. & Wilson, C. (2011). *Language Files: Materials for an Introduction to Language
and Linguistics* (11th Ed.). Columbus, OH: Ohio State University Press.

Michalke, M. (2015). Using the koRpus Package for Text Analysis. Retrieved from
https://cran.r-project.org/web/packages/koRpus/vignettes/koRpus_vignette.pdf

Motyl-Mudretzkyj, I. & Späinghaus, M. (2014). *Anders gedacht: Text and Context in the
German-Speaking World (3rd Ed.)*. Boston: Heinle Cengage Learning.

Saville-Troike, M. (2012). *Introducing Second Language Acquisition* (2nd Ed.). United Kingdom:
Cambridge University Press.

Schiller, A., Stöckert, C., Teufel, S., & Thielen. C. (1999). Vorläufige Guidelines für das

    Tagging deutsche Textcorpora mit STTS. Retrieved from http://www.sfs.uni-

    tuebingen.de/resources/stts-1999.pdf

Schmid, H. (1994). *Probabilistic part-of-speech tagging using decision trees*. Paper presented at

    the International Conference on New Methods in Language Processing, Manchester, UK.

Schmidt, T. (2014). Gesprächskorpora und Gesprächsdatenbanken am Beispiel von FOLK und

    DGD. In *Gespächsforschung- Online-Zeitung zur verbalen Interaktion*, 196-233.

Schmidt, T., & Westpfahl, S. (2013). POS für(s) FOLK – Part of Speech Tagging des

    Forschungs- und Lehrkorpus Gesprochenes Deutsch. *Journal for Language Technology*

    and Computational Linguistics. 1, 139-156.

Schumacher, P. B., & Hung, Y. (2012). Positional influences on information packaging: Insights

    from topological fields in German. *Journal of Memory and Language, 67*(2), 295-310.

Schwitalla, J. (2003). *Gesprochenes Deutsch Eine Einführung*. Berlin: Erich Schmidt Verlag.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27,

    53-94.

Richards, J. (2006). *Communicative Language Teaching Today*. United States of America:

    Cambridge University Press.

Roder, B., Schicke, T., Stock, O., Heberer, G., Neville, H., & Rosler, F. (2000). Word order

    effects in German sentences and German pseudo-word sentences. *Sprache &*

    Kognition, 19(1-2), 31-37.

Rosengren, I. (1993). Wahlfreiheit mit Konsequenzen – Scrambling, Topikalisierung und FHG

    im Dienste der Informationsstrukturierung. In Ries, Marga (Ed.), *Wortstellung und*

    Informationsstruktur (pp. 251-312). Tübingen: Niemeyer.

Vulanovic, R., & Kohler, R. (2009). Word order, marking, and parts-of-speech systems. *Journal of Quantitative Linguistics, 16*(4), 289-306.

Walter, D., & van Compernolle, R. (forthcoming). Teaching German Declension as Meaning: A Concept-Based Approach. *Innovation in Language Learning and Teaching.*

Appendix A - Events from FOLK

		FOLK			
Ereignis	Ereignis-ID	Sprechereignis-Art	Dauer	Sprecher	Erhebungsdatum
1	FOLK_E_00044	Alltagsgespräch: Bandbesprechung	0:31:39	4 Männer	2013
2	FOLK_E_00045	Alltagsgespräch: Bandbesprechung	0:27:19	2 Männer, 1 Frau	2013
3	FOLK_E_00077	Alltagsgespräch: Gespräch beim Friseur	0:20:29	2 Männer	2013
4	FOLK_E_00080	Alltagsgespräch: Pausenkommunikation im Theater	0:18:35	2 Frauen	2014
5	FOLK_E_00081	Alltagsgespräch: Pausenkommunikation im Theater	0:09:00	2 Frauen	2014
6	FOLK_E_00085	Alltagsgespräch: Pausenkommunikation im Theater	0:07:55	2 Frauen	2014
7	FOLK_E_00088	Alltagsgespräch: Pausenkommunikation im Theater	0:18:07	1 Mann, 1 Frau	2014
8	FOLK_E_00092	Alltagsgespräch: Pausenkommunikation im Theater	0:07:13	1 Mann, 2 Frauen	2014
9	FOLK_E_00197	Alltagsgespräch: Pausenkommunikation im Theater	0:15:41	1 Mann, 2 Frauen	2014
10	FOLK_E_00198	Alltagsgespräch: Pausenkommunikation im Theater	0:21:58	1 Mann, 1 Frau	2014

		FOLK			
Event	Event-ID	Speech Event: Type	Length	Speakers	Recording Date
1	FOLK_E_00044	Everyday speech: Band meeting	0:31:39	4 men	2013
2	FOLK_E_00045	Everyday speech: Band meeting	0:27:19	2 men, 1 woman	2013
3	FOLK_E_00077	Everyday speech: Conversation with the barber	0:20:29	2 men	2013
4	FOLK_E_00080	Everyday speech: Intermission communication at the theater	0:18:35	2 women	2014
5	FOLK_E_00081	Everyday speech: Intermission communication at the theater	0:09:00	2 women	2014
6	FOLK_E_00085	Everyday speech: Intermission communication at the theater	0:07:55	2 women	2014
7	FOLK_E_00088	Everyday speech: Intermission communication at the theater	0:18:07	1 man, 1 woman	2014
8	FOLK_E_00092	Everyday speech: Intermission communication at the theater	0:07:13	1 man, 2 women	2014
9	FOLK_E_00197	Everyday speech: Intermission communication at the theater	0:15:41	1 man, 2 women	2014
10	FOLK_E_00198	Everyday speech: Intermission communication at the theater	0:21:58	1 man, 1 woman	2014