

UNIVERZITA PALACKÉHO V OLOMOUCI  
FILOZOFICKÁ FAKULTA

**The Discourse Markers *I Mean* and *You Know* in Fiction and  
Subtitles and Their Czech Translation Counterparts**

Diplomová práce

**Autor:** Bc. Jana Mahnigová

**Studijní program:** Anglická filologie

**Vedoucí práce:** Mgr. Markéta Janebová, Ph.D.

Olomouc 2016

## PROHLÁŠENÍ

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně pod odborným dohledem vedoucí diplomové práce a uvedla jsem všechny použité podklady a literaturu.

V Olomouci dne 8. 12. 2016

Podpis .....

Poděkování patří především vedoucí diplomové práce Mgr. Markétě Janebové, PhD. za cenné rady, připomínky a pomoc při zpracování této diplomové práce. Dále bych ráda poděkovala mé rodině a blízkým, především za podporu a trpělivost.

## ANOTACE

Příjmení a jméno autora: Mahnigová Jana, Bc.  
Název katedry a fakulty: Katedra anglistiky a amerikanistiky, Filozofická fakulta  
Vedoucí diplomové práce: Mgr. Markéta Janebová, Ph. D.  
Počet znaků (včetně mezer): 200 515  
Počet stran: 101  
Název práce: The Discourse Markers *I mean* and *you know* and Their Czech Translation Counterparts  
Název práce v češtině: Diskurzivní ukazatele *I mean* a *you know* a jejich české překladové ekvivalenty

### Abstrakt

Předmětem této diplomové práce je analýza diskurzivních ukazatelů *I mean* a *you know* v paralelním překladovém korpusu InterCorp. Úvodní část se věnuje různým přístupům a pojetí diskurzivních ukazatelů a jejich vlastnostem a také charakterizuje mluvenou a psanou řeč. Analytická část se pak zabývá výsledky výzkumu a zkoumá překladové ekvivalenty *I mean* a *you know* v subkorpusu titulků a beletrie a také se věnuje nulové korespondenci a syntaktické pozici těchto výrazů.

**Klíčová slova:** diskurzivní ukazatele, *I mean*, *you know*, InterCorp, psaná a mluvená řeč

### Abstract

The aim of this diploma paper is to analyze the discourse markers *I mean* and *you know* in the parallel translation corpus InterCorp. Theoretical part provides approaches to discourse markers, describes the properties of discourse markers as well as characteristics of spoken and written language. Analytical part then focuses on the results from the research and discusses Czech translation equivalents of *I mean* and *you know* in the subcorpus of fiction and of subtitles. This part also deals with zero correspondence and the syntactic position of these two markers.

**Key words:** discourse markers, *I mean*, *you know*, InterCorp, spoken and written language

# TABLE OF CONTENTS

<b>INTRODUCTION</b> .....	1
<b>1 DISCOURSE MARKERS</b> .....	3
1.1 Terminology .....	3
1.2 Approaches to Discourse Markers .....	5
1.2.1 Cross-Linguistic Studies .....	13
1.3 Properties of DMs .....	14
1.3.1 Phonological Properties .....	16
1.3.2 Syntactic Properties.....	17
1.3.3 Semantic Properties.....	20
1.3.4 Functions of DMs.....	23
1.4 Characteristics of <i>I mean</i> and <i>you know</i> .....	26
1.4.1 Functions of <i>I mean</i> and <i>you know</i> .....	29
<b>2 SPOKEN AND WRITTEN LANGUAGE</b> .....	38
2.1. Prosody and Punctuation.....	39
2.2 Context-Bound and Context-Free .....	40
2.3 Spontaneity and Deliberate Working Over .....	40
2.4 Interactive and Non-Interactive .....	41
2.5 Interactional and Transactional .....	41
2.6 Informality and Formality .....	42
<b>3 METHODOLOGICAL PART</b> .....	44
3.1 Method .....	44
3.2 Data .....	45
3.2.1 Subtitles.....	45
3.2.2 Core.....	46
3.3 The Positioning of <i>I mean</i> and <i>you know</i> .....	49

3.4	Problems in the Analysis.....	51
3.5	Research Questions .....	53
<b>4</b>	<b>ANALYTICAL PART .....</b>	<b>55</b>
4.1	<i>I mean</i> .....	55
4.1.1	Translation Equivalents of <i>I mean</i> in Fiction.....	55
4.1.2	Translation Equivalents of <i>I mean</i> in Subtitles .....	57
4.2	<i>You know</i> .....	59
4.2.1	Translation Equivalents of <i>you know</i> in Fiction.....	59
4.2.2	Translation Equivalents of <i>you know</i> in Subtitles .....	62
4.3	Properties of Czech Translation Equivalents .....	64
4.4	Zero Correspondence .....	67
4.5	The Occurrence of <i>I mean</i> and <i>you know</i> in Direct Speech .....	71
4.6	Syntactic Positions of <i>I mean</i> and <i>you know</i> .....	73
4.6.1	The Positioning of <i>I mean</i> .....	73
4.6.2	The Positioning of <i>you know</i> .....	76
4.7	The Co-occurrence of <i>I mean</i> and <i>you know</i> with Other Expressions .....	78
	<b>CONCLUSION .....</b>	<b>81</b>
	<b>RESUMÉ .....</b>	<b>86</b>
	<b>APPENDICES .....</b>	<b>92</b>
	<b>WORKS CITED .....</b>	<b>96</b>
	<b>LIST OF TABLES AND LIST OF GRAPHS .....</b>	<b>101</b>

## INTRODUCTION

The aim of this study is to analyze linguistic expressions *I mean* and *you know* and their Czech translation equivalents in fiction and subtitles. This thesis regards them as discourse markers<sup>1</sup> since they have properties associated with such words. In particular, Brinton (2008, 1), characterizes discourse markers as “phonologically short items that are not syntactically connected to the rest of the clause (i.e. are parentheticals), and have little or no referential meaning but serve pragmatic or procedural purposes.” *I mean* and *you know* follow this definition as they stand outside the syntax of the sentence, have little or no propositional meaning but they perform important pragmatic functions.

Discourse markers have been so extensively studied in the recent decades that Fraser (1999, 932) referred to this area as a “growth industry in linguistics.” However, despite the large amount of studies dealing with discourse markers, little attention has been dedicated to *I mean* and *you know* and their Czech equivalents in the multilingual translation corpus InterCorp. To fill this gap, this thesis presents a corpus-based study of these two discourse markers with the focus on their translation since their Czech counterparts can give an insight into their pragmatic functions and meaning. In other words, the research aims to find the most common Czech translation equivalents of *I mean* and *you know* in the corpus of subtitles and fiction, representing spoken and written language respectively, and what their translation equivalents can reveal about their meaning and pragmatic functions. It also aims to find out whether there are any differences between their Czech equivalents in subtitles and fiction. Furthermore, the analysis examines the syntactic position of the markers and investigates whether the position of the marker changes or remains the same in translation. Last, it is investigated whether *I mean* and *you know* tend to co-occur with other elements or discourse markers and what it can indicate about their meaning.

The theoretical part discusses the issue of terminology and it outlines previous approaches and research on discourse markers and their findings. It discusses the characteristics of discourse markers, specifically, their syntactic, phonological and semantic properties and their pragmatic functions. Special attention is paid to the properties of *I mean* and *you know* on the basis of which this thesis regards them as discourse markers. Since the use of *I mean* and *you know* is investigated in subtitles and

---

<sup>1</sup> Although there are numerous different labels for these expressions, the present thesis prefers the term “discourse marker” for the reasons provided in the subsequent chapter.

fiction corpora, representing spoken and written language respectively, these two modes of language are briefly described and the differences pointed out.

In the methodological part, the data and methodology applied in the research are introduced and described in greater detail. The analytical part of this paper then discusses the results from the subcorpora and it deals with the translation equivalents of *you know* and *I mean*, their meaning and their position in the sentence. This part is accompanied by tables, charts and illustrative examples related to the analysis. The last section of this paper summarizes the findings and draws a conclusion based on the research.



# 1 DISCOURSE MARKERS

## 1.1 Terminology

As already mentioned, there has been an abundance of studies dealing with discourse markers. Since the 1970s and 1980s, when discourse markers were gradually becoming a popular topic in linguistics, a number of influential studies for DMs were published, namely Deborah Schiffrin's *Discourse Markers* (1987), Lawrence Schourup's *Common Discourse Particles In English Conversation* (1982), Bruce Fraser's *Pragmatic Formatives* (1987), etc. These and a number of others have provided a great amount of valuable information on this phenomenon, yet the matter of DMs is complicated as the approaches and perspectives adopted to study these items vary greatly.

As a consequence, there is a disunity among linguists on how to call this group of linguistic expressions. The following terms that are used in literature for this group of linguistic items are taken from Fraser's article (1999). Fraser (1999, 932) provides the following labels and the authors who use them: "cue phrases" (Knott and Dale, 1994), "discourse connectives" (Blakemore, 1987, 1992), "discourse operators" (Redeker 1990, 1991), "discourse particles" (Schourup, 1985), "discourse signalling devices" (Polanyi and Scha, 1983), "phatic connectives" (Bazanella, 1990), "pragmatic connectives" (van Dijk, 1979; Stubbs, 1983), "pragmatic expressions" (Erman, 1992), "pragmatic formatives" (Fraser, 1987), "pragmatic markers" (Fraser, 1988, 1990; Schiffrin, 1987), "pragmatic operators" (Ariel, 1994), "pragmatic particles" (Östman, 1995), "semantic conjuncts" (Quirk et al., 1985) and "sentence connectives" (Halliday and Hasan, 1976).

In addition, the issue of the disunity in terminology is further complicated by the fact that authors themselves are not consistent in the labels they use and in the course of time many of them change the name they call these expressions. For instance, Karin Aijmer initially uses the term "discourse particles" (2002) then she adopts the name "discourse markers" (2007) but later she applies the term "pragmatic marker" (2013). Fraser (1999, 936) changes the name "pragmatic formatives" to "pragmatic markers" to refer to expressions that "do not contribute to the propositional content of the sentences but signal different types of messages." The terminology is also complicated by the fact that some scholars are more specific in what they regard as DMs. Lenk (1998, 37) views

“discourse markers” as a subcategory of “discourse particles”. Likewise, Fraser (1996, 2) considers “discourse markers” as a subcategory of “pragmatic markers”.

Despite the intricacy of terms, this thesis prefers the name “discourse marker”<sup>2</sup>. First, it is widely used and many scholars agree that it refers to various linguistic expressions that have important pragmatic or procedural functions in structuring discourse. Furthermore, Jucker and Ziv (1998, 2) claim that this term has the widest currency and its advantage is that this label “enables us to include a variety of elements under a single conceptual umbrella.” This term is also used by Schiffrin, one of the first scholars who greatly contributed to the study of these expressions, and she applies this label for items that contribute to discourse coherence as they “add to its overall structure and interpretation” (Schiffrin 1987, 55).

Second, the name DMs seems to be generally accepted in literature and a number of authors who initially use different label then ultimately come to prefer this term providing compelling reasons for it. Namely, Schourup initially applies the term “discourse particle” in *Common Discourse Particles in English Conversation* (1985) but later in the paper “Discourse Markers” (1999) he presents plausible arguments for the preference of the label DMs. It is explained that the popularity of the term can be attributed to the fact that it has “a narrower and more precisely specifiable reference than discourse particle” while the label “particle” is used for elements that cannot be easily sorted to a well-established word class (Schourup 1999, 229). Moreover, the term DM is more frequently regarded as describing a functional class of items drawn from various syntactic classes (Schourup 1999, 226). Similarly, Blakemore firstly uses the label “discourse connective” (1987) but then argues for the preference of the term DM (2002, 1) claiming that it is more appropriate term since it emphasizes the fact that the item operates on the level of discourse and the term “marker” shows that “their meanings must be analyzed in terms of what they indicate or mark rather than what they describe.”

The terminology is interwoven with the definitions and approaches taken to study DMs. Therefore the great variation in terminology mirrors a certain discrepancy among authors concerning the definition of these items. Schourup (1999, 241-242) claims that a certain item may be viewed as a DM on one definition but not on another and he maintains that this “variation is to be expected in an area that has only recently become

---

<sup>2</sup> Henceforth DM.

a focus of intensive study and which bears on many different areas of discourse research, cognitive, social, textual, and linguistic.” To define DMs is as complicated as the terminology for there are numerous definitions of DMs which mostly depend on the approach adopted by the scholars to study them. Hence the next part looks closer at various perspectives on DMs and briefly outlines the most important theories and approaches applied to the research of DMs as well as areas within which DMs are studied.

## 1.2 Approaches to Discourse Markers

Among the firsts linguists who laid the foundations for subsequent study of DMs were M. A. K. Halliday and Ruqaiya Hasan. Their work is regarded as one of the most significant works for linguists studying discourse and discourse coherence (Rubio 2007, 24). In *Cohesion in English* (1976) they emphasize the difference between a text which forms a unified whole and a sequence of unrelated sentences and they suggest that the difference lies in cohesion which is defined as “relations of meaning that exist within the text, and that define it as a text” (1976, 4). Five distinct means of cohesion in English are identified: reference, substitution, ellipsis, conjunction and lexical cohesion.

Their definition of conjunctions is very close to DMs as they write that conjunctions provide “a specification of the way in which what is to follow is systematically connected to what has gone before”, but in contrast to other means of cohesion, conjunctions express different type of relation since “we are focusing attention not on the semantic relations as such, as realized throughout the grammar of the language, but on one particular aspect of them, namely the function they have of relating to each other linguistic elements, that occur in succession but are not related by other, structural means” (Halliday and Hasan 1976, 227). This definition is in accordance with the fact DMs display a relationship between two discourse segments, which is a characteristic regarded as crucial for the status of DMs by many scholars, for instance by Schiffrin, Fraser and Schourup. Furthermore, they suggest that conjunctive items are expressions drawn from the word classes of adverbs, adverb compounds and prepositional phrases, and include for instance *well, and, now, I mean, but, in other words* which is what other linguists studying DMs agreed on later.

Their work on cohesion and coherence was extremely important as many scholars later approached DMs from this point of view. Specifically, Deborah Schiffrin's book *Discourse Markers* (1987) gave a significant insight into the study of this phenomenon from the discourse coherence perspective. She first provides an operational definition in which she characterizes DMs as "sequentially dependent elements which bracket units of talk" (Schiffrin 1987, 31). She notes that units of talk are as various as tone groups, sentences, verses, propositions or actions. As for brackets, they are part of organizational apparatus of discourse as they can mark discourse units of talk (e. g. conversation, word) but also of social life and social organization in general (ibid. 36). However, towards the end of her book, she revises her operational definition and provides properties characteristic of markers. DMs should be syntactically detachable, utterance-initial, they should have a range of prosodic contours, operate at global as well as local levels and on different planes of discourse (Schiffrin 1987, 328). She studies eleven DMs and claims that they are not members of a single class but various word classes such as conjunctions (e.g. *and*, *but*, *or*, *because*), adverbs (*now*, *then*, *so*, *well*), or they can be lexicalized phrases (*y'know*, *I mean*). Furthermore, Schiffrin considers as DMs a wide range of items as she also includes interjections (*oh*) and she regards nonverbal gestures and paralinguistic features as possible DMs as well. In her analysis, she focuses only on these markers, however, she eventually admits that other items can have marker's use, such as verbs *see*, *look* and *listen* (Schiffrin 1987, 327-328).

According to Schiffrin, the use of markers is crucial for discourse coherence. To find coherence, speaker and hearer have to incorporate and jointly negotiate three components of talk: discourse structure, meaning and action (ibid. 30) and DMs facilitate this as they indicate the location of utterances within the emerging structures, meanings, and actions of discourse (ibid. 22). Put simply, DMs add to discourse coherence as they relate discourse units and they display that relationship which is an important aspect of DMs. She then suggests that DMs then operate on the local level, which concerns the coherence between adjacent units in discourse. In this aspect, Schiffrin agrees with Halliday and Hasan who also claim that these items function to relate linguistic elements that occur in succession. Yet she claims that DMs can operate on global levels as well, i. e. wider structures of discourse.

Another scholar who focuses on DMs from discourse coherence perspective is Redeker who prefers the term *discourse operators* and she defines them as:

a word or phrase – for instance, a conjunction, adverbial, comment clause, interjection – that is uttered with the primary function of bringing to the listener’s attention a particular kind of linkage of the upcoming utterance with the immediate discourse context. An utterance in this definition is an intonationally and structurally bounded usually clausal unit. (Redeker 1991, 1168)

As is evident from the quote, Redeker also emphasizes that DMs link adjacent units, just like Schiffrin does.

Similarly, according to Fraser, the notion of relation is necessary for determining the status of his pragmatic markers and approaches these items from a grammatical-pragmatic perspective to study “the pragmatic role played by terms expressing a semantic relationship between messages” (Fraser 2009, 294). He defines pragmatic markers as linguistic expressions forming a functional class which are part of a discourse segment and although they do not convey propositional meaning, they signal aspects of the message the speaker wants to communicate (ibid. 295). These linguistic items do not constitute a separate syntactic category, instead, they make up a pragmatic class which comprises of expressions drawn from different word classes, mostly conjunctions, adverbs or they can be prepositional phrases and some idioms such as *still and all* and *all things considered* (Fraser 1999, 934). Rather than propositional, Fraser (1999, 944) asserts that these expressions have procedural meaning, which he admits was firstly used by Blakemore (1987), and this type of meaning “specifies how the segment it introduces is to be interpreted relative to the prior, subject to the constraints mentioned earlier.”

His account provides classification of four kinds of pragmatic markers: basic pragmatic markers, commentary pragmatic markers, discourse structure markers and discourse markers. The last mentioned are then characterized as expressions which:

...impose a relationship between some aspect of the discourse segment they are a part of, call it S2, and some aspect of a prior discourse segment, call it S1. In other words, they function like a two-place relation, one argument lying in the segment they introduce, the other lying in the prior discourse. (Fraser 1999, 936)

For this reason, Fraser (1999, 940) regards conjunctions *since*, *because*, *and*, *although* and *however* as DMs as they relate two separate messages. The term segment is understood by Fraser as proposition, sentence, utterance or message therefore the scope of relationship can extend a sentence. Furthermore, the following example demonstrates that the segments linked by a DM do not have to be adjacent (Fraser 1999, 938):

1. *He drove the truck through the parking lot and into the street. Then he almost cut me off. **However**, these weren't his worst offenses.*

In (1), it is exemplified that the DMs *however* relates not only the prior segment but also other prior segments preceding these. In other words, they must represent a semantic relationship between some aspects of the discourse segments which can extend the scope of adjacent sentences.

As already mentioned, since Fraser views DMs from pragmatic perspective, he places particular importance on the fact that these expressions impose relationships on discourse segments. For this reason, he excludes items such *you know* and *well* which do not indicate a two-placed relationship, instead, they introduce a comment or separate message that relates to the following segment (Fraser 1999, 942). He also excludes *I mean* claiming that it does not fulfil the conditions for the status of DMs. Interestingly, he contradicts this statement as he mentions this marker as one of the examples of collateral markers, a subcategory of DMs. Moreover, he is not concerned with exclamation particles (*wow, gosh, damn, yippee*) and interjections (*hey, you there*) as DMs because they are not positioned in the host utterance (Fraser 2009, 97). He also disregards interjections *oh* and *ah* as he focuses only on verbal expressions as well as he disregards non-verbal DMs such as gestures and paralinguistic features such as prosodic features like stress, pauses and intonation (ibid. 298).

Though he takes a perspective on DMs different from the coherence-based approach, he also mentions coherence, asserting the interpretations of S1 as well as S2 must be compatible with the particular DM in order to regard a sequence as coherent (Fraser 1999, 941):

2. *The U.S. policy is crazy. **Furthermore**, I love you anyway.*

In (2), the interpretation of the two segments related is not compatible with the meaning of *nevertheless* and therefore the sequence is incoherent.

In his article "Discourse Markers", Lawrence Schourup (1999, 242) provides a definition of DMs as "a more or less open class of syntactically optional, non-truth conditional connective expressions" and argues that this holds true despite the inconsistency among authors on how to call these items and what is included to the

category of DMs. However, in his *Common Discourse Particles in English Conversation* (1982), Lawrence Schourup offers quite a different perspective on DMs, which deserves some elaboration. In contrast to other scholars, he emphasizes the importance of unexpressed thinking of the conversants as it has a direct influence on what is said or done. Therefore, to deal with functions of DMs he proposes a tripartite model which is defined as:

The covert thinking of the speaker, what that speaker has presently in mind and may, or not, disclose, will be referred to below as the private world; what is on display as talk and other behavior on the part of conversants and is thus available to both the speaker and any other(s) will be called the shared world; and the covert thinking of other conversants, which is invisible to the speaker, will be called the other world. This terminology is applicable equally to any of the participants in a conversation, so that what is private world for one conversant may be the other world for some other conversant. (Schourup 1982, 5)

He defines DMs as loosely as expressions such as forms and short phrases frequent in conversation which are known under the terms “discourse particles”, “interjections”, “discourse markers”, “hesitations” or “fillers” (ibid. 1). Specifically, he focuses on what the expressions such as *oh, like, well, you know, I mean, mind you, hey, aha* and *sort of* indicate about the unexpressed thinking of the conversants. It is demonstrated that they function within his tripartite model and mediate between the unexpressed thinking of conversants and their speech and external behaviours (ibid. 2). To explain, many interjections, for instance *oh, ah* or *aha*, which are called evincives in Schourup’s theory, perform two important functions in conversations: they indicate the real time moment of occurrence of covert thinking and thus they establish the timeliness of the reaction of the speaker and they also reflect the undisclosed thinking of the speaker (ibid. 16). The latter is exemplified in the sentence:

3. **Oh!** *I didn't make the phone call you asked me to.*

*Oh* in (3) suggests that a thought expressed in the following sentence crossed speaker’s mind and indicates that his failure to make the phone was caused by his forgetfulness and it was not an intention (ibid. 15-16).

As for other DMs, it is argued that the function of an item is a result of the interaction between its basic meaning and specific discourse context (ibid. 99). To demonstrate this dependence of functions of an item on its core use and on the

conversational context, Schourup (ibid. 99) puts forward that the core use of *you know* indicates that the speaker is not sure with whether he is understood by the hearer and therefore it functions to reassert control or get confirmation of understanding. Schourup (ibid. 104) concludes that the private world is significant in conversations which is clearly indicated by the high frequency of DMs that are used to bridge the private with the shared world and he adds that limiting our attention in linguistic analysis only to the shared world would ignore two thirds of the aspects of conversations

Schourup's approach resembles relevance-theory in the aspect that he emphasizes the mental state and how DMs help to communicate the intended interpretation which is present only in the private world. Relevance theory, which can be applied in the investigation of DMs, offers a way for understanding cognitive effects in communication comprehension. Sperber and Wilson (2012) criticize the coherence-based approach claiming that it has several shortcomings in the process of sentence interpretation, for instance it is unable to account for disambiguation or repetition and they also maintain that not every communicable message is fully encodable which is at the core of the coherence-based approach. This is in contrast to inferential model of communication which they suggest as an alternative. In this model, only knowing the meaning of the sentence uttered does not suffice for successful communication as what can be communicated is more than is in fact encoded (ibid. 38). To explain, they put forward that a successful inferential communication "involves a communicator ostensibly engaging in some behaviour (e. g. a piece of miming or the production of a coded signal) likely to activate in the addressee (via recognition or decoding) some specific conceptual structure or idea (ibid. 37). A principle of relevance, which is important for this theory as well, then suggests that a speaker engages in some behaviour or makes an utterance as his utterance is relevant enough to worth some processing effort. Hearers then expect every utterance to be relevant for them and thus infer to understand the intended message.

The intended interpretation process of an utterance is then influenced by "discourse connectives," which is their term for DMs. In particular, DMs facilitate the search for relevance as they "constrain the choice of contexts and cognitive effects" (ibid. 205). Sperber and Wilson (ibid. 204) illustrate this on the following examples:

4. *I prefer Edinburgh to London. I hate the snowy winters.*
5. *I prefer Edinburgh to London. **However**, I hate the snowy winters.*



By adding *however* in (5), information on how to manipulate the conceptual representation of utterance was provided. Specifically, the DM constrained the choice of context by contradicting and eliminating a potential expectation that would follow from the first part of utterance and at the same time it raised some expectation about the intended cognitive effects (ibid. 205). Blakemore (1987) deals with the same issue as well and she maintains that there are two types of linguistically encoded meaning, one of them is conceptual, which encodes concepts and contributes to truth conditions of the utterance, and the second is procedural. The latter one is characteristic of DMs, as their function is „to guide the interpretation process by specifying certain properties of context and contextual effects” (Blakemore 1987, 77). To put it differently, what differentiate DMs from other words is the fact that they encode procedural information which expresses how to process the conceptual meaning of the utterance.<sup>34</sup>

This chapter discussed relevant approaches to the analysis of DMs and their understanding of these expressions. The approaches differ, Halliday and Hasan discuss the cohesion of texts and see DMs as one of the means contributing to cohesion, while discourse-coherence perspective focuses on how markers add to discourse coherence. Fraser’s grammatical-pragmatic perspective views DMs as expressions with primary pragmatic role of relating discourse segments and relevance theorists are interested in how the interpretation of utterances is constrained by the use of DMs.

This leads to different terminology as well as diverse definitions of DMs. However, a certain amount of overlap can be detected. Specifically, in all approaches, DMs are seen as linguistic expressions, words or phrases from different syntactic classes and since they take characteristics from their class membership they do not form a homogenous morphosyntactic group. Instead, they constitute a functional class with important pragmatic functions. Authors also agree that DMs can signal relationship

---

<sup>3</sup> DMs can be also approached from the perspective of politeness theory developed by Brown and Levinson (1987). In communication, participants wish to maintain “face” which is “the public self-image that that every member wants to claim for himself” and which comprises of positive face and negative face (Brown and Levinson 1987, 61). Some DMs can be used as a strategy of maintaining face, for instance, *you know* is a marker of positive face as it expresses the speaker’s wish to make his wants accepted by others (Brinton 1996, 188) and the discourse markers *I think* or *I guess* are used to keep positive face for it expresses speaker’s uncertainty and helps to avoid disagreement (ibid 239).

<sup>4</sup> DMs have also been a common subject of many sociolinguistic studies, for instance Erman’s (2001) study of *you know* in the speech of adults and adolescents, or the use of DMs has been examined within different registers, e. g. Szczyrbak’s (2014) research on the use of DMs in police interviews. Such studies provide useful information on the phenomenon of DMs but since these studies are not crucial for this paper and because there is limited amount of space in this thesis, they will not be described.

between discourse units or other segments. Furthermore, Schiffrin, Redeker and Fraser agree that DMs do not have to be adjacent and the relation DMs express can extend to wider structures of discourse and therefore they can have not only local but also global scope.

As regards individual markers, the authors' stances differ. Fraser is concerned with the pragmatic roles played by these items. In contrast to Schiffrin, his account of DMs is more specifically defined as his understanding of DMs rests on the fact that they are only verbal expressions that signal a relationship between two segments hence he excludes interjection *oh* and markers such as *well*, *I mean* and *you know*. Schiffrin includes these, together with non-verbal expressions, paralinguistic features to the category of DMs. It is because her coherence model subsumes also participation plane as she argues that DMs can index utterances to texts as well as participants which is typical of markers *I mean* and *well*. Both Redeker and Schiffrin also explain that DMs function on more levels and thus suggest their multifunctionality in discourse coherence. This is what Fraser disregards and in addition to this, he does not include markers with interpersonal function to the category of DM. For this reason, Fraser's understanding of DMs will not be followed.

Relevance theory which centres around interpretation and focuses on the processes underlying interpretation of utterances and views DMs as constraining the context will be disregarded as well. It is because this thesis presents a corpus-based study focused on translation equivalents of *I mean* and *you know* and relevance theory is not considered significant for this study. For the same reason, Schourup's theory will not be followed.

Instead, this thesis draws on Schiffrin's (1987) understanding of DMs as she admits that functions of DMs can be oriented towards participants and this is the case of *I mean* and *you know*, which are central to this thesis. *I mean* is oriented towards the speaker while *you know* is directed at the hearer. Moreover, Schiffrin concludes that prototypical DMs share certain properties and these characteristics are also associated of *I mean* and *you know*. Namely, both of these markers are syntactically detachable, they can occupy utterance-initial position; they are multifunctional and they operate on more planes of discourse, namely participation framework, information state and ideational structure. As this thesis investigates the translation counterparts of these markers and cross-linguistic studies are important for the present study, they are focused on in the next section.

### 1.2.1 Cross-Linguistic Studies

A different perspective on DMs is found in cross-linguistic studies. In such studies, the intricacy of the translation of DMs has attracted considerable interest. It has been observed that DMs are difficult to translate since they may have no exact equivalents in other languages (Aijmer 2002, 1). Nonetheless the translation is an important tool that reveals a lot about DMs and their meanings as seen from their equivalents in different languages. Furkó (2014, 182) claims that a huge advantage of translation is that it helps to determine the functions of a particular DM in a wide range of bi- or multilingual contexts. Confirming the popularity of DMs in cross-linguistic studies, Karin Aijmer (2007, 34) underlines that they add to our knowledge of language universals and the way the lexical item is translated conveys information about the meanings and functions, specifically, single and infrequent translations may reveal new or emerging developments of a DM while the most conventionalized meanings or functions are reflected in the most frequent translations.

Zero correspondence, which according to Johansson (2007, 58) refers to the cases where there is no formal cross-linguistic correspondence, is frequently applied to the translation of DMs. Aijmer and Altenberg's corpus-based study (2001) discusses possible reasons for omission of DMs. One of the factors accounting for zero correspondence is that there are often no exact equivalents in the target language. For example, they claim that *well* is difficult to translate when there is no standard equivalents in the target language, instead there is a wide range of translations which indicate different shades of meaning and functions (Aijmer and Altenberg 2001, 33).

Nevertheless, they found out that even if there are exact equivalents of DMs available in the target language, DMs are often omitted in translation. Other factors are responsible, namely, zero correspondence can be a result of language-specific conventions or if the context is clear enough, adverbial connectors indicating relationship between discourse units tend to be omitted as the relationship between the sentences can be inferred (Aijmer and Altenberg 2001, 22). Furthermore, the fact that DMs do not affect the truth conditions of an utterance neither do they contribute to its propositional content represents another reason responsible for omission. They also briefly discuss zero correspondence with respect to grammaticalization<sup>5</sup> and they argue

---

<sup>5</sup> Grammaticalization is a diachronic process affecting DMs. This process is defined as “the change whereby lexical items and construction come in certain contexts to serve grammatical functions and once

that the proportion of zero correspondence of DMs is related to the degree of grammaticalization of the particle (Aijmer and Altenberg 2001, 32).

To conclude this chapter, translation of DMs can be a useful tool in cross-linguistic studies. Therefore, this thesis will investigate how *I mean* and *you know* are translated and what their Czech translations indicate about their meaning and functions. In the analytical part, Aijmer and Altenberg's cross-linguistic study (2001) will be used in the discussion of translation equivalents as their article provides a valuable insight into this matter for they address the issue of translation of DMs especially the issue of zero correspondence.

### 1.3 Properties of DMs

This chapter looks closer at the characteristics of DMs. Strictly speaking, most of the authors come to conclusion that a single definition is inadequate for differentiating DMs from other linguistic expressions and they set criteria to distinguish the marker's use from its literal use. These criteria usually consist of properties typical of DMs and like the definitions and terminology concerning DMs, the authors' opinions vary on what is regarded as a property of DMs. Schourup (1999, 241) notes even though most of the definitions concentrate on connectivity and non-truth-conditionality, "no definition seems likely to win universal acceptance in view of the unresolved theoretical differences and varying background assumptions that inform these definitions."

Nevertheless, determining properties of these items is important as it helps to differentiate marker's use of an expression from its literal use. The following characteristics serve as the basis for the criteria applied to the analysis for distinguishing the discourse marker use of *you know* and *I mean* from their literal use. Brinton (1996, 33-35) puts forward the following summary of features characteristic of DMs:

- 1) they appear more frequently in oral rather than written discourse
- 2) they are often stylistically stigmatized and negatively evaluated
- 3) they are short and often phonologically reduced

---

grammaticalized, continue to develop new grammatical functions" (Hopper and Traugott 2003 in Brinton 2007, 61).

- 4) they form a separate tone group
- 5) they are often restricted to sentence-initial position, or may always occur sentence initially
- 6) they appear either outside the syntactic structure or may be loosely attached to it and have no clear grammatical function
- 7) they are optional
- 8) they carry little or no propositional meaning
- 9) they are multifunctional, can operate on both local and global plane

The fact that DMs mostly occur in the oral discourse is simply explained by the characteristics of speech. In contrast to writing, speaking usually does not provide speakers with so much time for planning and DMs can provide clues for the hearer to understand the intended interpretation of the message. Stressing the importance of DMs in spoken discourse, Lenk (1998, 203) claims that these items have significant discourse structuring functions which are beneficial for the hearer. Consequently, DMs prevail in speech-related text types such as records of spoken language (court records), texts written to be spoken (sermons) or text imitating spoken language (plays) (Lutzky 2012, 24).

Moreover, Lutzky notes that the use of DM is related to informality (ibid.). Consequently, the predominance of DMs in oral discourse can be related to the fact that spoken language is frequently less formal. This is confirmed by Archer et al. (2012, 77) who explain that an informal conversation with no DMs whatsoever would be unnatural. Their association with informality may be one of the reasons for the fact that they may be negatively perceived and even criticized. Brinton (1996, 33) claims that when they occur in high frequency in oral discourse or when they appear in written formal discourse, their use is seen as a sign of dysfluency or carelessness. Others assert that only DMs that function interpersonally are occasionally stigmatized (Witczak-Plisiecka 2009, 210). One of them is *you know* about which Schourup (1982, 68) writes that English teachers condemn this DM as a “verbal garbage” or “anemic phrase” and regard it “as a 'crutch' used when one has nothing to say, or when one cannot, or will not bother to, find the proper words to express something.”

### 1.3.1 Phonological Properties

Their predominance in oral discourse is closely related to the next characteristics which concerns their prosody. Phonological properties of markers figuring in numerous definitions concern the facts that DMs do not typically carry stress, are very often phonologically reduced and constitute a separate tone group. According to Hansen (1998, 43), one of the most recognized property of DMs is that they are unstressed. This is confirmed in Mullan's study (2010, 120) in which *I think* when used as a DM with a discourse-organizational function was unstressed, phonologically reduced and spoken at a faster rate. Nevertheless, this is not agreed on universally. Admitting that DMs are often completely unstressed, Archer et al. (2012, 104) argue that markers may also bear nuclear tone and be prominent. In that case, the focus is placed on their additive function which is exemplified on the following excerpt:

6. A: *Tell us about your experiences*

7. B: *Well [pause] it all began when I...*

*Well* in sentence (7) is phonologically emphasized to indicate that a speaker will add more information and the listener will not have opportunity to speak for a while (Archer et al. 2012, 104).

As for the phonetical reduction of DMs, for instance, *you know* can be reduced to *y'know* and *because* to *cos*. Some authors attribute their reduction to their high frequency and to the fact that rather than a propositional they carry an interpersonal meaning (Archer et al. 2012, 104). In contrast, Schourup (1999, 233) maintains this does not have to be their defining property yet most of the DMs are generally reduced irrespective of their clause position and he explains that their phonological reduction correlates with their weak clause association. Formation of a separate tone unit is another consequence of their weak clause association. This means that DMs are prosodically independent and set off from their surrounding context by means of pauses and intonation (del Saz Rubio 2007, 91). However, Urgelles-Coll (2010, 24) argues that even this is not a necessary feature of DMs and illustrates this on the following sentences:

8. *However, this is an issue under discussion.*

9. *This is **however** an issue under discussion.*

While in (8) *however* belongs to an independent tone unit, in (9) this DM is completely integrated into the sentence. As can be seen on the examples, their occurrence in a separate tone unit, it is linked to their loose syntactic position (Lutzky 2012, 38).

### 1.3.2 Syntactic Properties

As regards their syntactic properties, most of the authors deny that the items labelled as DMs form a single homogenous syntactic category;<sup>6</sup> rather they make up a functional class. According to Fraser (1999, 944), DMs consist of heterogeneous expressions such as conjunctions, adverbials and prepositional phrases which do not make up a syntactic category, instead, he argues that DMs make up a pragmatic class. In the same way, Hansen (1998, 65-66) holds the view that instead of constituting an independent formal morpho-syntactic category, DMs represent a functional-pragmatic class. Schiffrin (1987, 40) writes that each item then brings characteristics from its class membership into the discourse class but this is complicated by the fact that some expressions cannot be easily assigned to a syntactic category, for instance, there is a disagreement concerning the word classes of *well* and *oh*.

What is common to most of DMs is their sentence position. Markers are outside the syntactic structure of the clause or loosely attached to it hence they can occur in sentence initial, medial or final positions. Their marginal position can be in writing indicated by a comma separating a DM (Urgelles-Coll 2010, 24). Schourup (1999, 233) maintains that the sentence placement of markers does not affect the functions they perform and demonstrates this on the following sentences where the truth conditions remain intact:

---

<sup>6</sup> It must be pointed out that due to grammaticalization, DMs lose morphological and syntactic characteristics of their word class, their form gets fixed and the marker moves from major (open) to minor (closed) word class (Brinton 2001, 149). For example, the verb *mean* of the DMs *I mean* loses its verbal properties.

10. (*After all/Now/However*), *corgis are an intelligent breed.*

11. *Corgis, (after all/now/however), are an intelligent breed.*

This is in agreement with Aijmer (2002, 29) who analyzed *actually* and *sort of* and found out that these expressions can occupy parenthetical or final positions with little difference in meaning.

However, DMs tend to appear utterance-initially. Schiffrin (1987, 328) views utterance-initial position as a common place of occurrence and regards it as one of the crucial characteristic of DMs. There are several reasons why the initial placement is preferred. Specifically, Schourup (1999, 233) explains that the tendency to position DM initially is linked to their use to restrict context before interpretation of an utterance is made. This resembles one of the points given by Aijmer (2002, 29-30) who specifies the reasons (based on Auer 1996) why the initial placement of DMs is important:

- 1) Initial position provides evidence to DM status.”
- 2) It fulfils pragmatic and interactional purposes which could not be attained with the item placed in its sentence-internal position.
- 3) It expresses the theme of the sentence (e. g. introduces topics, explains relation to the preceding context, shows a change of posture of the speaker or restricts the perspective).
- 4) From the cognitive perspective, the initial position of a DM enhances the processing of the utterance as it provides the context for the interpretation before the sentence is uttered.

The fact that DMs are independent of a sentence structure and do not carry any referential meaning is closely related to their optionality. Yet it must be pointed out that not all DMs are optional and could be simply removed, for instance Fraser (1999) includes conjunctions such as *because*, *whereas*, *since* and *while* to the category of DMs, but those items cannot be omitted as it would result in ungrammaticality since they bear important syntactic roles. However, Schiffrin (1987, 32) notes that the omission of a DM in its sentence initial position does not make the sentence ungrammatical. This is confirmed by Schourup (1999, 231) who elaborates on this and adds that DMs are not only optional in their syntactic sense but also semantic one.



When removed from a sentence, the relationship between the elements it indicates is not explicit yet it is still available and the utterances can be interpreted in roughly the same way which he demonstrates on the following examples (ibid.):

12. *The others are going to Stoke. **However**, I am going to Paris.*

13. *The others are going to Stoke. I am going to Paris.*

This is agreed on by Fraser (1999, 941) who claims that in such cases the explicit meaning of the first and the second segment is related by a DM. In addition to this, he points out that a DM can relate a non-explicit interpretation of S1 to the explicit interpretation of S2 and to illustrate this point he provides examples (14) and (15) where *in spite of* and *so* refer to an implied proposition associated with S1 (ibid. 942).

14. *We left late. **In spite of that**, we arrived on time.*

15. (Boss to assistant) A: *Box up my entire office.*

B: ***So**, he fired you too.*

He also draws the attention to sentences where the DM relates explicit interpretation of S1 to a presupposed proposition (16), to an entailed proposition (17) or the DM relates topics of S2 and S1 (18).

16. A: *I realize that Jack is sick. **But** you know Jack is not sick.*

17. A: *Here is a triangle.*

B: ***But** it has four sides.*

18. *John has been absent lately, hasn't he? **Before I forget**, when are you leaving?*

On the basis of these examples, he concludes that a DM does not only display a relationship between segments but rather “a DM imposes on S2 a certain range of

interpretations, given the interpretation(s) of S1 and the meaning of the DM” (Fraser 1999, 942).

Nonetheless, this does not mean that their syntactic and semantic optionality render them redundant. Scholars agree that they are in fact very important as they indicate a relationship between units and thus facilitate a coherent interpretation of the utterance. Andersen (2001, 42) acknowledges the importance of DMs in utterance interpretation as they specify how the utterance should be understood and thus they facilitate the processes of pragmatic inference which help the addressee to arrive at the intended meaning that a speaker wishes to communicate including his attitudes towards what is said. Moreover, DMs can also convey interpersonal meaning. Aijmer (2002, 5) claims that apart from structuring discourse, DMs also reveal attitudes, evaluation, types of speaker and other aspects of the communication situation and this aspect of DMs which is called indexicality is the most important property of DMs.

Accordingly, their absence in the utterance could aggravate the processing of the utterance and the hearer may encounter difficulties in understanding the intended message. Fraser (1999, 944) affirms this and writes that when a DM is deleted, the hearer is left without a lexical clue indicating the relationship between the two segments. The fact that discourse markers are not necessary for syntactic structure of the sentence but they have crucial function in spoken language was verified in Redeker’s experiment where she removed all discourse markers from a sound track of a television show which consequently slowed down listener’s comprehension (Redeker 1993 in Aijmer 2002, 15).

### 1.3.3 Semantic Properties

DMs are more difficult to be characterized from semantic point of view. As for their meaning, scholars agree that they have little or no propositional meaning,<sup>7</sup> rather they carry procedural meaning and most of DMs do not affect the truth conditions of an

---

<sup>7</sup> This is a result of the grammaticalization process in which the meaning of an expression is weakened and shifts from referential to non-referential meaning. To illustrate this on the discourse marker *I mean*, the verb *mean* is deprived of its full meaning and gets less concrete meaning (Brinton 2007, 62). Additionally, the meaning of a marker often goes through “subjectification” which expresses speaker’s belief or attitude to the utterance and the marker’s meaning may also contain some traces of the original meaning (ibid.).

utterance which is reflected in their optionality and syntactical independence. Proposition is best understood as “the unit which connects the language and the world” and something which bears truth conditions (Andersen 2001, 46). Fraser (1999, 946) puts forward that rather than to the propositional content, they encode procedural meaning for they contribute to the interpretation of an utterance. Similarly, Lenk (1998, 52) points out that DMs do not contribute anything to the proposition of the utterance but they are used in a strictly pragmatic manner indicating sequential and ideational relationships within the discourse.

Nevertheless, this is not always the case. For instance, Urgelles-Coll (2010, 25) illustrates on the examples that some DMs may influence the truth conditions of the proposition expressed by the sentences:

19. *He was really tired. **However**, the noise did not let him sleep.*

*He was really tired. The noise did not let him sleep.*

20. *John went to Paris and **therefore**, Mary went to Rome.*

*John went to Paris and Mary went to Rome.*

It is argued that while in (19) the truth conditions of both sentences are the same and the DM *however* can be omitted without change in propositional meaning, in (20) the truth conditions are affected by the DM and thus removing *therefore* would not convey the causal connection between the two segments (ibid.). Put differently, the DM *therefore* partially contributes to the truth conditions of the utterance. In the same way, Andersen (2001, 49) notes that not all DMs are dispensable with respect to their contribution to propositional meaning, as is the sentence with the DM *sort of*:

21. A: *You said you've always got someone who fancies you.*

B: *No, I didn't, I said I've always got someone who **sort of** fancies me.*

Andersen (2001, 48) puts forward that certain markers add to the propositional meaning of the utterance such as *sort of* in (21), maintaining that “fancying” and “sort of fancying” are not identical in terms of the propositional meaning as the DM hedges the statement. In addition to it, he maintains that *sort of* does not encode procedural

information as most DMs do, rather, it encodes conceptual information and affects the truth conditions of the utterance (Andersen 2001, 62).

To pin down the meaning of DMs, authors tried to define DMs with respect to their core meaning. According to Schourup (1999, 250) every DM has a constant semantic content of some kind and the core is determined when various discourse uses of a DM are compared and decided what these uses have in common. Fraser (1999, 945) likewise notes that every DM has a core which is then enriched by the discourse context and exemplifies this on the DM *but*:

22. *She's good looking. **But** he's ugly as sin.*

23. *He's good looking. **But** that isn't going to get him a job in this market.*

24. *He's late. B: **But** he's not late at all.*

25. *You say that Mary is coming. **But** we weren't talking about Mary at all.*

26. *A: James is not in his office. B: **But** I just saw him there.*

He explains that there are not five distinct meanings of the DM *but*, instead there is an interaction between the core meaning of DM, which is a simple contrast, and the context in which it occurs (ibid.). This is in agreement with Andersen (2001, 40), who notes that DMs contain relatively low degree of lexicality in contrast to their high degree of context sensitivity. Though defining the core of markers may provide some insight into their meaning, this issue is still very problematic since the core meanings suggested have a tendency to be abstract, general or summary (Aijmer 2002, 20) Therefore, Aijmer (2002, 23) puts forward that it is necessary to recognize functions of a particular marker.

### 1.3.4 Functions of DMs

It is generally acknowledged that a marker does not fulfil a single function, rather DMs are characterized as being multifunctional. An abundance of theoretical frameworks classifying functions of DMs have been proposed which is a consequence of the fact that describing and constraining their multifunctionality is not a simple issue for DMs can perform a large number of functions depending on the context and in some meanings they can overlap with other markers (Aijmer and Vandenberg 2011, 229).<sup>8</sup>

<sup>9</sup>This is an area where authors' stances diverge but many scholars propose that markers function primarily on two levels. Lutzky (2012, 39) gives a list of functions of DMs which operate on structural and interactional levels:

<b>Structural</b>	<b>Interactional</b>
Initiating function	Conveying positive or negative attitudes
Closing or conclusive function	Attention-catching
Turn-taking devices	Hesitation devices, fillers
Frame function, marking boundaries in discourse (e. g. topic changes/shifts, digressions...)	Face-threat mitigation (hedging)
Quotative function	Acknowledging function
Introducing parts of an adjacency pair	Qualifier function (signalling some kind of insufficiency)
	Intensifying function

**Table 1.** Functions of DMs according to Lutzky (2012, 39)

A particular marker can fulfil more functions in each group, for instance, a marker can signal a speaker change and thus perform a structural function but at the same time it can convey the attitude of the speaker (ibid. 38). Lutzky divides the functions into two levels but some authors agree on the textual level and further

<sup>8</sup> Aijmer and Vandenberg (2011, 224) argue that the pragmatic functions of DMs are the result of the semantic processes associated with the grammaticalization, a theory in which pragmatic functions develop from semantic meaning through delexicalisation.

<sup>9</sup> Correct assigning of functions may be an intricate task due to the overlap or similarities between functions. Lutzky (2002, 22) explains that such an intricacy is a consequences of grammaticalization.

subdivide the interactional level. Namely, Andersen (2001, 26) perceives their functional complexity as divided into three groups subjective, interactional and textual functions which resemble Lutzky's categorization. Her structural function corresponds to Andersen's textual function since it communicates "what the speaker perceives as the relation between sequentially arranged units of discourse" (Andersen 2001, 66), in other words, it is the connection between propositional content or communicated assumptions. However, what Lutzky sees as interactional functions Andersen further distinguishes the communicative content of utterances and speaker's informative intention into interactional and subjective functions. While subjective function expresses the attitude of the speaker towards a proposition (ibid. 66), interactional function conveys "the speaker's conception of the hearer's relation to the proposition" (ibid. 69).

Similar classification to Andersen's can be found in Erman's study of *you know* (2001). Erman also agrees that DMs perform textual functions but he proposes two other domains. Namely, a social domain which is concerned with the addressee and the management of discourse and a metalinguistic domain in which markers are focused on the message and speaker's attitude towards it (Erman 2001, 1341).

Similar classification is given by Brinton who stresses that in spite of being grammatically optional and carrying non-propositional meaning, DMs bear important pragmatic functions and their absence would result in the unnaturality and disjointedness of discourse (Brinton 1996, 35). She provides a number of functions and then, drawing on Halliday's functions of language, she describes how these functions fit into two modes of language: textual and interpersonal (ibid. 38). The interpersonal mode covers "the expression of the speaker's attitudes, evaluations, judgments, expectations, and demands as well as of the nature of the social exchange, the role of the speaker and the role assigned to the hearer" and is divided into two other functions of DMs (ibid.):

- a) subjective
- b) interpersonal

The textual mode mostly functions to structure text and the entire discourse and subsumes the following functions:

- c) initiating and closing discourse

- d) turn-taking
- e) signalling topic shifts
- f) indicating new and old information
- g) constraining the relevance of adjoining utterances
- h) structuring utterances as a text on a global level
- i) repair making

Brinton's understanding of functions of DMs as working on two levels corresponds to domains recognized by Schiffrin who describes DMs as indexing utterances to the participants, which parallels interpersonal level, and also indexing utterances to the text, paralleling the textual level (Brinton 1996, 39).

To conclude, the terminology, definition and characteristics of DMs is a complex issue and there is a diversity of approaches taken to study DMs which results in the fact that there is no universally agreed set of properties that would be common to all DMs. However, my conception of DMs is that they are various words or phrases from different word classes which have little or no propositional meaning, are syntactically optional and since they are usually not integrated into their host utterance, they can occupy sentence initial, medial or final position. They are more frequent in oral rather than written discourse and they tend to be phonologically reduced. They do not usually carry grammatical functions yet they are very important for their discourse or pragmatic functions which help in the interpretation of an utterance. Their functions fall into two main domains, textual and interpersonal.

Furthermore, in some cases it might be difficult to determine whether an item is or is not a DM. To deal with this, I agree with Hansen (1998, 358) who sees the properties as structured around prototypes. Jucker and Ziv (1998, 2-3) propose this as well and explain that since discourse marker appears to be a fuzzy concept, there are more prototypical members of the class that possess more criterial features and there are peripheral members displaying less characteristic properties. Therefore, the next chapter looks at the two DMs essential in this study, *I mean* and *you know*, and what properties they have. Focusing only on the above stated criteria, I will disregard their phonological properties since they will be analyzed in corpora where no prosodic features are available.

#### 1.4 Characteristics of *I mean* and *you know*

Although some authors exclude them from the group of DMs, for instance, Fraser does not classify them as DMs as they do not signal a relationship between two discourse segments which he regards as a crucial condition for the status of DMs, most scholars, such as Schiffrin (1987), Müller (2005), Fox Tree and Schrock (2002) regard them as DMs and this view is also followed in this thesis. It is because *I mean* and *you know* have the properties which are characteristic of DMs. They are syntactically detachable therefore their omission does not render the sentence ungrammatical and they can be placed freely in the sentence. They do not carry propositional meaning and do not change the truth conditions of an utterance but they perform important pragmatic functions.

Adopting sentences from the subcorpus, let us now exemplify how these properties apply to *I mean* and *you know*. In (27a), *I mean* is placed clause-initially but its position can be changed to clause final (27b), which confirms the fact that the marker is independent of the syntax of the sentence. Therefore, its omission does not affect the grammaticality or the meaning of the sentence as (27a) and (27c) are interpreted in the same way. However, *I mean* specifies that what follows is an adjustment of the prior message and when the marker is deleted (27c), this is not indicated which demonstrates that the marker carries pragmatic functions.

27. a. ***I mean***, *I'm glad that girl gettin a chance.* [JKT\_CON]

b. *I'm glad that girl gettin a chance*, ***I mean***.

c. *I'm glad that girl gettin a chance.*

The same applies to *you know* which in (28a) is placed clause-medially but its position can be shifted to initial (28b) or final (28c). The marker can be just as well removed (28d) as it stands outside the syntactic structure of the host sentence and its omission does not influence the grammaticality nor changes the truth conditions of the sentence. But when deleted, the interpersonal function conveyed by *you know* is absent.

28. a. *Not that he was a*, ***you know***, *left-footer or anything like that.*  
[TPNG\_GOO]

b. ***You know***, *not that he was a left-footer or anything like that.*



c. *Not that he was a left-footer or anything like that, you know.*

d. *Not that he was a left-footer or anything like that.*

The reason that these two markers were chosen for this study is that they are syntactically and functionally similar, both occur with a high frequency in spoken discourse and a number of authors study these two markers together as well. Schiffrin (1987, 267) gives three reasons for treating these two markers together: the use of both of them is based on their semantic meaning, their functions are complementary and both are socially sanctioned. Likewise, Fox Tree and Schrock (2002) discuss these two markers together claiming that their use arises from their basic meanings and they are also alike in terms of their functions. They also argue that two markers appear commonly in conversation as their functions are linked with spontaneous and unplanned conversation (Fox Tree and Schrock 2002, 745). In the same way, Quirk et al. (1985, 1112) subsume these two markers in the same group because they both convey the speakers' comments on the content of the main clause or display the speakers' views on the way they are speaking. Romero Trillo (2015, 8) selects these two expressions because their use as discourse markers significantly outnumbers their literal use in the *London-Lund Corpus*. Since they are syntactically different from other DMs, for instance from *well*, *so*, *right*, *oh* and *right*, they sometimes get different labels, namely *comment clauses* (Redeker 1990, Quirk et al. 1985) or *pragmatic parentheticals* (Brinton 2007).

As for their morphosyntactic properties, both markers consist of a subject and a verb, instead of being single-word items, and this form is morphologically fixed and the verb does not inflect which is in contrast to their literal use. They generally constitute a separate tone unit and they both fit in the first type of Quirk's classification of comment clauses which consist of a transitive verb that is syntactically deficient as it lacks a nominal that-clause as object (Quirk et al. 1985, 1112). Although they are not integrated into their host utterance and can occupy initial, medial or final position, Redeker (1990, 374) claims that comment clauses, especially *you know*, often occur at the beginning of direct quotes as they show the transition to quoted speech (29) while *I mean* signals different type of transition therefore it often appears between a speech error and its repair (30) or to present an explanation (31). Redeker (1990, 374) demonstrates this on the following examples:

29. He says, *you know*, pack and go! That's it! You're out o'here!

30. It goes up/*I mean* down.

31. And then it goes into this *SERIES* of really bizarre pictures about running water ... *I mean* mind you there are hardly an/there were like ... maybe *FIVE* subtitles.

As a result of grammaticalization, both DMs have been bleached of its lexical literal meaning. Yet it is the literal meaning from which their discourse meaning is derived. In case of *you know*, the literal meaning indicates that the hearer is conscious of the information presented by the speaker (Fuller 2003a, 188). *You know*, containing a second person pronoun used as an indefinite general pronoun similar to *one* and the verb *know* indicating the cognitive state in which one has information about something, has therefore two possible meanings in which the information is either available to the recipient of talk or the information is generally available (Schiffrin 1987, 267). As for *I mean*, Brinton (2007, 43) argues that its literal meaning is “to intend (to do something)” is usually accompanied by the *to* complement and illustrates this in the sentence (32).

32. *I didn't mean to be rude last Wednesday.*

But she claims that this structure is less frequent in modern English than it used to be in the earlier periods because the verb *mean* is substituted by other verbs such as *intend* (Brinton 2007, 44). On the other hand, its discourse meaning is “to signify, to intend to convey a certain sense” and this is common in modern English (*ibid.*). In particular, the marker *I mean* modifies two aspects of meaning of speaker's prior utterance which are propositional information or speakers' intentions since the predicate *mean* is polysemous (Schiffrin 1987, 295-296). According to Schiffrin (1987, 267), their literal meanings have an impact on the use of the marker, namely she claims that “*you know* gains attention from the hearer to open an interactive focus on speaker-provided information and *I mean* maintains attention on the speaker.” Interestingly, Schiffrin (1987, 309) argues that *I mean* also maintains hearer's attention because this marker “instructs the hearer to continue attending to the material of prior text in order to hear

how it will be modified.” The present thesis will examine whether the Czech translation equivalents reflect this as well.

### 1.4.1 Functions of *I mean* and *you know*

It is the literal meaning of these DMs that affects their functions. From this follows that the main difference between these two markers is that *you know* is oriented to the addressee’s attention or reception while *I mean* is speaker-oriented and thus it functions primarily in the modification or adjustment of the speaker’s prior utterance (Fox Tree and Schrock 2002, 743). According to Quirk et al. (1985, 1114-1115), these markers focus on participants and he lists the following functions: they hedge as they indicate the speaker’s tentativeness over the truth value of the matrix clause, express the speaker’s certainty, show the speaker’s emotional attitude towards the content of the matrix clause or they are used to attract the hearer’s attention. When these two markers co-occur together, their functions are complementary (Schiffrin 1987, 309).

Building on the core meaning of the marker, Fox Tree and Schrock (2002) are more specific as they list five functions of *I mean*: interpersonal, monitoring, turn management, organizational and repair. First, the marker with interpersonal function refers to the speaker’s need to adjust their speech in order to express exactly what they intend to express and this function includes positive politeness as it indicates casual talk as well as negative politeness as it distances speaker from what he is saying and decreases face threat (Fox Tree and Schrock 2002, 741). Second, *I mean* is used for monitoring addressee’s understanding after the speaker forewarned an adjustment (ibid. 742). Third, the use of the marker with turn management function enables speakers to forewarn upcoming adjustments at any point in the turn and this function correlates with the marker’s position, namely *I mean* in a turn-initial position indicates the speaker’s contribution to an adjustment to the his preceding turn and to exemplify this, Fox Tree and Schrock provide the following example:

33. A: *They tend not to be so dramatic, do hey,*

B: *- I I think it is true that, . a sort of*

A: ***I mean** you’re not going to get a sort of medal for uh, . drafting a beautiful new bill or something like that*

In (33), the speaker A uses *I mean* to signal an adjustment of his prior turn, disregarding the contribution of the speaker B (ibid. 741). *I mean* with organizational function can contribute to the organization, such as to introduce commentary, justification, modification or in topic shifts (ibid. 742).

Fox Tree and Schrock argue that *I mean* also performs a repair function which “conforms with the marker’s basic meaning to forewarn upcoming adjustments” (ibid. 742). This seems to be the most frequent function performed by this marker. Brinton (2007) considers this function as very common as well and Schourup (1982, 57) likewise argues that *I mean* helps to clarify the nature of utterance when a particular aspect of correction is unclear. In research, Romero-Trillo (2015) found out that *I mean* is most frequently used to accomplish self-initiated repair. This is confirmed by Szczyrbak’s study (2014) where *I mean* is mainly used for modification of the ongoing interaction.

On the other hand, *you know* generally performs different functions than *I mean* because of its different core meaning. To explain, Schiffrin (1987, 295) maintains that the speaker is in the position of information provider and therefore he depends on hearer’s reception and *you know* can help him to do so as it has a variety of functions which range from elicitation of hearer’s attention or confirmation of a proposition to relinquishment of the floor. Szczyrbak’s research (2014) shows that *you know* mostly performs organization function and it is commonly used to elicit understanding or signal uncertainty and the same was found in Romero-Trillo’s study (2015) where *you know*’s most frequent function was sympathetic-circularity function.

This is in agreement with Fox Tree and Schrock (2002, 741) who maintain that:

...*you know* informs addressees that an inference is requested, but does not provide information about what kind of inference. That is, rather than forewarning a clarification, the *you know* invites addressees to infer something about what was just said.

Fox Tree and Schrock then discuss the particular functions of *you know* within the same categories as proposed for *I mean*. As for the interpersonal function, *you know* is used to draw addressee inferences or interpretations, either at word level or the interpersonal level and just like *I mean*, *you know* can lessen speaker’s commitment to a face-threatening utterance (ibid. 737). In the turn coordination, *you know* provides the hearer

with an opportunity for inferences from the talk at any point during the conversation and this function correlates with the marker's position, for instance, turn initial *you know* gives the speaker time to consider his inferences when the speaker is talking (ibid. 738). As regards repair function, they claim that when speakers have difficulties to express themselves, they use *you know* to encourage addressees to infer the intentions. As an illustration of this function, they list the following example (ibid. 739):

34. A: *well this brew I made, I . I picked a bottle up, just after you'd gone actually, . there was a filthy mess of yeast, - where it had pushed the cap off . and it was a filthy . you know, not not nasty but, quite a k thick creamy sort of scum of yeast, on thiy um it was dried, **you know**,.*

B: *on the floor .*

A: *on th- on thiy well on thiy s- you know on thiy hatchway there*

B: *oh yeah*

According to Fox Tree and Schrock (2002, 738), *you know* in this function is used to stall time while speakers repair problematical speech.

This DM can facilitate monitoring of listener understanding when the speaker is not sure whether he is understood. In other words, *you know* is used by the speaker to invite addressee's inferences as it is demonstrated in (35) where *m* is a reponse to A's contribution.

35. A: *Faulkner's uh relaxed, but not too relaxed, **you know***

B: *m*

Last function which is identified is organizational function which subsumes three subcategories: topic shifts, emphasis and reference (ibid 740).

Nevertheless, Fox Tree and Schrock's account is inadequate to determine a particular function of a marker as it is difficult to see what falls into their categories. It is because detailed criteria or conditions to distinguish among the functions are missing and in some cases, they provide only sentence-length description of the function. Furthermore, they list insufficient number examples which would support their claims. In case of *I mean*, they provide only one example when the marker functions in turn management and they even do not demonstrate when *I mean* functions in repair which,

as they claim, is its primary function. The same applies to *you know*, for which they list three examples yet none of them concerns the interpersonal function which as they emphasize is the main function of this marker. Since their classification fails to give clear criteria, which are crucial for correct assigning functions, and consequently valid analysis, their account of functions will not represent a basis for analysis in this thesis.

A far more reliable classification of functions of *I mean* is proposed by Brinton who provides the following taxonomy (Brinton 2007, 39-53):

<b>Functions</b>	<b>Subcategories</b>	<b>Reformulation</b>
<b>Appositional meaning</b>	Repair (used for editing mistakes or self-repair)	<i>What I mean to say</i>
	Reformulation (to reformulate preceding utterance)	<i>In other words, what I am saying amounts to the following</i>
	Explicitness (To make preceding utterance more precise or explicit)	<i>Namely, that is</i>
	Exemplification (Providing examples)	<i>For instance</i>
<b>Cause</b>	(To give an explanation)	<i>Because, I am saying this because</i>
<b>Speaker attitude</b>	Emphasis or veracity of an utterance	
	Evaluation or judgement	
	Sincerity	<i>I am serious when I say</i>
<b>Interpersonal meaning</b>	(Markers of the speaker's attention of addressee and expressions of hope that the speaker was understood by the hearer)	<i>I am implying more than I am saying</i>

**Table 2.** Brinton's classification (2007, 39-53) of functions of *I mean*

Brinton defines each function and illustrates it on examples thus her taxonomy represents more useful account of *I mean's* functions than Fox Tree and Schrock's classification (2002). Although she offers a reformulation of the meaning in a particular function which is especially helpful, the reformulation of two subcategories of speaker attitude are not given. What is also problematic about her account is that to the category

of interpersonal use function she includes clauses such as *you know what I mean, if you understand what I mean* (36) and *you see what I mean* (37).

36. *It is because she isn't successful...if you understand what I mean.*

37. *"If it was, then conceivably Congressman Metcalf resented that – Well. You see what I mean."*

Although these phrases have an interpersonal function as they are oriented toward the addressee, this thesis does not consider them as DMs since in such cases *I mean* is a complement of and is related to another clause.

Another shortcoming of her model is that she regards *I mean* in its literal use as an instance of sincerity function. In her examples (38) and (39) (Brinton 2007, 52), the verb *mean* carries the complement *it* which is characteristic of the non-discourse marker use of the expression. Likewise, *I mean* in (40) is complemented by the noun phrase *every point*.

38. *Mistress Page: I mean it not; I seek you a better husband.*

39. *I'm in this for the distance. I mean it.*

40. *...he will lose unless he concentrates on every point, and I mean every point.*

To find a suitable model for classification of functions, attention was dedicated to Erman's taxonomy (2001) who suggests that DMs function in three domains, namely textual, social and metalinguistic and demonstrates the functions on *you know*.

Text monitors	Social monitors	Metalinguistic monitors
Discourse markers	Turn taking markers (highlighting)	Approximators
Editing markers (including repair markers and hesitation markers)	Turn yielding markers (confirmation seeking)	Hedges
	Comprehension securing markers	Emphasisers

**Table 3.** Erman’s classification (2001, 1341) of functions of *you know*

In particular, text monitoring markers are oriented towards the organization of discourse as they indicate transitions between propositions, arguments, states, events or text therefore Erman calls them *discourse markers* (Erman 2001, 1342). The markers in this category can operate either at the clause level or at the textual level. Editing markers consist of hesitation markers which are used for getting time and repair markers which indicate repair. Although both categories of editing markers can be placed freely in the clause, editing markers “usually occur after function words, within the phrase after a determiner, the speaker obviously doing lexical search, or after a con/disjunct at the beginning of the clause for the sake of planning the overall continuation of it” (ibid. 1344). Hesitation markers often co-occur with repetition (41) or with repetition and *like* (42).

41. *Come on Jane. Advertise your business on, on, **you know**, on tape.*

42. *My sister really needed that. She cos like, cos **you know** like these interesting lighters when she went to France?*

Moving to social monitors, when a marker is used with turn-yielding or confirmation seeking function, it is supposed to trigger listener’s response but when it is used with the turn-taking and highlighting function, this enables speaker to select himself as the speaker instead of being selected by a previous speaker (ibid. 1345-1346). The third subfunction of social monitors involves comprehension-securing in which speaker is concerned whether he has been correctly understood by others. Erman



argues that the due to its lexical meaning, *you know* is suitable for this function and he demonstrates the function on the following example:

43. A. *I hate it when you see someone being sick. They go (mimicking vomiting sound)*  
B. *No.*  
A. *I watched erm, **you know** Warren? He was being sick right and he was hanging over the banisters like this in these flats...*

The third domain is associated with the markers that are used to “underscore the illocutionary force of the utterance as a whole” (ibid. 1347). This concerns DMs with emphatic function which enable speaker to stress the illocutionary force of the utterance. Metalinguistic monitors include hedges which are used when the speaker does not want to fully commit himself to the utterance or he wants to downtown the effect of the utterance. This group involves markers such as *I suppose, I (don't) think, I guess* but *you know* is not included as it does not function as a hedge. Nevertheless, because of its lexical meaning, *you know* readily offers itself for the function of approximator which is according to Erman (2001, 1348) performed when speakers are “giving the listener/s a rough but sufficiently exact idea about a certain state of affairs for the general purpose of the conversation.”

At first glance, Erman's understanding of functions as operating in three domains of language and its subcategories seems logical and suitable for analysis as he supports his functional categories with a plenty of examples. But even his model is not completely perfect. First, Erman (2001, 1344) demonstrates on (44) the use of marker with textual function as it indicates inserted comment. However, in the example, *you know* can be used with confirmation seeking function as well. Second, he does not distinguish between the discourse marker's use and the literal use of *you know* as he includes sentences (45) and (46). In these two examples, *you know* is complemented by the object therefore it does not comply to the properties of discourse markers.

44. A: *Shelley, come round to me right, and she was, stroking Dempsey and he walked past wagging his, **you know** when they put the tail down [and]*  
B: *Yeah.*  
A: *(continues)*

45. A: *Stop the tape. Don't even tape this.*

B: *I don't know if I'll be allowed, I'll have to ask cos, **you know** how it is.*

46. A: *I hate it when you see someone being sick. They go (mimicking vomiting sound)*

B: *No.*

A: *I watched erm, **you know** Warren? He was being sick right and he was hanging over the banisters like this in these flats...*

Such shortcomings make his account inadequate for the objective analysis of functions of *you know*.

As it was demonstrated, there are numerous theories applied to clarify the functions of a particular DM, yet this aspect of DMs still remains somewhat intricate for a number of reasons. Some authors are not precise and explicit in defining and delimiting their categories and do not demonstrate their categories on examples, as Fox Tree and Schrock. Other scholars, such as Erman and Brinton, do not differentiate between discourse marker use and literal use of the expression. Moreover, their models for classification of functions are often not perfect as the borders between particular categories are not clear-cut. And since wider context is missing in the data from InterCorp, this would lead to subjective assigning of functions which would result in unreliable research.

This thesis also follows Huang's view (2011, 29) who maintains that in most studies it is often not clear how the functions of DMs are determined because it could be done by guessing or logic and in some cases it can be problematic to assign a function since the researchers cannot read the speaker's mind. Similarly, Scheler and Fischer (1997, 668) suggest that judgments on DMs are often subjective and even with a perfect model for assigning functions of DMs a certain level of dubious cases can be found. Moreover, they insist that even in real setting humans correctly understand only 80% - 90% of discourse meanings (ibid.). This problem is further complicated by the fact that most items can be used as DMs and can also have a "standard" semantic reading, for instance *now* can be a temporal adverb or a DM with a variety of discourse readings and it might be difficult to differentiate between them (Stede and Schmitz 2000, 127). Likewise, Aijmer and Vandenberg (2011, 228) state that some items have clearly

distinguishable clause-constituent functions from pragmatic functions, while the distinction between other items might not always be so straightforward as in the case of the word *surely* which use as an epistemic adverb is hardly distinguishable from its discourse marker use.

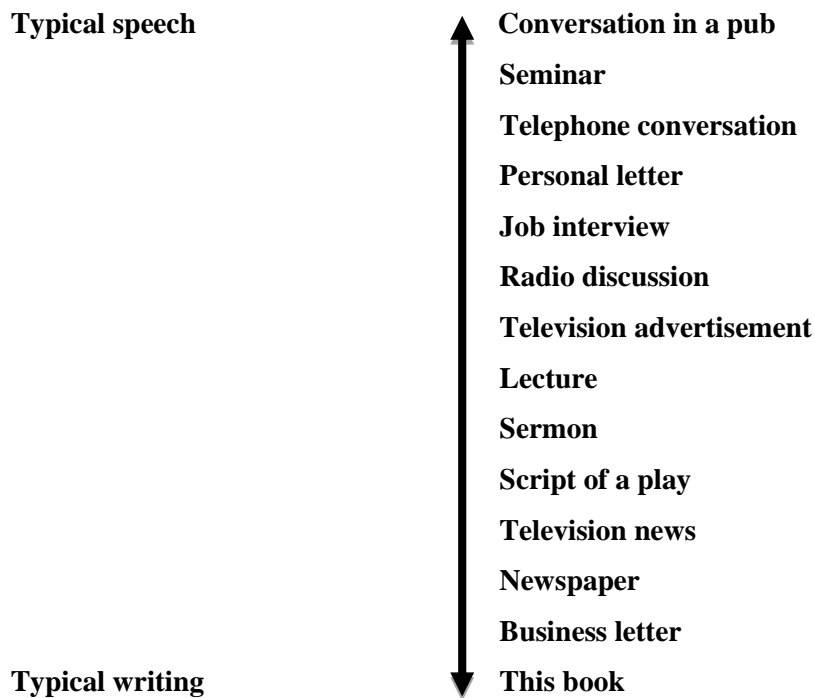
For these reasons, this study would not focus on the analysis of functions of *I mean* and *you know*, rather, the objective of the analysis will be to determine what Czech translations can reveal about the meaning of these two markers with respect to their primary functions. This thesis will examine whether the translation equivalents confirm that *I mean* primarily realizes textual function, such as modification and repair, and whether Czech translations reflect *you know*'s interpersonal function and the marker is used to get confirmation of understanding or sympathy.

## 2 SPOKEN AND WRITTEN LANGUAGE

This study investigates the use of DMs in the corpora of subtitles and fiction and compares their use and Czech translations in spoken and written language. For this reason, this section briefly characterizes these two modes of language and points out the differences between them. First, speech and writing must be defined and distinguished.

Crystal (1995, 291) writes that “speech uses the medium of ‘phonic substance’, typically air-pressure movements produced by the vocal organs, whereas writing uses the medium of ‘graphic substance.’” Although this explains some of the differences between these two modes yet the distinction between spoken and written language should not be made only according to the medium of their production because then a novel read aloud would be regarded as spoken language and a transcribed conversation as a written language. The distinction should not be based on their original modality of composition either because the characteristics of written and spoken language overlap in some respects depending on what type of writing is being compared to what kind of speech (Jahandarie 1999, 132-133). To illustrate this point, a university lecture may with its linguistic properties and content resemble written language rather than spoken and online chatting or text messaging can have more in common with speech than with writing.

Therefore, the suggestion to see spoken and written language not as a dichotomy but rather as a continuum which is proposed by Leech et al. (1982) seems to be reasonable. Leech et al. (1982, 140) argue that it is better to see written and spoken language as a continuum from “typical” speech to “typical” rather than two distinctive “watertight categories.” He presents the continuum below (ibid.):



**Figure 1.** Continuum of speech and writing proposed by Leech et al. (1982, 140)

This is confirmed by Crystal (1995, 293) who states that differences between spoken and written language cannot be taken as absolute differences but rather as trends and he presents similar continuum. The continuum by Leech et al. with typical speech on one end and typical writing on the other is essentially helpful for two reasons. First, it accounts for the cases where certain text type might contain features typical of speech as well as writing and explains the overlap of characteristics of these two mediums. Second, focusing on typical speech and typical writing certainly helps to pin down the differences between them. Therefore, the following differences are characterized with respect to typical instances of speech and writing.

## **2.1. Prosody and Punctuation**

The first difference which concerns prosody and punctuation is a result of different production methods of speech and writing. The fact that spoken language is produced by the vocal organs enables the speaker to make use of paralinguistic features and employ different intonation patterns, rhythm, pitch, volume stress and other prosodic patterns. Speech can be accompanied by other nonverbal signals available to a speaker, such as gestures, posture, facial expressions or haptics. These features cannot be

engaged in writing yet they can partially compensated for by using various graphical features such as punctuation, capital letters, paragraphs, italicization, underlining, bold-facing as well as periods, question marks or exclamation marks which can indicate intonation (Jahandarie 1999, 134).

## **2.2 Context-Bound and Context-Free**

Speakers can draw on the non-verbal features as well as the immediate situational context hence the spoken language tends to be less explicit. Biber et al. (1999, 1044) state that speech is typically less lexically dense, has a low degree of grammatical elaboration and specification of meaning in contrast to writing as the participants of conversation share a great deal of contextual background. For this reason, spoken language also contains more deictic expressions which relate to the situation or the speaker. The shared knowledge among participants attributes for the high frequency of pronouns in conversation, especially personal pronouns which directly refer to the participants, as the speaker assumes that hearer identifies the reference of pronouns (ibid.).

On the contrary, written language is “desituated” and is only slightly affected by the physical and social situation of its production and reception (Chafe 1994, 45). In order to make up for the lack of immediate spatio-temporal context, writers are more explicit and specific. In general, written language is more lexically dense, includes more nouns as well as longer premodified noun phrases and the packaging of information is more concentrated in contrast to spoken language (Brown and Yule 1983, 16). Interestingly, the fact that written texts generally tend to contain more explanatory details holds true even if the only receiver is the writer themselves (Jahandarie 1999, 138).

## **2.3 Spontaneity and Deliberate Working Over**

This difference is discussed by Chafe (1994, 43) who claims that lack of time for careful planning in conversation results in quick activation of new topics and ideas while in writers can afford to work over the text several times before presenting to the readership. Writing is not constrained by time thus it is more carefully structured as the

writer has more time to plan, organize, revise and rewrite. The planning also allows for more precise words, complex syntactic structures, longer sentences, subordination and higher information density.

In contrast, speech is constrained by time hence the speaker has to speak and plan simultaneously or during a certain limited amount of time and he is deprived of the advantage to reorganize or return to what he has already uttered. Consequently, speech is typically marked by the occurrence of errors, incomplete sentences, pauses, repetition, false starts (ibid. 139). The spontaneity of speech also attributes for the use of DMs that speakers employ in order to modify their utterance, make repairs and reformulation. Similarly, DMs in speech may be used as hesitation markers or verbal fillers in order to provide speakers with some time for planning what to say next.

## **2.4 Interactive and Non-Interactive**

Due to the physical absence of the addressee, typical writing is not as interactive as speech and since the writer does not always know who the addressee will be, he must be more explicit and precise in writing. In contrast, speech is interactive for the physical presence of the addressee enables the speaker to monitor the reaction of the hearer and get an immediate feedback and if needed, the speaker is able to straightaway adjust or clarify his speech. To do so speakers can draw on some DMs, namely, the marker *I mean* can help the speaker to signal modification or repair of previous utterance. Moreover, speakers make use of monitoring and interaction features to show that they are aware of the addressee's presence and his reactions and encourage them to actively participate (Leech et al. 1982, 139). For instance, *you know* can help speakers to attract hearer's attention, request his confirmation or elicit inferences.

## **2.5 Interactional and Transactional**

Scholars agree that speech and writing differ in terms of functions they perform, the former having interpersonal function and the latter is used to convey information. Brown and Yule (1983, 13) maintain that apart from a few exceptions,

...in daily life in a literature culture, we use speech largely for the establishment and maintenance of human relationships (primarily interactional use), we use

written language largely for the working out of and transference of information (primarily transactional use).

The distinct functions then have an effect on the linguistic characteristics of each mode. The characteristics of writing is suited for transmission of facts and knowledge which influences its linguistic properties and Sindoni (2014, 31) lists following features that predominate in writing: more content and lexical words, longer words and more varied vocabulary, complex noun clauses and nominalization and the language tends to be more precise and specific.

This is not the case of speech, the purpose of which is to establish relationships, to show sympathy or create intimacy, this is reflected in the high frequency of personal pronouns, expressions indicating personal feelings, opinions and predominance of personal references and emphatic particles (Jahandarie 1999, 139-141). One of the ways to establish a relationship is by using DMs as some of them can function interpersonally, for instance to convey attitudes, get hearer's attention or as hedges. This is in agreement with Bazzanella who maintains that DMs in informal conversation carry out primarily a phatic function as informal conversation is mainly phatic (Bazzanella 1990, 630 in Aijmer 2002, 177).

## **2.6 Informality and Formality**

Typical spoken language contains more informal linguistic features than writing. It is because speech is used in less formal situations and is largely used for interpersonal purposes. Informal features that are common in speech include lengthy coordinate sentences, nonsense vocabulary, slang, contracted forms and some DMs. Namely, *well*, *you know*, *you see*, *actually*, *sort of* are DMs that are typical of spoken language (Stenstrom 1990 in Aijmer 2002, 177). These properties are not characteristic of writing which tends to be more formal and makes use of multiple subordinations in one sentence, complex syntactic patterns and longer sentences (Crystal 1995, 291).

This chapter outlined major differences between spoken and written language and it was noted that these two cannot be defined as two distinct categories with clear-cut borders since there is an overlap of characteristics among different types of speech and writing. For this reason, it is better to view speech and writing as continuum as is proposed by



Leech et al. (1982) and describe their differences and characteristics with respect to “typical” speech and “typical” writing. These are influenced by the different manner of production of speech and writing as well as by their different functions. In particular, interpersonal function of speech and its spontaneity and informality accounts for the use of DMs as they can carry out phatic function and have textual functions, such as repair, reformulation, exemplification etc.

### 3 METHODOLOGICAL PART

#### 3.1 Method

This thesis provides analysis of 2,000 tokens of *I mean* and *you know* in spoken and written language and their translation equivalents and syntactic positions. The method used in the research is a contrastive study which systematically compares two or more languages in order to determine their similarities and differences (Johansson 2003, 31). In such studies, the use of multilingual corpora is particularly common as it provides a wide variety of data in a number of languages. Johansson (2007, 51) characterizes parallel corpora as “collections of texts in two or more languages which are parallel in some way, either by being in a translation relationship or by being comparable in other respects” and emphasizes the value of multilingual corpora claiming that:

The use of multilingual corpora, with a variety of texts and a range of translators represented, increases the validity and reliability of the comparison. It can be regarded as the systematic exploitation of the bilingual intuition of translators, as it is reflected in the pairing of source and target language expressions in the corpus texts. (2007, 52)

The main asset of cross-linguistic perspective is that translation patterns can “make meanings visible through translation” (Johansson 2007, 57). Therefore, this thesis focuses on Czech translation counterparts as the way a lexical expression is translated can reveal the meaning and functions of a word which is especially useful in case of DMs which carry little or no propositional meaning. A special attention will be devoted to zero correspondence, examining the claim made by Aijmer and Altenberg (2001) that omission is a common strategy of DMs even if there are suitable translation equivalents available in the target language because these expressions do not have any propositional meaning. Moreover, a quantitative analysis will be carried out as the frequency of translation equivalents will be counted and compared.

Furthermore, *I mean* and *you know* will be analyzed in the speech and writing respectively in order to find out whether their translation equivalents differ in these two subcorpora. As the data drawn from InterCorp, which is introduced in the following section, do not include audio recordings of speech, spoken language will be represented by subtitles.

Since the data available in parallel corpora enable us to examine the syntactic position of DMs in target as well as source texts, the present thesis will also analyze whether the position of the marker remains unaltered or whether is changed in the translation. Last, it will be investigate whether *I mean* and *you know* tend to co-occur with other expressions or DMs.

## 3.2 Data

InterCorp<sup>10</sup> is a translation parallel corpus which contains original texts and their aligned translation counterparts in one or more languages therefore it represents a suitable source of data needed for the purpose of this thesis. A wide range of text samples are available, such as core (mostly fiction), political commentaries, newspaper and legal texts, proceedings of the European Parliament and film subtitles. Czech is so called pivot language which means that Czech text, either translation or source, is aligned with one or more foreign language texts (Rosen 2016). The texts are automatically aligned except for core which is manually aligned which reduces the chance for misalignment. The alignment is carried out on the sentence level and this means that a sentence in the source language is aligned with a corresponding sentence or sentences in the target language (Cvrček 2014).

### 3.2.1 Subtitles

The data for this analysis of subtitles were taken from InterCorp which includes film subtitles from [www.opensubtitles.org](http://www.opensubtitles.org). It must be pointed out that users contributing to this website are often not professional translators and this may impinge on the quality of the translated texts. This thesis regards subtitles as a representation of speech although they are not a perfect representative of speech, they only approximate spoken language. However, this thesis follows Toury's view who maintains that, rather than as translated written texts, subtitles tend to be regarded as translated utterances and very often also function as such (Toury 1995 in Karamitroglou 2000, 34).

---

<sup>10</sup> The corpus is available at <http://ucnk.ff.cuni.cz/intercorp/>.

The present thesis analyzes altogether 1,000 instances of the two discourse markers in subtitles<sup>11</sup> and thus a subcorpus had to be made. When creating a subcorpus of English subtitles, the following steps were taken: at the website, InterCorp version 8<sup>12</sup> was chosen, then the subcorpus was given a name, alignment with Czech was checked and as a text type were chosen subtitles. English was set as a source language as well as original language and the subcorpus was created. The size of the subcorpus in tokens is 55,217,046. After that, a query was entered and in order to obtain data from subtitles after year 2000 with English as a source language, the following formula for *I mean* was entered in the CQL slot:

```
[word="I"] [word="mean"] within <div origyear>="2000" & group="Subtitles" & srclang="en" />
```

InterCorp generated randomly shuffled 18,989 hits for *I mean*. These were downloaded in xlsx format and then had to be manually sorted so as to distinguish *I mean* in its literal use and in its discourse marker use. This was done according to the criteria given in section 3.2.3 of this thesis. Then several categories were created in order to get results for the research questions listed in section 3.5.

The same procedure was applied to *you know* but the words searched in the formula had to be modified so as to get results for this marker. The changes concerned only the words in brackets, i. e. [word="you"] [word="know"], otherwise the rest of the formula remained same. When the whole formula was entered, InterCorp generated 33,027 hits for *you know*.

### 3.2.2 Core

The data from core are regarded as a representing written language. To create the subcorpus of fiction, the selection of the individual books had to be done manually. Books with English as a source language published after 1960 were chosen from InterCorp version 9 and only one book by an author selected in order to avoid idiolects

---

<sup>11</sup> Ideally, only one variety of English should be analyzed. However, this could not have been done as subtitles contain mostly films with American English whereas fiction mostly British English.

<sup>12</sup> InterCorp version 8 contains the same amount of subtitles as version 9.

and certain style of translators. The selection of books consisting of 50 titles was used for creating the subcorpus of fiction and this subcorpus was aligned with Czech. The size of the subcorpus in tokens is 6,394,150. After that, the search for *I mean* and *you know* was carried out, InterCorp generated 899 randomly shuffled tokens for *I mean* and 2,486 for *you know*. Again, these were downloaded in xlsx format, sorted according to the criteria and analyzed.

### 3.2.3 Criteria for determining discourse marker use of *I mean* and *you know*

Since the data included also literal use of *I mean* and *you know*, these had to be excluded from the analysis. To distinguish between discourse marker use and non-discourse marker use, the following criteria were applied:

1. The omission of the DM does not render the sentence ungrammatical. It can be deleted but it contributes to the interpretation of the unit.
2. When omitted, the truth conditions of an utterance remain the same
3. The DM is not followed by a complement, e. g. noun, adjective, pronoun, to-complement, nominal group, subordinate clause.
4. It is independent of syntactic structure and thus it can occupy initial, medial or final position (flexibility of position), the DM is frequently separated by commas.
5. The DMs do not co-occur with auxiliary verbs and is morphologically inert.

#### **The following phrases were discarded as well:**

*What I mean is, I meant, by this/that I mean, you know/see what I mean, as you know, if you know, you know what?, Do you know...?, you know that ...*

Then I went through the data one by one, the cases where *I mean* and *you know* were not considered as DM were discarded until the number of 500 tokens of *I mean* and 500 tokens of *you know* in subtitles as well as core was reached, altogether 2,000 occurrences of these two markers were collected.

On the basis of the criteria, the following examples illustrate what was excluded from the study. Particularly, sentence (47) was disregarded as the verb is followed by the complement *it* which indicates the literal use of *I mean*. Similarly in (48) the subordinate clause is dependent on the verb *mean* and the verb in (49) is in the past tense. As for *you know*, in example (50) *you know* is related to the subordinate clause and sentence (51) contains non-discourse meaning and represents subordinate clause since it is introduced by the conjunction *as*.

47. *I mean it, slow down.* [JEEPERS\_01]<sup>13</sup>

48. *I mean what I say.* [COUNTESS\_09]

49. *Maybe it wasn't what I meant.* [ARS\_HI]

50. *Dad, you know what happened to the guy in the parking lot.* [CONTRACT\_06]

51. *As you know, I can be a mean old judgmental bastard.* [JF\_COR]

What fulfilled the conditions is demonstrated on example (52) and (53) where *I mean* stands outside the syntactic structure of its host utterance, does not carry any complement and could be omitted without affecting the truth conditions of the sentence. However, it carries important function as it indicates that what follows is an explanation of prior meaning of the sentence. Similarly, *you know* in (54) is syntactically detachable and its omission would not change the truth conditions of the utterance. In (55), *you know* indicates its discourse marker use by being separated by commas and by being placed outside the syntax of the host utterance. Moreover, its deletion would not change the meaning of the sentence.

52. *I hope it's alright I'm being here, **I mean**...it's at the middle of the night!*  
[MATADOR\_05]

---

<sup>13</sup> For the list of abbreviations of InterCorp texts quoted in this thesis see Appendix 1 and 2.

53. “*But the more people the better really – I mean, Michael Corner and his friends wouldn’t have come if he hadn’t been going out with Ginny* – “[JKJ\_THR]

54. *You’re on for noon, you know.* [JG\_CLI]

55. “*They make nests, you know, gorillas,*” *said the angel, pouring another drink and managing to hit the glass on the third go.* [TPNG\_GOO]

### 3.3 The Positioning of *I mean* and *you know*

To determine the position of these two markers within a clause, Macaulay’s classification (1991) was used. He notes that DMs can occur freely in the sentence and defines the positions:

By initial position is meant either the first position in the clause or immediately following a coordinating conjunction or a discourse marker such as *well*; by medial position is meant any position preceded and followed by any constituent other than a coordinating conjunction or a discourse marker; and by final position is meant a position followed by no constituent other than a terminal tag such as *and that*. (Macaulay 1991, 156 in Macaulay 2002, 754)

Applying his classification in the analysis, *you know* in (56), is clause-initial, while the marker in (57) is identified as clause-final. Clause-initial position is also assigned when the DM is preceded by other DMs, a greeting, interjection or by a name.

56. *You know, Mary, I can’t understand why you put yourself in that guy’s hands.*  
[DL\_CHA]

57. “*I’ve gone right off this towel, you know,*” *said Zaphod handing it back.*  
[DA\_RES]

As for clause-medial position, this position is assigned in cases where the DMs is inserted and interrupts a phrase, such as in (58) where *you know* is placed within a verb

phrase. Similarly, *you know* positioned clause-medially in (59) is inserted within a prepositional phrase.

58. *By the side of the road there was an old woman selling...um...you know...milk and so the Roman soldier went to her and asked her how much a cupful would cost.* [GD\_BIR]

59. *“Or you go around with sawn-off shotguns and everyone’s got these cars with, you know, knives and guns stuck on –”* [TPNG\_GOO]

As clause-medial uses of the markers are also regarded cases where the speaker starts a sentence and then changes his mind and restarts all over again with the marker or when he hesitates. This is typical of *I mean* (60) but it sometimes occurs with *you know* as well (61).

60. *“You lie down on your – I mean, you lean over the bank, you know, and slough the things about in the water.”* [JKJ\_THR]

61. *“Like a...you know, like a sit-in,” suggested Tom.* [ZS\_ONB]

The DMs were counted as clause-medial when they occurred before an apposition (62) and (63).

62. *I was thinking, the first thing we should do is Expelliarmus, you know, the Disarming Charm.* [JKR\_HAR]

63. *...she’s become a lesbian but that’s all right these days and is doing a dissertation on the films of Sergio Leone as seen from a feminist perspective, and our Stan, you know, our Sandra’s twin, I told you about him last time...* [TPNG\_GOO]



### 3.4 Problems in the Analysis

Before proceeding to the analytical part, some issues must be mentioned. First, during the research, some problems were encountered, specifically the misaligned sentences occurred in the data as subtitles in InterCorp are automatically aligned. Such tokens as well as the cases where the context was missing and only the marker itself appeared were excluded from the analysis.

Second, it must be pointed out that although subtitles are regarded as a representative of spoken language, their translation differs from conventional translation because acoustic as well as visual channel must be accommodated in the process of translation which puts spatiotemporal constraints on the process of subtitling (Karamitroglou 2000, 10). This has an impact on the quality of translation as this may emerge in the higher number of zero correspondences. According to Johansson (2007, 61) another reason for which translation corpora must be approached with caution in cross-linguistic research are so called translation effects which refer to the fact that “translations may reflect features of the source language.”<sup>14</sup> Although the wide variety of texts chosen in the analysis should partially eliminate this problem, this factor still must be taken into account because the influence of English as a source language on Czech translation may affect the results of the analysis. Moreover, a translation may reflect a certain style or carelessness of a translator but this should be eliminated as well by the wide range of text samples.

Third, an issue concerning DMs many other analysts deal with is that it is sometimes problematic to differentiate between the discourse marker use and the literal use of a DM. Despite the established criteria, in some cases it was ambiguous whether *I*

---

<sup>14</sup> According to Johansson (2007, 61), this can be simply resolved in the bidirectional translation corpus which can control translation effects. He explains that in such corpus, translation equivalents can be verified as: “the comparison can proceed in either direction, e.g. from original text to translation or from translation to the sources in the original text” (ibid. 52). Then, the mutual correspondence of translation counterparts can be then statistically measured. Mutual correspondence calculates “the frequency with which (grammatical, semantic and lexical) expressions are translated into each other” in a bi-directional translation corpus (Altenberg 1999, 254 in Johansson 2007, 26). Although this method offers itself for the investigation of translation counterparts of *I mean* and *you know*, it could not be used for the insufficient amount of data available in InterCorp. First, there is a lack of data for subtitles translated from Czech to English. Second, the translation counterparts for *I mean* and *you know* from Czech to English would be impeded on by non-discourse instances of these lexical expressions.

*mean* is a discourse marker or not. Therefore, sentence (64) and (65) were excluded as *I mean* bears noun phrases as a complement and is not separated by commas which indicates this is literal use of the phrase *I mean*. In contrast, *I mean* in (66) is regarded as a discourse marker as it is separated by commas and it can be omitted without changing the meaning. It is also because Brinton (2007, 45) considers such instances as discourse marker use and she demonstrates this on the example (67).

64. *The past, **I mean** the real past, matters less than we pretend.* [JB\_SEA]

65. ***I mean** me and the cowboy.* [CD\_CIR]

66. *That smell, **I mean**.* [CP\_CHO]

67. *“I’ll see you in the morning.” She laughed. “**I mean**, afternoon.”*

Next, in majority of cases Macaulay’s classification of clause positions (2002) proved useful yet there were some cases where the determination of the position was difficult as in (68) and (69).

68. *“I was saying to him, you should probably ... because he’s so talented, **you know**, you could, like, invite him round or –”* [ZS\_ONB]

69. *“I am not forgotten, **you know**, no, I still receive a very great deal of fan mail...”* [JKR\_HAR]

*You know* in (68) was identified as being in clause-medial position while in (69), the position was marked as clause-final as it is considered as relating back to the preceding clause and the following clause is introduced by the expression *no*.

Last issue concerns direct speech. The DMs were categorized as belonging to direct speech when the quotation marks were included either in the English sentence or in Czech translation. However, if the marker appeared in a longer stretch of text, the quotation marks did not appear there as InterCorp gives only the preceding and following sentence.

### 3.5 Research Questions

The analysis of DMs focused on the following research questions:

**1. What are translation equivalents of *I mean* and *you know*? What are the most frequent ones?**

As already mentioned, translation can reveal the meaning and function of DMs so this question aims to find out what the translation counterparts indicate about the meaning and function of *you know* and *I mean*. The attention will be also paid to Czech translations in order to determine whether they have the same properties as DMs.

**2. Do translation equivalents differ in the corpus of subtitles and fiction, representing spoken and written language respectively?**

This question will investigate whether the translation counterparts in fiction and subtitles are the same and possible differences between their translations will be pointed out and explained.

**3. As for fiction, what is the frequency of occurrence of these two discourse markers in direct speech (as indicated by quotation marks)?**

The thesis regards fiction as a representative of written language but it sets to find out whether the DMs occur in the direct speech as they are characteristic of spoken rather than written language.

**4. Do DMs tend to be omitted more frequently in subtitles than in fiction? Is this tendency confirmed by both *I mean* and *you know*?**

Since DMs carry little or no propositional meaning, a certain degree of zero correspondence is expected but it will be studied whether the amount is higher in subtitles or in fiction and possible reason will be discussed. Moreover, it will aim to find out whether the tendency to omit DMs can be confirmed by both markers as their primary functions differ. The possible differences between omission of *I mean* and *you know* will be considered.

**5. What is the most frequent sentence position of these two markers? Does the position of the marker remain the same in the translation?**

This question will test whether the general tendency of DMs to be positioned clause-initially is confirmed by both *I mean* and *you know*. It will also analyze whether there is a correlation between the sentence position of a DM and its function and if the position remains unchanged in translation.

**6. Do the two markers tend to appear together with other discourse markers or other expressions?**

The co-occurrence of *I mean* and *you know* with other DMs or other expressions will be covered as well. Namely, it will be explored what the clustering with certain elements can reveal about the function or meaning of DMs.

## 4 ANALYTICAL PART

The present chapter provides a detailed description of the results from the analysis and discusses the answers to the research questions. First, translation counterparts and zero correspondence are addressed, then syntactic positions are dealt with and finally the attention is paid to clustering.

### 4.1 *I mean*

#### 4.1.1 Translation Equivalents of *I mean* in Fiction

Translation equivalents of *I mean* identified in fiction subcorpus are shown in Table 4.

Translation equivalents of <i>I mean</i> in fiction	$\Sigma$	%
zero correspondence	147	29,4%
<i>myslím; tím myslím; myslím tím</i>	88	17,6%
<i>chci (tím) říct</i>	86	17,2%
<i>totiž</i>	74	14,8%
<i>viš/víte; viš, co myslím</i>	18	3,6%
<i>tedy</i>	13	2,6%
<i>teda</i>	12	2,4%
<i>no</i>	8	1,6%
<i>teda myslím; teda chci říct</i>	8	1,6%
<i>já jen; já jenom</i>	7	1,4%
<i>přece</i>	6	1,2%
<i>rozuměj/rozumíš</i>	6	1,2%
<i>vlastně</i>	5	1,0%
<i>jenže pochop, pochopte</i>	4	0,8%
<i>viš, jak to myslím; viš myslím si</i>	4	0,8%
<i>ale vždyť</i>	3	0,6%
<i>chápeš/chápej</i>	2	0,4%
<i>mám dojem</i>	1	0,2%
<i>mám na mysli</i>	1	0,2%
<i>prostě</i>	1	0,2%
<i>přemýšlel jsem</i>	1	0,2%
<i>podívej</i>	1	0,2%
<i>třeba</i>	1	0,2%
<i>tak mě napadlo</i>	1	0,2%
<i>tím chci naznačit</i>	1	0,2%
<i>to jako</i>	1	0,2%
<b>TOTAL</b>	<b>500</b>	<b>100%</b>

**Table 4.** Czech translation equivalents of *I mean* in the subcorpus of fiction

As is evident from the table that, apart from zero correspondence which is discussed in the next subchapter, there is a wide variety of translations of this DM yet in majority of cases *I mean* tend to be translated as a clause with covert subject *já* (=I). The most frequent Czech translation counterparts involve the verb *myslet* (=to mean), either itself or accompanied by the pronoun *tím* (=it) and almost the same percentage is represented by the clause *chci říct* (=I want to say, I intend to say). It is also very frequently translated as the particle *totiž* (=in fact, because) or *teda/tedy* (=then) or as combination of these particles and the verb *myslím* or *chci říct*.

The investigation of translation equivalents was mainly motivated by the fact that translation can reveal the meaning and function of DMs. As already mentioned, Schiffrin (1987) maintains that *I mean* functions to adjust the propositional information of speaker's prior utterance or to modify his intentions. This is reflected by Czech translations such as *myslím* (=I mean); *tím myslím* (=by this I mean); *chci říct* (=I want to say); *teda/tedy* (=then); *totiž* (=in fact); *vlastně* (=actually) which can be used in repair or modification of both propositional meaning and intentions. For instance, in (70), the speaker uses *I mean* to reformulate his preceding utterance while the marker in (71) is used to exemplify the intention of the speaker. The translations in (72) and (73) indicate that the marker is used to specify or explain speaker's intentions.

70. **I mean**, I'm saying, but you have a natural femininity so it doesn't really...it's... [THE REAL L\_10]

*Já vím, ale ty máš v sobě tu přirozenou... **myslím tím**, že máš v sobě přirozenou ženskost, takže to není tak...*

71. How does your wife feel about it, Mr Swallow, **I mean** is she with you about the kids and the school and the house and all? [DL\_CHA]

*Co na to vaše žena? **Myslím** na ty problémy s dětmi a školou a domem a tak.*

72. **I mean** – you were so young. [ZS\_ONB]

***Totiž** – byla jste tak mladá.*

73. “**I mean**, there's no problem, I look young, I look good, but...” My voice trailed off. [JA\_CAL]

*“Já jenom... se mnou žádný problem není, vypadám mladě, dobře, ale...”  
Nedořekla jsem.*

The list of translation equivalents encountered in the analysis also contains cases where the translation counterpart is oriented towards the hearer, such as *víš/víte* (=you know); *víš, co myslím* (=you know what I mean); *rozuměj/rozumíš* (=you see); *pochop/pochopte* (=you see); *chápeš/chápej* (=you see/understand) or *víš, myslím si* (=you know I think). They make up 6,8% (34 instances) of all translation equivalents in the data analyzed which means that these were not so common as *chci říct* and *myslím* but they indicate that *I mean* can also carry interpersonal function as it can be used to attract hearer's attention or express speaker's need to be understood by the hearer. This confirms the claim by Schiffrin (1987, 309) who writes that *I mean* can be used to maintain the attention of the hearer as this marker makes the hearer more attentive so as to find out what kind of adjustment or modification will follow. This is shown in (74) and (75) where *I mean* is oriented to attracting the hearer's attention and eliciting his sympathy.

74. *I mean, you can hear everything in the next trailer, so our neighbours had suffered through this with us.* [JG\_CLI]

*Víte, v sousedních přívěsech je všechno slyšet, a tak naši sousedi si protrpěli všechno s námi.*

75. *I mean, just look around.* [TPNG\_GOO]

*Víte, jak to myslím, stačí se přece rozhlídnout.*

#### 4.1.2 Translation Equivalents of *I mean* in Subtitles

As regards the translation equivalents in subtitles, the two most common translations are the same as in fiction, i. e. the clauses *myslím (tím)* (=I mean, by this I mean) and *chci říct* (=I intend to say). These are followed by the particles *teda, tedy* (=then) and *vlastně* (=actually). All encountered Czech counterparts are summarized in Table 5.

Translation equivalents of <i>I mean</i> in subtitles	Σ	%
<i>zero correspondence</i>	281	56,2%
<i>myslím; tím myslím; myslím tím</i>	82	16,4%
<i>chci (tím) říct</i>	53	10,6%
<i>teda</i>	20	4,0%
<i>vlastně</i>	10	2,0%
<i>tedy</i>	9	1,8%
<i>mám na mysli</i>	7	1,4%
<i>ale vždyť</i>	7	1,4%
<i>viš/víte</i>	6	1,2%
<i>takže</i>	4	0,8%
<i>totiž</i>	3	0,6%

<i>no</i>	3	0,6%
<i>ale</i>	2	0,4%
<i>prostě</i>	2	0,4%
<i>přece</i>	2	0,4%
<i>teda myslím; teda chci říct</i>	2	0,4%
<i>mám dojem</i>	1	0,2%
<i>já míním</i>	1	0,2%
<i>viš jak</i>	1	0,2%
<i>já jen</i>	1	0,2%
<i>říkám</i>	1	0,2%
<i>přemýšlel jsem</i>	1	0,2%
<i>podívej</i>	1	0,2%
<b>TOTAL</b>	<b>500</b>	<b>100,0%</b>

**Table 5.** Czech translation equivalents of *I mean* in the subcorpus of subtitles

Another research question addressed the possible differences between the Czech equivalents in fiction and subtitles. It can be concluded that the translation equivalents do not differ considerably in subtitles and fiction; *I mean* is most frequently translated with the verb *myslet* (=to mean), namely as *tím myslím; myslím tím; myslím* (=I mean; by this I mean; I mean this; I think or I mean), which accounts for 17,6% in fiction and 16,4% in subtitles and *chci říct* (=I intend to say) which is represented by 17,2% in fiction and 10,6% in subtitles.

However, an interesting finding emerged in the subcorpora of fiction as the list of Czech translations contains longer and more complex equivalents than in subtitles. The list of translation counterparts that are encountered in fiction such as *tak mě napadlo* (=it occurred to me); *tím chci naznačit* (=by this I indicate); *to jako*; *viš, jak to myslím* (=you know how I mean it); *viš, myslím si* (=you know I think); *přemýšlel jsem* (=I was



*thinking*); *jenže pochop* (=but you see, but you know); *pochopte* (=you see); *rozuměj/rozumíš* (=you see) did not occur in subtitles at all. The frequency of each of these translations is not high, they appear only once or twice in the analyzed subcorpora, yet altogether they account for 2,4% (12 instances) of all translation equivalents and this indicates a difference between translation of subtitles and fiction which arises from the fact that subtitles are constrained by time and space. Consequently, the translation counterparts in subtitles are usually shorter and single-word items, such as *takže* (=so); *no* (=well); *ale* (=but); *prostě* (=just); *přece* (=yet); *já jen* (=I just), etc. Since fiction is not limited in this way, the translators can use longer and more complex translation equivalents. Another factor responsible for this is that subtitles from InterCorp are very often not translated by professionals and thus they opt for the most common words while fiction is translated mostly by professionals who can choose more creative and explicit translation counterparts.

Another difference concerns particles *totiž* (=in fact) and *teda* (=then) as equivalents of *I mean*. *Totiž* covers 14,8% of tokens in fiction while in subtitles it makes up only 0,6%. Instead, *teda* (20 tokens, 4%) is the next most frequent equivalent after *myslím (tím)* and *chci říct* in subtitles. To compare, *teda* accounts for only 2,4% (12 tokens) in fiction. The reason is that subtitles approximate spoken language which is usually more informal and colloquial than written language. According to *A Dictionary of Standard Czech Language* (Havránek et. al 2011), *teda* is colloquial and slightly less formal than *totiž*, thus it predominates in subtitles.

## 4.2 *You know*

### 4.2.1 Translation Equivalents of *you know* in Fiction

As for *you know*, the following table summarizes the Czech equivalents which are found in the subcorpus of fiction and their frequency.

Translation equivalents of <i>you know</i> in fiction	Σ	%
<i>víte/víš</i>	208	41,6%
<i>zero correspondence</i>	167	33,4%
<i>vždyť víš/víte; však víš/víte</i>	26	5,2%
<i>chápeš/chápete; však chápete</i>	13	2,6%
<i>abyste věděl/a</i>	9	1,8%
<i>no víte přece; to přece víte</i>	7	1,4%
<i>vždyť to znáš</i>	6	1,2%
<i>jak víš/víte; jak dobře víš</i>	6	1,2%
<i>no, jo</i>	5	1,0%
<i>víš/víte co</i>	4	0,8%
<i>znáte to</i>	4	0,8%
<i>hele/heleď</i>	4	0,8%
<i>že jo?; že ano?; co?</i>	3	0,6%
<i>ehm</i>	3	0,6%
<i>rozumíš/rozumíte</i>	3	0,6%
<i>poslyš/poslechněte</i>	3	0,6%
<i>prosím tě</i>	2	0,4%
<i>to je přece jasný</i>	2	0,4%

<i>no však víš</i>	2	0,4%
<i>abych pravdu řekla</i>	2	0,4%
<i>věřte tomu; věřte</i>	2	0,4%
<i>povím ti; řeknu ti</i>	2	0,4%
<i>vlastně; teda</i>	2	0,4%
<i>nevíte; nemyslíte</i>	2	0,4%
<i>možná by ses divil</i>	1	0,2%
<i>myslím</i>	1	0,2%
<i>vidíte</i>	1	0,2%
<i>no víš</i>	1	0,2%
<i>jistě jste slyšel</i>	1	0,2%
<i>představte si</i>	1	0,2%
<i>nebo jsi zapomněl</i>	1	0,2%
<i>jestli vás to zajímá</i>	1	0,2%
<i>prostě</i>	1	0,2%
<i>jen si pamatuj</i>	1	0,2%
<i>mezi námi</i>	1	0,2%
<i>copak nevíš</i>	1	0,2%
<i>věřili byste</i>	1	0,2%
<b>TOTAL</b>	<b>500</b>	<b>100%</b>

**Table 6.** Czech translation equivalents of *you know* in the subcorpus of fiction

As Table 6 shows, the most frequent translation counterparts of *you know* include the verb *vědět* (=to know) or *chápat* (=to understand) with the subject *ty* (=you – tu form) or *vy* (=you – vous form) which is covert. These are sometimes accompanied by *vždyť* or *však* (=but). It can be concluded that *you know* is in majority cases translated with the

verb *vědět* (=to know) as when all translation equivalents containing the verb *vědět* are considered, they cover 52,6% of all translation equivalents.

The Czech equivalents mirror the interpersonal function of *you know*. Except for interjections *ehm*; *vlastně* (=actually), *teda* (=then), *no* (=well) and *jo* (=yeah), all of the translations are oriented toward the addressee. The results show that *víš/víte* (=you know) and *chápeš/chápete* (=you see) are commonly used to elicit understanding and sympathy which is demonstrated in (76). The comprehension securing function of the marker is also indicated by the translation equivalents *že jo?*; *že ano?*; *co?* (=right?; am I right?; is that right?) and sentence (77) shows speaker's intention to get a confirmation to his statement while sentence (78) demonstrates his purpose to elicit inference from the hearer. In several sentences, such as (79), *you know* was used to attract hearer's attention which was reflected by its Czech counterpart *hele/heleď* (=look) and *poslyš* (=listen).

76. **You know**, it looked sorta funny. [JKT\_CON]

*Vypadalo to strašně srandovně, **chápeš**?*

77. Still, you take what you can get, **you know**? [IB\_SON]

*Člověk holt musí vzít zaviděk tím, co je po ruce, **že ano**.*

78. "Guys like Barry, they have so much rage against women, **you know**."

[JA\_CAL]

*"Chlapi jako Barry v sobě dusí moře nenávisti k ženám, **co**?"*

79. **You know**, this restaurant does have an indoor section. [DS\_SEC]

***Heleď**, tahle restaurace má stolky i uvnitř.*

As the examples demonstrate, *you know* translated as *chápeš/chápete* or as a question tag is used to request confirmation or elicit inferences while the translation of the marker as *vždyť víš* or *však víš* (=but you know) suggests shared knowledge between the speaker and the hearer. Sentences (80) and (81) show that there is some background information among the participants and the speaker assumes that the hearer has the indicated information.

80. *It's not to, you know* – [RAMPAGE\_09]

*Není to tak že, vždyť víš* –

81. *Do you think, veggie girl. You know...* [SEE NO EVIL\_06]

*Ta vegetariánka, však víš...*

Sometimes, *you know* was used only as a verbal filler or hesitation marker which was revealed by its translation as the interjection *ehm*:

82. “Well,” Theodore would say at last, getting to his feet, “I think perhaps we ought to..., *you know*..., just see what the other side of the lake has to offer.”

[GD\_BIR]

*"Poslyš," řekl po chvíli Theodor a vstal. "Měli bychom už asi jít a podívat se, co se dá ... ehm ... objevit na druhém břehu jezírka."*

#### 4.2.2 Translation Equivalents of *you know* in Subtitles

The translation counterparts were analyzed in subtitles as well and these are presented in Table 7.

Translation equivalents of <i>you know</i> in subtitles	Σ	%
<i>zero correspondence</i>	294	58,8%
<i>víte/víš</i>	143	28,6%
<i>vždyť víš; však víš/víte</i>	21	4,2%
<i>chápeš/chápete</i>	19	3,8%
<i>no</i>	5	1,0%
<i>jasný; jasně</i>	4	0,8%

<i>jo?; že jo?; že?; ne?</i>	4	0,8%
<i>prostě</i>	3	0,6%
<i>nevíte</i>	2	0,4%
<i>rozumíš</i>	2	0,4%
<i>víš co</i>	1	0,2%
<i>vždyť to znáš</i>	1	0,2%
<i>dobře</i>	1	0,2%
<b>TOTAL</b>	<b>500</b>	<b>100%</b>

**Table 7.** Czech translation equivalents of *you know* in the subcorpus of subtitles

As shown in Table 7, all the translation equivalents reflect the interpersonal function of *you know*. Translation equivalents of *you know* with highest frequency in subtitles involve the verb *vědět* (=to know), specifically conjugated forms *víš* or *víte* (=you know, T-V forms) which are sometimes accompanied by *vždyt'/však* (=but), *jak* (=as) or *přece* (=but). These amount to 33% in subtitles and thus this the most frequent translations include the verb *vědět* which is the same as in fiction. Czech translation *chápeš/chápete*, aimed at eliciting sympathy and understanding, occurred in subtitles as well. Both subcorpora also contained particles *no* (=well) and *jo* (=yeah) as well as question tags *že jo?*, *jo?*; *že?*; *že ano?* and *ne?* (=right?; am I right?; isn't it?) but they were not so common.

Concerning the differences among the translation counterparts in fiction and subtitles, similar tendency that was observed in case of *I mean* applies to *you know*. As already mentioned, the list of the translation counterparts of *I mean* in fiction contains longer and more complex expressions in comparison with subtitles. This is confirmed by *you know* as a couple of its translations in fiction are represented by subordinate clauses such as *abyste věděla* (=you know); *abych pravdu řekla* (=to be honest) or *jestli vás to zajímá* (=if you are interested). These and several others that are more complex and longer, such as *no to přece víte* (=but you know that); *nebo jsi zapomněl* (=or have you forgotten); *jistě jste slyšel* (=you have surely heard); *to je přece jasný* (=but that is obvious); *jak víš/víte* (=as you know); *prosím tě* (=please); *věřili byste* (=you would believe it); *poslyš/poslechněte* (=listen); *copak nevíš* (=don't you know) and *jen si pamatuj* (=just remember) were identified only in fiction but not in subtitles. Altogether, these complex translations constitute 7,8% of Czech counterparts in the subcorpus of fiction.

Moreover, the list of translation counterparts of *you know* contains 37 items in fiction in contrast to only 13 items in subtitles. Again, this tendency is attributed to the fact that the translation of fiction is carried out mostly by professionals who can opt for more explicit and creative translations which is not the case of subtitles. Moreover, shorter and single-word translations of DMs in subtitles are a result of the fact that translation in subtitle is limited by time and space.

### 4.3 Properties of Czech Translation Equivalents

This section discusses similarities and differences between *I mean* and *you know* and their most common Czech translation counterparts as one of the interesting findings which emerged from the analysis concerns the fact that Czech translation counterparts of *I mean* and *you know* do not always have the same properties as English discourse markers. First, the syntactic structure of most frequent translations *myslím (tím)* (=I mean this, by this I mean) and *chci (tím) říct* (=by this I intend to say) partially resembles the syntactic structure of *I mean* as these translations consist of a verb and the subject *I*, which is covert in Czech translation, and they are also oriented to the speaker. However, this is not the case of particles *teda*, *tedy*, *vlastně* and *totiž* (=actually) which are not focused on the speaker but rather on the modification of the meaning of the text.

Second, one of the characteristics of the DM *I mean* is that it does not carry any complement which is not always the case of its Czech translations. In the data, sentences with Czech counterparts of the DM *I mean* are mostly used without a complement (83) but cases where the Czech equivalents is followed by some complement occurred as well which is demonstrated in (84) and (85) where an object clause connected by the conjunction *že* (=that) is related to *chci říct* (=I intend to say) or *tím myslím* (=by this I mean) which behave as main clause. Besides, *myslím* and *chci říct* are very often accompanied by the object *tím* (=it), as in (84) and in (85).

83. *I mean*, we were out in the open, but nobody else was around. [SB\_HE]  
*Chci říct, byli jsme venku, ale poblíž nikdo nebyl.*

84. *I mean*, the pattern is a little different on each side. [LB\_CAP]  
*Chci tím říct, že ten vzor je na každé straně trochu odlišný.*

85. *I mean*, he's running around like he's on fire. [TALLADEGA\_06]  
*Tím myslím, že tam pobíhá, jako by byl v plamenech.*

These Czech translation counterparts which are followed by a subordinate clause are integrated to the sentence structure. Since another subordinate clause is attached to them, these equivalents occur in clause-initial position and they cannot be omitted as it would lead to ungrammaticality. This is in contrast to discourse marker use of *I mean*.

This does not apply to the particles *teda*, *tedy* (=then), *totiž* (=in fact) and *vlastně* (=actually) which are syntactically independent of the host utterance and can therefore occupy any clause position or they can be omitted without change of meaning. In the analyzed data, they also tend to appear initially and they signal that what follows is the modification of the meaning of prior utterance which is demonstrated in (86).

86. *I mean*, the armory has been broken into a few times. [NARC\_02]

*Teda*, sklad zbraní byl párkrát vyloupený.

Third, when Czech translation equivalents with the prototypical syntactic properties of DMs are compared to *I mean*, we see that what some of them they have in common is their meaning. Particles *tedy*; *teda* and *totiž* carry little or no propositional meaning and their omission would not render the sentence ungrammatical nor meaningless. However, just like *I mean*, they contribute to the interpretation of the utterance. Brinton (2007, 39) claims that *I mean* helps the speaker to establish and negotiate meaning with the hearer and this holds true also for these translation counterparts. For instance, Kolářová (1998; my translation) writes that *tedy* and *teda* do not have a specific meaning but these particles play an important role in the comprehension of the text as they help to establish the meaning of an utterance. In contrast, the meaning of *chci říct* and *myslím* is not as propositionally empty as the meaning of these particles as they indicate the intentions of the speaker to a greater extent than *teda*; *tedy* or *totiž*.

The syntactic structure of most common Czech translations of *you know*, which are *víš/víte* (=you know) and *chápeš/chápete* (=you see), resembles the English DM *you know* as they consist of a covert subject *you* and a verb and are oriented towards the hearer. Like *you know*, the verbs *víš/víte* or *chápeš/chápete* are in present tense and in most of the cases they do not carry a complement (87). As for the syntactic position, the Czech counterparts *víš/víte* and *chápeš/chápete* can be freely placed in the sentence as they are not integrated to the host utterance.

Concerning the meaning, when *you know* is translated as *víš/víte* or *chápeš/chápete*, in such cases the equivalents carry a little propositional meaning and do not change the truth conditions of an utterance. Rather they function interpersonally as they relate to the hearer.

87. *But, **you know**, you do realize how much you love home.* [BREAK\_06]

*Ale, **víš**, uvědomíš si, jak moc miluješ domov.*

88. *But rocks are like, **you know**, land. It's like these rocks are a kit.* [CP\_CHO]

*Ale kameny, **víš**, to je něco jako pozemek. Jako by ty šutry byly stavebnice.*

In sentences (87) and (88), the Czech equivalents can be omitted since they do not affect the grammaticality and the meaning of the sentence. Likewise, translation counterparts such as *ehm; no; prostě; hele/heled'*, which have no propositional meaning, rather they indicate speaker's hesitation or stalling time as in (89) or his need to attract hearer's attention, can be omitted without changing the grammaticality or meaning of the sentence. The same can be said about the meaning of question tags which are aimed at eliciting hearer's confirmation or sympathy which sentence (90) shows, rather than carrying propositional meaning.

89. *Or, **you know**, was there some kind of...* [IMPULSE \_08]

*Nebo, **prostě**, bylo tam něco...bylo tam něco, co bylo znatelně odlišný?*

90. *Well, maybe that short leash is her safety net, **you know**?* [MOTIVES \_07]

*No, možná to zkrátí oka v její ochranné síti, **ne**?*

Most of the Czech equivalents have the same characteristics as *you know*. Nonetheless, cases where the counterparts did not have the properties typical of DMs occur in the data as well. Namely, sentences like (91) where the conjunction *že (=that)* follows the verb *víš/víte* appear in the subcorpora even though they are rare.

91. *They make them up as they go along, **you know**.* [LB\_CAP]

***To** **víš**, že si je vymýšlejí za pochodu?*

As this sentence shows, the verb *víš* bears the object *to (=it)* and is related to a subordinate clause introduced by *že (=that)*. Furthermore, cases where the translation equivalents *víš/víte* are modified by *však, přece* or *vždyť (=but)* and thus did not behave like *you know* are found (92) as well as Czech translations where *you know* is modified by the adverb *dobře (=well)* and also the conjunction *jak (=as)* as in (93).



92. *You know, her, in her private life.* [CP\_CHO]

*Však víš, vidět ji někde v jejím soukromí.*

93. *My dear wife always does, you know.* [JL\_ALO]

*Jak dobře víš, má drahá žena umí být velmi přesvědčivá.*

This is in contrast to *you know* which when used as a DM is always placed outside syntactic structure of the host utterance and is not modified by any other element. In such cases, the translation equivalents do not behave like *you know* as when the marker is translated as superordinate clause, it is not outside the syntactic structure and its position is to clause-initial. Consequently, the omission would result in ungrammaticality as there is another element related to the verb. Furthermore, the equivalents in sentences (91), (92) and (93) are not propositionally empty as they indicate some shared or background knowledge between the hearer and the speaker. Therefore, their omission in Czech would also affect the meaning of the sentence.

This part presents an interesting finding that the properties of Czech translation equivalents are not always in accordance with properties of prototypical DMs. In other words, *I mean* and *you know* are sometimes translated to Czech not as DMs but as literal use of these expressions.

#### 4.4 Zero Correspondence

In the analyzed data, high percentage of zero correspondence was identified for both markers in subtitles as well as fiction. The omission of *you know* and *I mean* was determined when no equivalent occurred in the Czech translation and this is illustrated in sentences (94) and (95).

94. *I mean, if you have something to say to me in the gym, well, then, fine.*

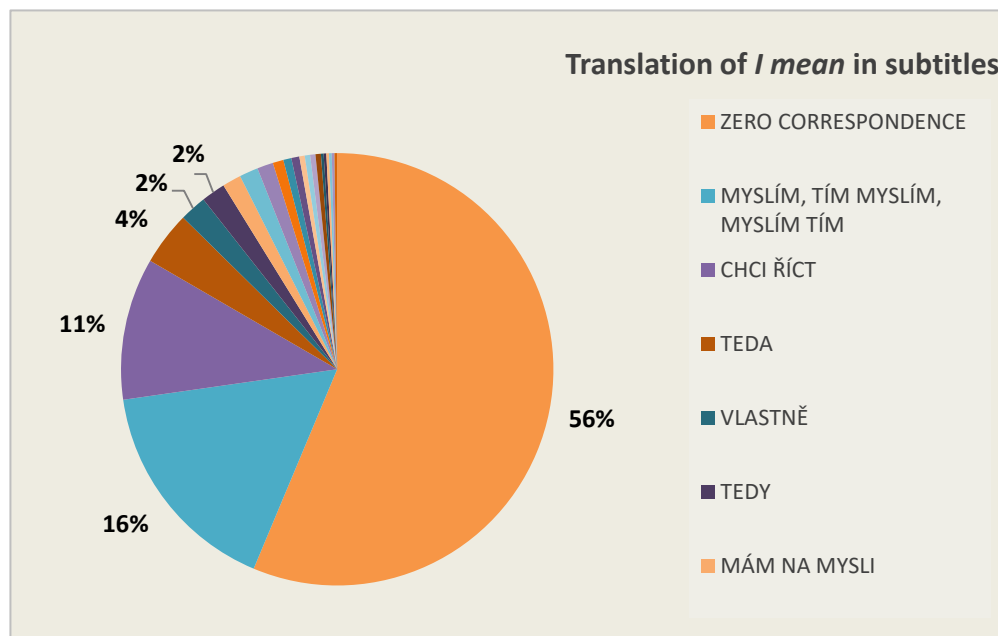
[FAHRENHEIT \_04]

*Když chcete něco říct v posilovně, tak prosím.*

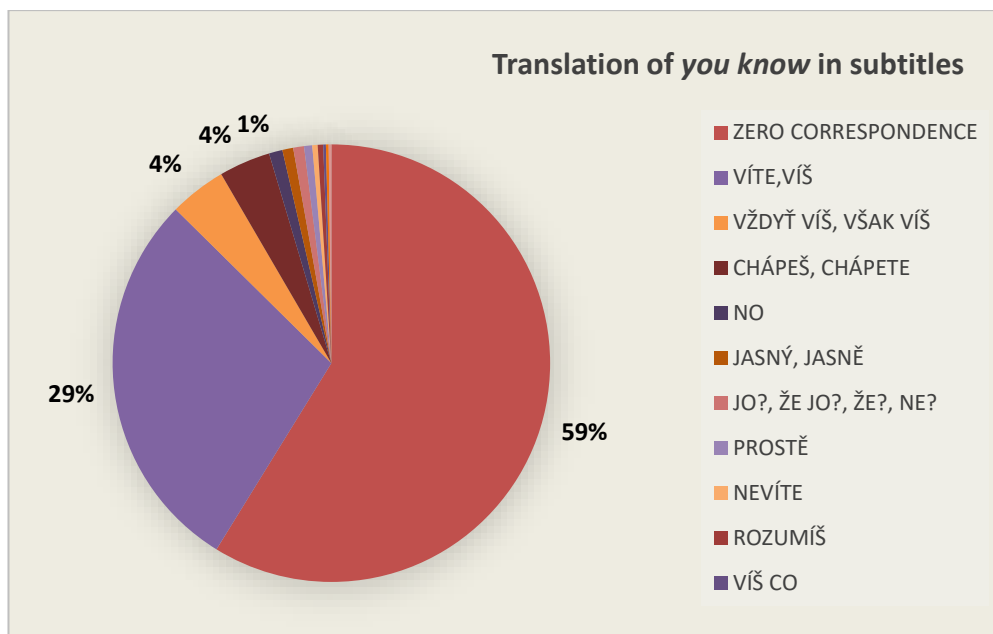
95. “Well, I wouldn’t overstress that angle, **you know**,” he said finally, “one’s never alone with a rubber duck.” [DA\_RES]

“Tenhle moment bych nepřeceňoval,” ozval se konečně, “s gumovou kachničkou není člověk nikdy sám.”

Graph 1 and Graph 2 below that show the proportion of omission to other translation equivalents demonstrate that omission was a predominant strategy in the translation of both markers in subtitles.



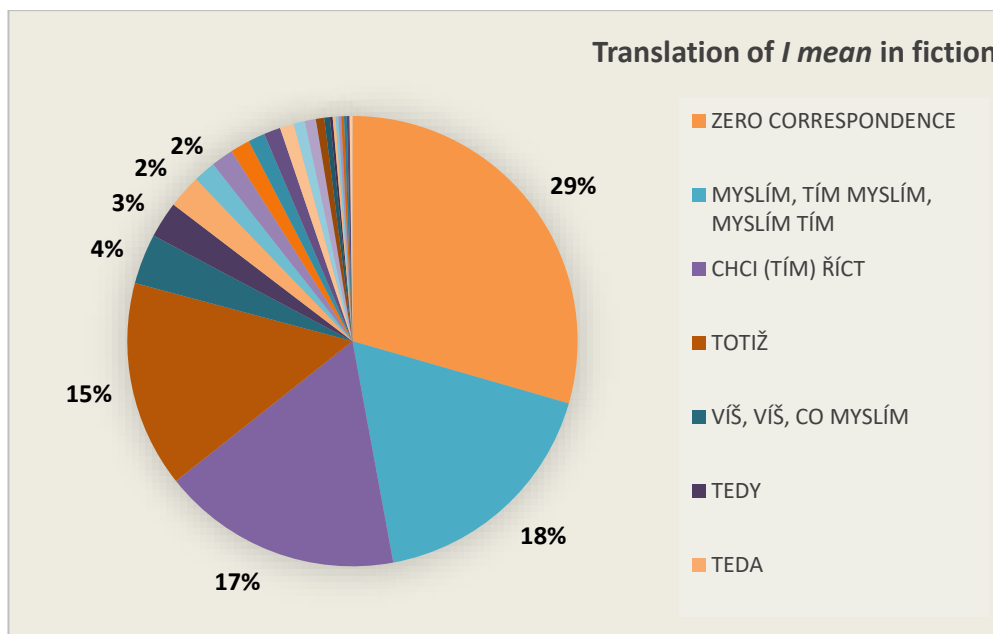
**Graph 1.** Percentage of Czech translation equivalents of *I mean* in the subcorpus of subtitles



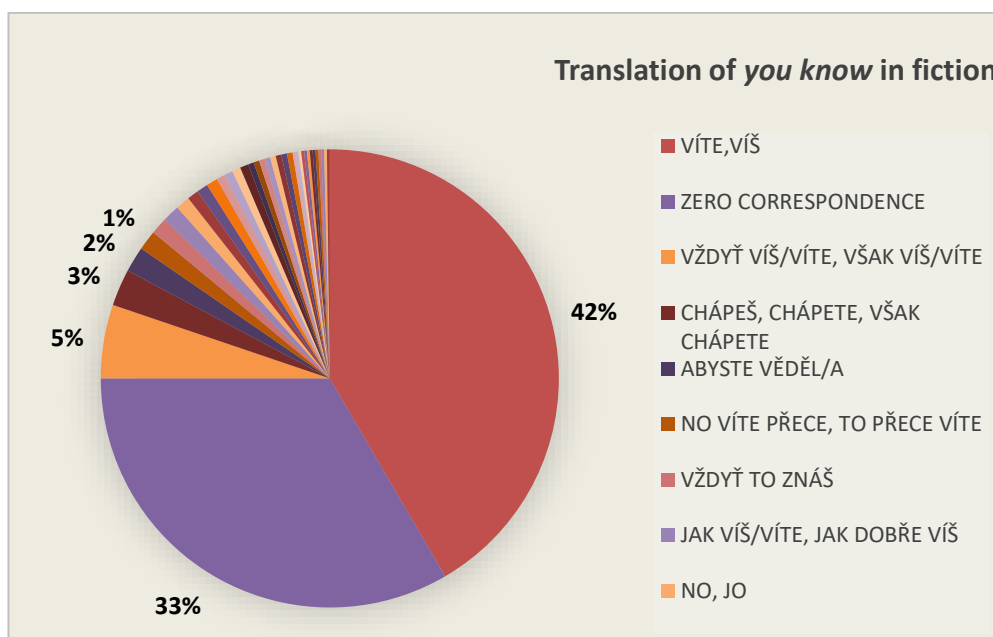
**Graph 2.** Percentage of Czech translation equivalents of *you know* in the subcorpus of subtitles

Due to such a substantial amount of omission of both *I mean* and *you know*, possible reasons for zero correspondence must be considered. First, this strategy is often applied when DMs often have no suitable translation equivalents in the target language. However, this factor can be dismissed in the present analysis because both discourse markers have translation equivalents in Czech. Second factor that can contribute to the frequency of zero correspondence is spatiotemporal constraint that limits the process of translation of subtitles. Audiovisual translation must fit in a certain amount of space and appears on the screen only for a limited amount of time and this may make translators leave out words that are grammatically unnecessary and semantically optional.

This raises question as to how much these two factors influence the translation of *I mean* and *you know*. To determine the extent, the frequency of zero correspondence of these two markers in subtitles can be compared to fiction. The proportions of translation counterparts are summarized in Graph 3 and Graph 4.



**Graph 3.** Percentage of Czech translation equivalents of *I mean* in the subcorpus of fiction



**Graph 4.** Percentage of Czech translation equivalents of *you know* in the subcorpus of fiction

When compared to subtitles, the percentage of omission of DMs lowers almost by a half in fiction. Therefore, it can be concluded that a significant factor contributing to omission is a spatiotemporal constraint in the translation of subtitles.

Nevertheless, even in fiction, zero correspondence was very frequent as it represents a third of translation counterparts of *I mean* and *you know*. Since the spatio-

temporal constraint and clearer audiovisual context as factors responsible for omission are excluded in fiction, the analysis confirms Aijmer and Altenberg's (2001, 33) claim that omission is frequent translator strategy of DMs as they do not contribute to the propositional content of the utterance nor do they affect its truth conditions. This is confirmed by both *I mean* and *you know* in fiction as well as subtitles.

Last but not least, one of the research question examined possible differences between the frequency of zero correspondence between *I mean* and *you know*. Generally, both markers tend to be omitted but a subtle difference is observed. In subtitles, the omission of *you know* is higher by 3% than the omission of *I mean*. Similar tendency is observed in fiction where omission of *you know* is higher by 4% when compared to *I mean*. Although the difference is not large, it can be explained by the primary functions of these markers. Since *I mean* primarily carries textual function, it is more important for addressee's comprehension and thus the omission is slightly lower. On the contrary, *you know* operates more interpersonally as it aims to elicit sympathy or get confirmation, therefore, it is not so crucial for addressee's understanding and the number of zero correspondences is thus larger.

#### **4.5 The Occurrence of *I mean* and *you know* in Direct Speech**

Last but not least, the investigation focused on the proportion of the occurrence of DMs in direct speech in fiction. Although fiction represents writing, what attracted an interest was to what extent these two markers appear in direct speech as indicated by quotation marks. Interestingly, it was found out that *I mean* appears in direct speech in 65% of instances and *you know* in 75% of cases.<sup>15</sup>

The percentage is higher for *you know* as it is primarily interpersonal and is oriented towards the hearer, therefore it appears in majority of cases in direct speech when the speaker directly addresses the hearer. In contrast, *I mean*'s main function is textual and therefore it does not necessarily require an addressee. Instances where *I mean* did not occur in direct speech mostly involved cases where the speaker either

---

<sup>15</sup> It must be pointed out that this number may be even higher. It is because when the DMs appear in direct speech in larger stretches of language, InterCorp gives only the preceding and following sentence and the quotation marks are not displayed in the data.

leads an inner monologue or when he addresses readers and provides explanations as in (96) and (97).

96. *This must have been the case, **I mean**, I must have been in some way conscious, for he would not have had the strength, surely, to get me to my feet unassisted, much less to haul me from the beach to my bedroom door, slung across his back, perhaps, or dragging me by the heels behind him.* [JB\_SEA]

*Muselo to tak být, musel jsem být do určité míry při vědomí, rozhodně by neměl dost síly, aby mě bez pomoci dostal na nohy, natož aby mě dovedl z pláže až do ložnice, nepřehodil by si mě přes rameno, ani by mě neodtáhl za sebou za nohy.*

97. ***I mean**, aside from the ethical questions involved (and my experience is that Peach's girls fret far more about ethics than any other group of people I've ever encountered), what would be the point?* [JA\_CAL]

***Chci říct**, že pokud odhlédneme od etické stránky (a ze zkušeností vím, že Broskviččina děvčata jsou na etiku mnohem háklivější než kterákoliv jiná skupina lidí, jakou jsem kdy poznala), v čem by byl problém?*

The high percentage of occurrence of the DMs in direct speech supports the claim that *I mean* and *you know* are highly characteristic of spoken language and it is confirmed by both markers in both subcorpora. The DMs frequently occur in subtitles, a representative of spontaneous conversation, and appear in high percentage in direct speech in fiction. Furthermore, the fact that *you know* and *I mean*, which are typical of spoken language, but appear in fiction also indicates that spoken and written language cannot be considered as two completely distinct categories with clear-cut borders. Instead, particular writing or speech types should be regarded as continuum with overlapping characteristics.

## 4.6 Syntactic Positions of *I mean* and *you know*

*I mean* and *you know* were approached also from the syntactic point of view, specifically, their syntactic position was analyzed. In addition to this, the investigation also set out to determine whether the syntactic position changes in the Czech translation.

### 4.6.1 The Positioning of *I mean*

Since DMs are not syntactically incorporated in their host utterance, they can appear in any position in the sentence. However, as the table indicates, *I mean* is predominantly placed in a clause-initial position in both fiction and subtitles. In particular, this comprises 80% of instances in fiction and 95% of instances in subtitles which is in striking contrast to clause-medial position in which the marker appears very rarely as there are only 52 (10%) and 14 (3%) cases in fiction and subtitles respectively. In fiction, the marker also appeared in final position, namely in 10% which is in contrast to subtitles where it occurs only in 2%.

Syntactic position of <i>I mean</i>	Initial	Medial	Final	Position in translation	
				Remained	Changed
Fiction	399	52	49	310	43
	80%	10%	10%	88%	12%
Subtitles	478	14	8	211	8
	95%	3%	2%	96%	4%

**Table 8.** The syntactic position of *I mean* in the subcorpora

The syntactic position of *I mean* supports generally held view, which was mentioned earlier, that DMs tend to be positioned initially so that they can restrict context before interpretation of an utterance is determined. This is fully in agreement with *I mean* as its sentence position parallels the primary function of *I mean* which is to introduce a repair or reformulation (98). Moreover, *I mean* placed clause-initially forewarns elaboration or exemplification (99).

98. ***I mean***, that's not illegal. [ZS\_ONB]

*To není zakázaný.*

99. "***I mean***, your husband, Monty, for example, said Kiki, boldly." [ZS\_ONB]

*"Totiž, třeba váš manžel. Monty, například," spustila Kiki směle.*

However, *I mean* functions in a similar way when it is positioned clause-finally. The need of the speaker to demonstrate that he has just provided some modification of his preceding utterance is demonstrated in sentence (100).

100. *After driving over here by yourself, I mean.* [ARS\_HI]

*Teda myslím po tom, co jste se sem rozjela sama.*

As Table 8 indicates, the positioning in fiction follows similar tendency as in subtitles, i. e. *I mean* generally occupies clause-initial place and it is rarely positioned clause-medially. However, a subtle difference concerns final position in fiction where the frequency of the marker is higher by 8% than in subtitles. This raises the question as to what the reason for this tendency can be. This can be possibly explained by the characteristics and differences between the speech and writing. In speech, the hearer is under time pressure and he must process the utterance and arrive at the intended meaning in a given time. Therefore, the marker tends to be placed more frequently initially so as to restrict context before the interpretation is achieved. In contrast, when processing the written language, the addressee is not limited by time, he can reread the text several times or he can return to reread it anytime, thus *I mean* can occupy final position without hindering the interpretation. However, further research would be needed to confirm this.

As regards the position in translation, it must be noted that in majority of cases the syntactic position of the marker remained unchanged in translation. In subtitles, the position of *I mean* remained the same in 96% while in fiction it was not changed in 88% of cases. However, 43 equivalents of *I mean* in fiction and 8 in subtitles changed their placement in Czech Translation. The changes of positioning of *I mean*'s counterparts are captured in Table 9.



The positions of <i>I mean</i> in source texts which were changed in translation	Initial	Medial	Final	Total
Fiction	12	0	31	<b>43</b>
Subtitles	5	0	3	<b>8</b>

**Table 9.** The positions of *I mean* in source texts which were changed in translation

When *I mean* occupied clause-final position in source text, the position of its counterpart was changed in 31 instances in Czech translation. The position is usually shifted to clause-initial or medial position in Czech, as it is shown in (101) and (102), even though the grammatical status of the Czech translation equivalent does not require it. In other words, the counterpart can be placed freely in the sentence but it is shifted to initial or medial position in translation.

101. “If the guy bring it back, ***I mean***.” [ARS\_HI]

“***Tedy*** pokud ho ten člověk přinese zpátky.”

102. “That’s impossible, ***I mean***.” [CP\_CHO]

“Ono to ***totiž*** nejde.”

In several cases, *I mean*’s change of position is caused by a grammatical status of Czech translation equivalent, especially when the translation involved verbs *chci říct* and *myslím* which were complemented by some element. Namely, since in (103) *I mean* is translated as superordinate clause followed by a subordinate clause, its translation is restricted to initial. Similarly, *myslím* in (104) is complemented by a subordinate clause therefore it must be placed initially.

103. “Stick insects do, ***I mean***.” [TPNG\_GOO]

“Ne, ***já chtěl říct***, že ptačím zobem se živěj pakobylinky.”

104. About how to do it, ***I mean***...[TM\_C]

***Myslím*** jak se to dělá...

## 4.6.2 The Positioning of *you know*

Unlike *I mean*, the positioning of *you know* does not adhere to the general tendency of DMs to be placed clause-initially. *You know* predominates in clause-final positions and it accounts for 46% of instances in fiction and 48% in subtitles. In fiction, the proportion of *you know* in clause-initial and clause-medial is the same, covering 27% of cases. Similar position is observed in subtitles where *you know* appears in clause-initial position in 22% and in clause-medial position in 30%.

Syntactic position of <i>you know</i>	Initial	Medial	Final	Position in translation	
				Remained	Changed
Fiction	136	135	229	274	59
	27%	27%	46%	82%	18%
Subtitles	110	151	239	195	11
	22%	30%	48%	95%	5%

**Table 10.** The syntactic position of *you know* in the subcorpora

Since the marker performs primarily interpersonal function, its syntactic position is not crucial for correct understanding of the utterance. Put differently, *you know* can be placed more freely in the sentence, unlike *I mean*, as it does not forewarn upcoming modification. Nevertheless, the most frequent position is clause-final which indicates that the marker is mostly used to elicit sympathy or receive confirmation of understanding after an utterance was made. Similarly, the marker in clause-initial position can be used to achieve the same or it can be used to attract hearer's attention.

Though the Czech translation counterparts of *you know* can occupy initial, medial and final position, the equivalents generally remained in the same clause position as in English sentences. The position in subtitles remained the same in 95% of cases while in fiction it remained in 82%. But a certain number of cases which involved changes of position in translation emerged and these are depicted in Table 11.

The positions of <i>you know</i> in source texts which were changed in translation	Initial	Medial	Final	Total
Fiction	7	17	35	<b>59</b>
Subtitles	2	4	5	<b>11</b>

**Table 11.** The positions of *you know* in source texts which were changed in translation

Like *I mean*, the change of position involved mostly cases where *you know* in English appeared clause-finally and this position was shifted to clause-initial in Czech.

105. *This is a very odd gun, **you know**.* [TPNG\_GOO]

*Poslyš, tohle je opravdu podivná zbraň.*

106. *In an ideal world, **you know**, there'd be a sunflower growing in this.*  
[OUTCASTS\_10]

*Víš, v ideálním světě by tady rostla slunečnice.*

Changes of position involved such cases instances where *you know* was in source texts placed clause-finally and this position shifted to clause-initial position in Czech. This shift was sometimes triggered by grammatical status of the translation because *you know* was translated as a main clause to which a subordinate clause was connected (107) and (108).

107. *He was very ill, **you know**, and he looked quite small and pathetic, he looked like a little bourgeois who has come out to collect he newspaper or call the cat.* [LB\_CAP]

*Sami víte, že tehdy už byl těžce nemocný, vypadal tak připosraženě a dojemně, zkrátka jako měšťák, který si přišel vyzvednou noviny nebo přivolat kočku.*

108. *We have Spanish lessons here, **you know**.* [JG\_BRE]

*Víš přece, že tu máme kurzy španělštiny.*

But there occurred instances where the translation equivalent was syntactically independent and the position of translation equivalent was still shifted to initial (109).

109. “*It’s almost like living it all over again, you know,*” she says, not turning around. [SK\_CAR]

“*Víte, je to, jako bych to všechno znovu prožívala,*” říká a zůstává ode mě odvrácena.

It can be concluded that *I mean* is characteristically placed in clause-initial position and this placement adheres to its primary function to introduce upcoming repair or reformulation and this also confirms the tendency to position DMs initially. This is not supported by *you know* which often appears in clause-final position where it is directed at eliciting hearer’s sympathy or confirmation of understanding. Their translation equivalents mostly occupied the same positions and in many cases where the position changed this was due to the grammatical status of the Czech translation.

#### 4.7 Co-occurrence of *I mean* and *you know* with Other Expressions

In order to extend our understanding of *I mean* and *you know*, the investigation was directed towards their co-occurrence with other DMs. Only elements placed adjacently to the DM were considered and the results are captured in Table 12.

Clustering		<i>Well</i>	<i>Oh, well</i>	<i>I mean</i>	<i>You see</i>	<i>You know</i>	Total
<i>I mean</i>	Fiction	5	1	xxx	1	1	8
	Subtitles	7	0	xxx	0	8	15
<i>You know</i>	Fiction	15	0	0	0	xxx	15
	Subtitles	35	2	6	0	xxx	43

**Table 12.** Co-occurrence of *I mean* and *you know* with other discourse markers

The clustering of both markers was much more common in subtitles than in fiction. This can be attributed to the characteristics of conversation which is spontaneous and unplanned so the speaker employs more DMs as gap fillers or as hesitation markers. This is in correspondence to the meaning of the marker *well* with which both *I mean* and *you know* co-occur with the highest frequency. In example (110), the markers *well* and *you know* function as verbal fillers as the speaker needs to gain some time while searching for the suitable words to express precisely what he intends whereas in (111) the DMs indicate speaker's tentativeness and hesitation to express exactly what he wants.

110. *Hey, Velma, I never meant to, well, you know... pick you last.* [SCOOBY DOO\_02]

*Velmo, nikdy jsem nechtěl... vždyť víš... nechat tě až na konec.*

111. *Well, you know, I've been thinking.*

*Víte, tak trochu jsem přemýšlel.* [CIRQUE\_09]

Though not to such an extent as with the marker *well*, both markers tended to cluster with each other. Altogether, *I mean* was placed adjacently to *you know* 14 times in subtitles. This supports Schiffrin's claim (1987) that the functions of *I mean* and of *you know* are complementary which is demonstrated in the following examples.

112. *You and me, I mean, you know... [G-FORCE\_09]*

*Myslím tím nás dva...*

113. *You know, I mean, look, you've already opened, and I could help you close.* [SEX DRIVE\_08]

*Víš, chci říct, koukni, už jsi to s ní začal a já ti to pomůžu ukončit.*

Sentence (112) shows that the speaker specified or modified his intention with the marker *I mean* and then used *you know* to elicit confirmation or sympathy from the hearer. Likewise, by using *you know* in (113), the speaker aims to attract hearer's attention and make sure he understands the utterance which reformulation is introduced by *I mean*.

Interestingly, in the analysis it was found out that *I mean* frequently co-occured with *no*, *yes* or more colloquial *yeah*. In subtitles, *I mean* cluster with these expressions 17 times and in fiction 14 times. The same tendency is not observed in the case of *you know* which is captured in the Table 13.

Co-occurrence with expressions of agreement or disagreement			
<i>I mean</i>		<i>You know</i>	
Fiction	Subtitles	Fiction	Subtitles
14	17	1	3

**Table 13.** Co-occurrence of *I mean* and *you know* with expressions of agreement or disagreement

The clustering of *I mean* with *yes*, *yeah* and *no* again corresponds with the primary function of this marker. The instances where *I mean* co-occurs with expressions of agreement or disagreement reflect speaker's intention to modify his intention or propositional information, two main functions mentioned by Schiffrin (1987). Sentences (114) and (115) show that the speaker's utterance was not correctly understood by the hearer and he reformulates or specifies it in order to be understood.

114. "No, ***I mean***," said Levi, as they reached a landing and a damaged door, patched up with a panel of unpainted wood, "pay a call." [ZS\_ONB]

"Ne, ***já jsem myslel***," vysvětloval Levi, když dorazili na podestu, k poškozeným dveřím vyspraveným kusem nenatřeného dřeva, "že jsem přišel na skok. Jako na chvíli na návštěvu."

115. "No, ***I mean***, I can't believe anybody would read my Guide and actually follow its advice, can you? [DL\_DEA]

"Ne, vážně, nechce se mi věřit, že by si rádce někdo přečetl a pak se ho slovo od slova držel, co myslíte?"

## CONCLUSION

The aim of the present thesis was to analyze discourse markers *I mean* and *you know* in the multilingual translation corpus InterCorp. Cross-linguistic research was selected for the objective of this thesis in order to find out what the translation of DMs can show about their meaning and functions. In particular, the investigation focused on Czech translation equivalents in the subcorpora of fiction and subtitles which were regarded as a representative of writing and speech respectively and altogether 2,000 instances of these two DMs were examined with respect to their translation equivalents, syntactic position and clustering with other linguistic expressions. Attention was also paid to the occurrence of *I mean* and *you know* in direct speech in fiction.

The theoretical part dealt with approaches to study DMs and explained the preference for the term DMs and clarified why this thesis follows Schiffrin's (1987) view of DMs. Since *I mean* and *you know* conform to her understanding of DMs as being syntactically detachable, multifunctional and typically placed in clause-initial position, they were chosen for this study as they also follow other properties characteristic of DMs which were described subsequently. These properties, namely, syntactic and semantic optionality, free placement in the sentence and lack of propositional content, then served as a basis for criteria applied to distinguishing discourse marker use from literal use of these expressions which were clarified in the methodological part. These two markers were also selected for this study as they are both characteristic of spoken language, their functions are complementary and are syntactically similar.

The research shows that the translations of DMs can capture different shades of their meaning which is especially important finding as it extends our knowledge of *I mean* and *you know*. It has been identified that *I mean*'s basic function is mirrored by all the Czech translations as they are oriented to speaker signalling modification of his intentions or propositional information, reformulation or explicitness. This is confirmed by two *I mean*'s most common translation equivalents which were *myslím* and *chci říct* sometimes complemented by the pronoun *tím*. In fiction, *myslím* represented 17,6% of instances and *chci říct* 17,2% of instances and the results were similar in subtitles where *I mean* was translated as *myslím* in 16,4% and as *chci říct* in 10,6% of cases. The same function was performed by the particles *totiž*; *vlastně*; *teda* and *tedy* which were also very frequent. Interestingly, translations such as *viš*, *co myslím*; *viš/víte*;

*rozuměj/rozumíš; pochop/pochopte, or víš, myslím si* which aim at eliciting addressee's sympathy or understanding occurred in the data as well. This suggests that *I mean* can also be oriented to the hearer and can perform interpersonal function which confirms one of the properties of discourse markers which is that they are multifunctional.

The translation counterparts of *you know* suggest that this marker primarily performs interpersonal function. Except for marginal cases which involved translation of this marker as interjection, the interpersonal function manifested itself in the Czech translations as all of the counterparts were oriented to addressee. And while the translation of *I mean* was evenly distributed into two or three equivalents, the translation of *you know* was most frequently represented by a translation equivalent *víš/víte* which accounted for 41,6% in fiction and 28,6% in subtitles. However, the verb *vědět* was very often complemented by *vždyt'; však; co; no* or was in subordinate clause and when all such instances were counted, the percentage rose to 52,6% in fiction and 33% in subtitles. Moreover, translation counterparts *hele/heleď* and *poslyš* clearly indicated the function of *you know* as attention-catcher. Comprehension-securing function of this marker was demonstrated in the cases where *you know's* counterparts were represented by question tags such as *že jo?; že ano?; co?; že? and ne?* which were aimed at eliciting hearer's confirmation of understanding. The interpersonal function of *you know* also manifested itself in translations such as *vždyt' víš* or *však víš* suggesting previous shared knowledge between the speaker and the hearer.

The most frequent translation counterparts of *I mean* as well as *you know* were the same in fiction as well as subtitles except for one difference concerning translation equivalents of *I mean*. The translation counterpart following *chci říct* and *myslím* was *totiž* in fiction (representing in 75 instances) but in subtitles next frequent equivalent was *teda* which was found in 20 instances while *totiž* was in subtitles represented by only 3 tokens. *Teda* was not even so common in fiction as it represented only 12 tokens. According to *A Dictionary of Standard Czech Language* (Havránek et Al. 2011), *teda* is a colloquial expression and since subtitles as a representative of spontaneous speech are less formal than fiction, it is therefore understandable that the occurrence of *teda* is higher there. Other difference between speech and writing concerns the fact that when translations in fiction and subtitles were compared, the research revealed that there is a greater variety and complexity of translation equivalents in fiction than in subtitles where translation counterparts generally consisted of single-word items. The tendency



was attributed to the fact that the translation of subtitles is constrained by time and space which makes translators opt for shorter equivalents.

The research also reports a high proportion of zero correspondence which supports the claim that DMs tend to be omitted as they do not affect the truth conditions of an utterance and carry no propositional meaning. In fiction, omission represented 29% of translation equivalents of *I mean* and for *you know*, zero correspondence represented 33%. The number of zero correspondence was larger for both markers in subtitles than in fiction and time and space pressure in audio-visual translation was identified as a contributing factor. In subtitles, the omission of *I mean* reached 56% and in case of *you know* it was 59%. The omission was altogether higher for *you know* than for *I mean*, although the difference was subtle. The difference was attributed to the functions of each marker; while *I mean* has primarily textual function which is more important for the interpretation and comprehension of an utterance, *you know*'s interpersonal function does not play such a crucial role in comprehension, thus it tended to be omitted more frequently.

The study also focused on DMs in fiction and revealed that *I mean* is found 327 times (65%) in direct speech and *you know* 373 times (75%). Therefore, this confirmed that DMs are typical of speech as they occurred in fiction in direct speech and also supported the claim that linguistic characteristics of speech and writing overlap.

The results of this study indicate that primary functions of *I mean* and *you know* affect the syntactic position of these two DMs. Namely, *I mean* occupied predominantly clause-initial position in 88% of instances in fiction and 96% in subtitles and that is in accordance with the basic function of the marker which is to signal upcoming modification or reformulation. At the same time, the position of *I mean* confirmed the tendency for DMs to be positioned clause-initially as it facilitates the processing of the utterance.

The same can be said about *you know* as its positioning corresponded to the marker's interpersonal function. *You know* in clause-final position was used to elicit confirmation of understanding or elicit sympathy after an utterance was made and the marker occurred in this position in 46% in fiction and 48% in subtitles. To capture addressee's attention but also to get his understanding, *you know* was placed clause-initially covering 27% of all occurrences in fiction and 22% in subtitles. Not uncommon was clause-medial position (27% in fiction and 30% in subtitles) which indicated the speaker's hesitation what to say next but also confirmation seeking.

In general, it was observed that the syntactic position of DMs has a tendency to remain the same in Czech translation. As regards *I mean*, its position remained unchanged in 88% in fiction and 96% in subtitles and *you know*'s position remained the same in 82% in fiction and 95% in subtitles. Cases where the position of markers changed were marginal but involved clause-final or medial occurrences of DMs which were shifted to clause-initial position in Czech. This was often due to the grammatical status of Czech translation as the markers were translated as superordinate clauses and a subordinate clause was related to them. However, further study would be needed in order to find out whether this is not due to translation effects.

Another aspect where the function of DMs manifested itself concerned co-occurrences of DMs with other expressions. Altogether the clustering was more frequent in subtitles suggesting the fact that DMs are characteristic of spoken language. Both *I mean* and *you know* clustered with the marker *well* and in such sentences the combination of markers functioned as hesitation markers or verbal fillers. Although not to such an extent, *I mean* and *you know* also appeared together which suggested the complementariness of their functions as their combination indicated seeking for understanding after a modification or specification had been made.

Unlike *you know*, *I mean* clustered with expressions of agreement or disagreement. The combinations of *I mean* with *yes*, *yeah* or *no* represented 14 instances in fiction and 17 instances in subtitles and the co-occurrences again indicated the primary function of the marker. Cases where *I mean* clustered with *no* showed that the speaker had not been correctly understood and thus he disagreed with the hearer and then presented reformulation of his previous utterance or exemplification. Similarly, sentences where *I mean* occurred with *yes* or *yeah* this was followed by exemplification or explicitness.

Another interesting finding emerged from the closer analysis of Czech translation counterparts of the DMs. In some cases, Czech equivalents were syntactically independent and did not carry any complement thus their omission would not result in ungrammaticality and they were semantically optional. But there were translations of *I mean* and *you know* which did not possess the same properties as English DMs. In particular, there were instances when *I mean* and *you know* were translated as a clause, which consisted of covert subject and verb, and were accompanied by complements *tím* or *co*, adverbs *vždyť* or *však* or were connected to a subordinate clause introduced by conjunction *že*. In such cases, translation equivalents

could not be omitted as it would affect the grammaticality of the sentence. Moreover, their deletion would also affect the truth conditions of an utterance as they carried propositional information.

All these findings enhance our understanding of *I mean* and *you know* in speech and writing. The research also extends our knowledge of their use and their meaning as seen through their translation counterparts. However, since this study was limited by the absence of a parallel corpus of spoken English aligned with Czech, subtitles were regarded as representative as speech. But their translation is influenced by time and space thus further study could assess DMs and their Czech translation in speech which translation is not constrained at all. Another limitation of the study was that Czech, as a target language, could reflect features of English. This could have impacted the syntactic position in translation. Therefore, in order to exclude possible translation effects, it would be interesting to examine syntactic position of DMs in Czech original texts.

## RESUMÉ

Tato diplomová práce se zabývá tzv. „discourse markers”, tj. diskurzivními ukazateli či markery. Předmětem je analýza dvou diskurzivních ukazatelů *I mean* a *you know* a jejich českých překladových ekvivalentů v překladovém korpusu InterCorp, konkrétně pak v korpusu beletrie a v korpusu titulků zastupující psaný a mluvený jazyk. Zároveň je pozornost věnována jejich pozici ve větě a výrazům, se kterými se často objevují. Tato práce také zkoumá v jakém poměru se *I mean* a *you know* objevují v přímé řeči v beletrii.

Teoretická část objasňuje terminologii, přístupy k diskurzivním ukazatelům a také popisuje vlastnosti diskurzivních ukazatelů. Přestože toto téma bylo v posledních desetiletích populární a bylo o něm napsáno mnoho, stále panuje neshoda týkající se nejen terminologie, ale i toho, které výrazy vlastně do této skupiny patří. Existuje nespočet termínů, které se používají pro tyto výrazy, avšak tato práce užívá termín „discourse marker”, protože je velmi často preferovaný a uznávaný a mnoho autorů, kteří prvně užívají jiný název, nakonec začnou preferovat tento. Zároveň je tento termín aplikován Schiffrinovou, jejíž pojetí diskurzivních markerů tato diplomová práce následuje.

Různá pojetí a odlišné přístupy k diskurzivním markerům jsou uvedeny a porovnány v kapitole 1.2, která také specifikuje důvody, proč se pro tuto studii jako nejlepší jeví koncepce od Schiffrinové. Schiffrinová (1987, 328) považuje diskurzivní ukazatele za syntakticky odlučitelné výrazy, které jsou nejčastěji umístěny na začátku věty, jsou multifunkční a operují na lokální i globálních úrovni diskurzu. Do kategorie diskurzivních ukazatelů zahrnuje i ty výrazy, které fungují na tzv. participation plane, to znamená, že mohou být orietovány na mluvčí či na posluchače, což platí pro *I mean* a *you know*. Z tohoto důvodu tato práce upřednostňuje pojetí Schiffrinové.

Co se týče vlastností diskurzivních markerů, ty jsou specifikovány v kapitole 1.3. Přestože i v této oblasti se názory autorů poměrně liší, panuje mezi autory shoda, že se jedná o výrazy z různých slovních druhů, čili diskurzivní ukazatelé tvoří syntakticky heterogenní skupinu. Brinton (1996, 33-35) pak předkládá shrnutí vlastností: „diskurzivní ukazatele jsou běžnější v mluveném projevu, jsou často stylisticky stigmatizovány a negativně hodnoceny, jsou krátké a často foneticky redukováné, často se objevují v první pozici ve větě nebo se mohou objevit pouze na tomto místě, stojí mimo syntax věty nebo jsou k větě volně připojeny a nemají žádnou gramatickou

funkci, nejsou povinné, nesou málo nebo žádný propoziční význam, jsou multifunkční a mohou působit jak na lokální tak na globální úrovni.“ Tyto vlastnosti jsou poté podrobně rozebrány.

Tato studie se zaměřuje na *I mean* a *you know*, protože oba výrazy mají tyto vlastnosti. Dalšími důvody pro výběr těchto dvou výrazů je skutečnost, že jsou syntakticky podobné, jsou charakteristické pro mluvený jazyk a mnoho autorů zkoumá oba tyto markery společně, protože se jejich funkce doplňují. Zároveň se autoři shodují, že tyto markery jsou si podobné v tom, že jejich funkce vychází z jejich doslovného významu. Dle Schiffrinové (1987, 267) *I mean* je orientováno na mluvčího a to konkrétně na modifikaci jeho záměrů či propozičního významu předchozí věty, opravu nebo explicitaci významu jeho projevu, zatímco *you know* je orientováno směrem k posluchači a plní spíše interpersonální funkci, protože mluvčím jím chce získat pozornost či porozumění od posluchače nebo získat potvrzení o pochopení. Nicméně, bližší klasifikace funkcí je komplikovaná, protože hranice mezi jednotlivými kategoriemi funkcí často nejsou jasně vymezené a autoři často nerozlišují mezi užitím výrazů jako diskurzivní ukazatele od jejich doslovného užití. Navíc ve většině studií není jasné, jak byly funkce určeny, v mnoha případech mohlo jít o pouhé odhadnutí a často je také přiřazování funkcí problematické, jelikož lingvisti nemohou číst mysl mluvčímu (Huang 2011, 29), a tak je určování funkcí často subjektivní (Scheler a Fischer 1997, 668). Z těchto důvodů, se tato studie nezaměřuje na bližší přiřazování funkcí *I mean* a *you know*, ale spíše na jejich překlad.

Kapitola 1.2.1 pojednává o mezijazykových studiích a jejich přínosu pro diskurzivní ukazatele. Překlad může odhalit význam diskurzivních markerů a také určit jejich funkce. Proto tato studie analyzuje na překladové protějšky *I mean* a *you know* a zjišťuje, co vyjadřují o jejich funkci či významu. Jedna možnost při překladu představuje tzv. nulovou korepodenci, což podle Johanssona (2007, 58) poukazuje na případy, u kterých v překladu nenalezneme formální mezijazykovou korespondenci. Aijmerová a Altenberg (2001, 33) se pak zabývají důvody pro nulovou korespondenci a jako jednu z příčin uvádějí případy, kdy cílový jazyk neobsahuje vhodné překladové ekvivalenty či nulová korespondence může poukazovat na mezijazykové rozdíly. Dalším důvodem pro vynechání adverbciálních konektorů v překladu je skutečnost, že kontext je dostatečně jasný a posluchač si může vztah mezi větami vyvodit (ibid. 22). Nicméně i v případech, kdy překladový jazyk obsahuje protějšky, nulová korespondence je častá a důvodem je podle Aijmerové a Altenberga (2001, 22) fakt, že

diskurzivní ukazatele neovlivňují pravdivostní hodnotu věty a nenesou propoziční význam.

Protože si jedna z výzkumných otázek klade za cíl zjistit, jestli jsou nějaké rozdíly mezi překladovými protějšky v psaném a mluveném jazyce, v kapitole 2 je pozornost věnována i jazykové charakteristice těchto dvou forem komunikace. Jelikož se charakteristika různých typů psané a mluvené řeči překrývá, spíše než jako na dvě pevně oddělené komunikační formy řeči, je vhodnější nahlížet na ně jako na kontinuum s „typickým“ psaným projevem na jednom konci a s „typickým“ mluveným projevem na konci druhém (Leech et al. 1982, 140). Bližší charakteristika psaného a mluveného projevu pak vychází z tohoto pojetí. Co týče diskurzivních markerů, ty jsou běžnější pro ústní projev, protože ten bývá spontánní a neplánovaný, a tak mluvčí často využívají například *I mean* pro opravu či úpravu jejich záměrů a předchozích informací. Mluvená řeč je také méně formální a při každodenní komunikaci se spíše využívá pro udržování mezilidských vztahů než k přenosu informací, což je spíše typické pro psanou řeč (Brown and Yule 1983, 13). K tomuto účelu mimo jiné mohou posloužit diskurzivní ukazatele, *you know* může být například užito k získání posluchačovi sympatie, přilákání jeho pozornosti či získání jeho potvrzení o porozumění.

Popis zvolené metodologie přináší kapitola 3. Pro výzkum byla použita metoda kontrastivní analýzy a data byla převzata z paralelního korpusu InterCorp, který je dostupný na <http://ucnk.ff.cuni.cz/intercorp/> a obsahuje velké množství textů zarovnaných s češtinou. Prvně byl vytvořen subkorpus titulků po roce 2000 se zdrojovým jazykem angličtina a také subkorpus beletrie po roce 1960. Data musela být poté protříděna, protože obsahovala tyto dva výrazy v jejich doslovném významu. Jako diskurzivní užití *I mean* a *you know* byly považovány výrazy, jejichž vynechání neovlivnilo gramatiku věty a navíc po jejich odstranění pravdivostní hodnota věty zůstala stejná. Zároveň diskurzivní ukazatel nebyl doplněn podstatným jménem, přídavným jménem, infinitivem či vedlejší větou a ani se neobjevoval spolu s pomocnými slovesy. Diskurzivní ukazatel mohl být umístěn volně, protože nebyl syntakticky propojen s větou. Také byly vyřazeny případy, kdy byl daný diskurzivní ukazatel součástí jiného, konkrétně se jednalo o tyto výrazy: *what I mean is, by this/that I mean, you know/see what I mean, as you know, if you know, you know what*. Celkem bylo vybráno 500 příkladů *I mean* a 500 *you know* v subkorpusu titulků a to samé v subkorpusu beletrie, dohromady tedy bylo zanalyzováno 2 000 příkladů. Metodologická část také přináší problémy při analýze a je zakončena výzkumnými otázkami.

Kapitola 4 rozebírá výsledky analýzy. Tato část se nejprve zabývá překladovými ekvivalenty *I mean* v beletrii a v titulcích. Nejčastějšími českými ekvivalenty jsou *myslím* (17,6% příkladů v beletrii a 16,4% v titulcích) a *chci říct* (17,2% příkladů v beletrii a 10,6% v titulcích), které jsou někdy doprovázeny zájmenem *tím*. Často je *I mean* překládáno jako *totiž* (zastupující 14,8% překladu v beletrii ale pouze 0,6% v titulcích) a *teda* (2,4% v beletrii a 4% v titulcích). Nejčastější české ekvivalenty odrážejí význam *I mean*, protože směřují na mluvčího a na modifikaci jeho záměrů či propozičního významu věty. Překlady také obsahovaly ekvivalenty jako *viš, co myslím; viš/víte; rozuměj/rozumíš; pochop/pochopte* nebo *viš, myslím si*, které naznačují, že tento marker může také plnit interpersonalní funkci, protože tyto překlady míří na získání posluchačova porozumění a sympatie.

Stejně tak bylo zjištěno, že překladové protějšky *you know* vyjadřují význam tohoto diskurzivního ukazatele a naznačují, že tento marker plní interpersonalní funkci, protože jeho nejčastější překlady jsou orientovány směrem k poslouchači a to zejména na získání jeho potvrzení o porozumění, pozornosti a sympatie. Nejčastěji byl přeložen do češtiny jako *viš/víte* (41,6% v beletrii a 28,6% v titulcích), nicméně pokud přičteme překladové ekvivalenty, kdy *viš/víte* bylo doprovázeno slovy *vždyt'; však; co; no* nebo byl přeložen jako vedlejší věta, vzrostou tyto čísla na 52,6% v beletrii a 33% v titulcích. *You know* s účelem získání posluchačova porozumění bylo zřejmé v případech, kdy tento marker byl přeložen jako tázací dovětek *jo?; že ano?; co?; že? and ne?*. Překladové ekvivalenty *hele/heleď* a *poslyš*, které se v datech také objevily, ačkoli ne v takovém poměru jako *viš/víte*, naznačují, že tento ukazatel může sloužit k přilákání posluchačovy pozornosti. Překlady obsahující citoslovce, jako například *ehm* or *no*, poukazovaly na užití ukazatele za účelem získat více času pro rozmyšlení, co bude následovat, či naznačovaly zaváhání mluvčího.

Dalším zajímavým zjištěním bylo, že české překladové ekvivalenty se vždy nechovají stejně jako *I mean* a *you know*. V některých případech mají české protějšky stejné vlastnosti jako anglické diskurzivní ukazatele, tj. nejsou syntakticky připojeny ke větě, mohou být vynechány, aniž by to ovlivnilo význam či gramatiku věty, nicméně přispívají k její interpretaci. Ovšem vyskytly se i překlady, kdy *I mean* a *you know* v českém překladu byly spojeny s vedlejší větou či nesly doplněk, tj. nebyly přeloženy jako diskurzivní ukazatele, ale spíše se jednalo o doslovné užití těchto markerů.

Překlady těchto diskurzivních ukazatelů se v subkorpusu beletrie a titulků příliš nelišily a nejčastější české ekvivalenty pro *I mean* a *you know* byly stejné v obou

subkorpusech. Nicméně beletrie obsahovala daleko více překladových ekvivalentů a ty také byly komplexnější a různorodější narozdíl od titulků, které zahrnovaly spíše krátké a jednoslovné ekvivalenty. Jako vysvětlení se nabízí skutečnost, že překlad titulků je omezený časem i prostorem, a proto překladatelé musí volit kratší výrazy.

Tento faktor je zároveň jednou z příčin nulové korespondence, která byla častá pro oba markery. V titulcích nulová korespondence představovala 56,2% případů u *I mean* a 58,8% u *you know*. Jelikož byla v beletrii nižší (29,4% příkladů pro *I mean* a 33,4% pro *you know*), potvrdilo se, že časové a prostorové omezení při překladu titulků je jeden z důvodů pro vynechání diskurzivních ukazatelů. Protože byl poměr nulové korespondence vysoký i v beletrii, potvrdil se výrok od Aijmerové a Altenberga (2001), kteří tvrdí, že nulová korespondence je častá u diskurzivních ukazatelů, protože nenesou žádný propoziční význam a nepřispívají k pravdivostní hodnotě věty.

Výzkum se soustředil i na syntaktickou pozici *I mean* a *you know* a bylo zjištěno, že jejich umístění ve větě je ovlivněno jejich významem a funkcí. *I mean* převážně zaujímal iniciální pozici (80% v beletrii, 95% v titulcích), což je v souladu s primární funkcí tohoto markeru, protože naznačuje nadcházející modifikaci či opravu. Zároveň pozice *I mean* potvrzuje jednu z vlastností diskurzivních ukazatelů, podle které jsou tyto výrazy umístěny na počáteční pozici ve větě. *You know* sice nebylo umístěno v takovém poměru na iniciální pozici, avšak i jeho umístění ve větě je v souladu s funkcí tohoto markeru. Nejčastěji bylo na finální pozici (46% v beletrii, 48% v titulcích), kdy bylo použito pro získání posluchačova potvrzení o porozumění či získání jeho sympatie poté, co mluvčí učinil výrok. *You know* v iniciální pozici (27% v beletrii, 22% v titulcích) bylo použito převážně k získání pozornosti, ale také k získání posluchačova porozumění. *You know* v mediální pozici (27% v beletrii, 30% v titulcích) pak umožnilo mluvčímu získat čas k plánování toho, co bude v rozhovoru následovat, či tento marker v mediální pozici indikoval zaváhání mluvčího.

V překladu pak umístění *I mean* a *you know* zůstalo nezměněné ve většině případů (*I mean* zůstalo na stejné pozici v 88% v beletrii a 96% v titulcích, zatímco *you know* se nezměnilo v 82% v beletrii a 95% v titulcích). Nicméně toto mohlo být ovlivněno efekty překladu. Příklady, kdy se pozice markerů v překladu změnila, se týkaly především mediální a finální pozice, které v překladu byly změněny do iniciální pozice. Toto bylo často způsobeno syntaktickou strukturou českého ekvivalentu, protože pokud *I mean* či *you know* bylo preloženo jako hlavní věta, na které byla závislá věta vedlejší, musel být český ekvivalent umístěn v pozici iniciální.



Bylo zjištěno, že funkce a význam těchto dvou diskurzivních ukazatelů má vliv i na výrazy, se kterými se často objevují. *I mean* se pojilo s výrazy *yes*, *yeah* and *no* (14 příkladů v beletrii a 17 příkladů v titulcích), což také indikuje význam tohoto markeru, kdy mluvčí nebyl správně pochopen, a tak vyjadřuje souhlas či nesouhlas a poté následuje oprava signalizovaná markerem *I mean*. Oba markery se často objevují s diskurzivním ukazatelem *well* a takové případy naznačovaly potřebu mluvčího získat čas pro plánování nebo mluvčího zaváhání. *I mean* a *you know* se také vyskytovaly společně, což naznačuje jejich vzájemné doplňování funkcí. Celkově byla kombinace *I mean* a *you know* s dalšími diskurzivními markery častější v titulcích, což potvrzuje skutečnost, že tyto výrazy jsou charakteristické pro mluvený projev.

Tento fakt je potvrzen i tím, že *I mean* a *you know* se v beletrii ve velké většině vyskytovalo v přímé řeči, tj. 65% *I mean* v beletrii bylo umístěno v přímé řeči a pro *you know* je toto číslo ještě vyšší, protože se jednalo o 75% případů. Jelikož se tímto potvrdilo, že diskurzivní ukazatelé, které jsou typické pro mluvený projev, se objevují často i v psaném projevu, toto zjištění zároveň podporuje výrok Leech et al. (1982, 140), kteří tvrdí, že charakteristika mluvené a psané řeči se překrývá.

## APPENDICES

### Appendix 1: Subcorpus of Fiction from InterCorp

Where the bibliography of the books cited in this thesis is incomplete in the InterCorp metadata, their first English and Czech editions are given.

- [DA\_RES] Adams, Douglas. 1980. *The Restaurant at the End of the Universe*. Translated by Jana Hollanová as *Restaurant na konci vesmíru*. 1999. Praha: Hynek.
- [JA\_CAL] Angell, Jeannette. *Callgirl: Confessions of an Ivy League Lady of Pleasure*. 2004. Translated by anonymous as *Dvojí život: Když se den změní v noc*. 2006. Praha: Ikar.
- [IB\_SON] Banks, Iain. 1998. *A Song of Stone*. Translated by Pavla Horáková as *Píseň kamene*. 2002. Praha: Volvox Globator.
- [JB\_SEA] Banville, John. *The Sea*. 2005. Translated by Richard Podaný as *Moře*. 2006. Praha: Picador.
- [LB\_CAP] Bernières, Louis de. 1998. *Captain Corelli's Mandolin*. Translated by Viktor Janiš as *Mandolína kapitána Corelliho*. 2002. Praha: BB Art.
- [SB\_HE] Brown, Sandra. *Hello, Darkness*. 2003. Translated by Marie Válková as *Zdravím tě, temnoto*. 2005. Praha: Ikar.
- [CD\_CIR] Day, Cathy. *The Circus in Winter*. 2004. Translated by Milena Pellarová and Šimon Pellar as *Cirkus v zimě*. 2005. Praha: Ikar.
- [GD\_BIR] Durrell, Gerald. 1969. *Birds, Beasts and Relatives*. Translated by Zora Wolfová as *Ptáci, zvířata a moji příbuzní*. 1979. Praha: Mladá fronta.
- [DF\_SLA] Francis, Dick. 1973. *Slay Ride*. Translated by Jaroslava Moserová-Davidová as *Chladná zrada*. 1998. Praha: Olympia.
- [JF\_COR] Franzen, Jonathan. *The Corrections*. 2001. Translated by Jan Jiráček as *Rozhřešení*. 2004. Praha: Ikar

- [JG\_BRE] Grisham, John. *The Brethren*. 2000. Translated by Richard Podaný as *Bratrstvo*. 2000. Praha: Ikar / Knižní klub.
- [JG\_CLI] Grisham, John. *The Client*. 1993. Translated by Josef Orel and Marie Orlová as *Klient*. 1995. Praha: Ikar / Knižní klub.
- [JI\_WID] Irving, John. *A Widow for One Year*. 1998. Translated by Milada Nováková as *Rok vdovou*. 2000. Praha: Euromedia / Knižní klub.
- [JKJ\_THR] Jerome, K. Jerome. 1889. *Three Men in a Boat (to Say Nothing of the Dog)*. Translated by J. Z. Novák as *Tři muži ve člunu (o psu nemluvě)*. 2007. Praha: Odeon.
- [SK\_CAR] King, Stephen. 1974. *Carrie*. Translated by Ivan Němeček as *Carrie*. 2007. Praha: Beta – Pavel Dobrovský.
- [JL\_ALO] Lindsey, Johanna. 2004. *A Loving Scoundrel*. Translated by Renáta Tetřevová as *Zamilovaný ničema*. 2005. Praha: Euromedia Group.
- [DL\_CHA] Lodge, David. 1978. *Changing Places: A Tale of Two Campuses*. Translated by Mirek Čejka as *Hostující profesori*. 1980. Praha: Odeon.
- [DL\_DEA] Lodge, David. 2008. *Deaf Sentence*. Translated by Richard Podaný as *Nejtíšší trest*. 2009. Praha: Mladá fronta.
- [TM\_C] McCarthy, Tom. 2011. *C*. Translated by Michaela Marková as *C*. 2011. Praha: Odeon.
- [CP\_CHO] Palahniuk, Chuck. 2001. *Choke*. Translated by Richard Podaný as *Zalknutí*. 2009. Praha: Odeon.
- [TPNG\_GOO] Pratchett, Terry and Neil Gaiman. 2006. *Good Omens*. Translated by Jan Kantůrek as *Dobrá znamení*. 1997. Praha: Talpress.
- [JKR\_HAR] Rowling, Joanne Kathleen. 2003. *Harry Potter and the Order of the Phoenix*. Translated by Pavel Medek as *Harry Potter a Fénixův řád*. 2004. Praha: Albatros.

- [ARS\_HI] Siddons, Anne Rivers. *Hill Towns*. 1993. Translated by Hana Parkánová-Whitton as *Bezpečné výšiny*. 2004. Praha: Ikar.
- [DS\_SEC] Silva, Daniel. 2007. *The Secret Servant*. Translated by Klára Míčková as *Tajná služba*. 2007. Ostrava: Domino.
- [ZS\_ONB] Smith, Zadie. 2006. *On Beauty*. Translated by Petra Diestlerová as *O kráse*. 2006. Praha: BB Art.
- [JKT\_CON] Toole, John Kennedy. 1980. *A Confederacy of Dunces*. Translated by Jaroslav Kořán as *Spolčení hlupců*. 1985. Praha: Odeon.

## **Appendix 2: Subcorpus of Subtitles from InterCorp**

The corpus of subtitles in InterCorp contains data from [www.opensubtitles.org](http://www.opensubtitles.org). However, InterCorp does not provide any other information except for the year of publication and the name of a film.

- [BREAK\_06] The Break-Up, 2006
- [CIRQUE\_09] Cirque du Freak: The Vampire's Assistant, 2009
- [CONTRACT\_06] The Contract, 2006
- [COUNTESS\_09] Countess, 2009
- [ECLIPSE\_10] Eclipse, 2010
- [FAHRENHEIT\_04] Fahrenheit 9/11, 2004
- [G-FORCE\_09] G-Force, 2009
- [IMPULSE\_08] Impulse, 2008
- [JEEPERS\_01] Jeepers Creepers, 2001
- [MATADOR\_05] The Matador, 2005
- [MOTIVES\_07] Motives 2, 2007

[NARC\_02]           Narc, 2002

[OUTCASTS\_10]       Outcasts, 2010

[RAMPAGE\_09]        Rampage, 2009

[SCOOBY DOO\_02]    Scooby-Doo, 2002

[SEE NO EVIL\_06]    See No Evil, 2006

[SEX DRIVE \_08]     Sex Drive, 2008

[TALLADEGA\_06]     Talladega Nights: The Ballad of Ricky Bobby, 2006

[THE REAL L \_10]    The Real L Word: Los Angeles, 2010

## WORKS CITED

- Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a Corpus*. Philadelphia: John Benjamins Pub.
- Aijmer, Karin. 2004. "Pragmatic Markers in Spoken Interlanguage." *Nordic Journal of English Studies* 3: 173-190.
- Aijmer, Karin. 2007. "The Meaning and Functions of the Swedish Discourse Marker Alltså – Evidence from Translation Corpora." *Catalan Journal of Linguistics* 6: 31-59.
- Aijmer. 2014. *Corpus Pragmatics: A Handbook*. Cambridge University Press.
- Aijmer, Karin and Bengt Altenberg. 2001. "Zero Translations and Cross-Linguistic Equivalence: Evidence from the English-Swedish Parallel Corpus." In: *From the Colt's Mouth, and Other Places: Studies in Honour of Anna-Brita Stenström*. Edited by Hasselgren, A. and L. E. Breivik, 19-41. Amsterdam: Rodopis.
- Aijmer, Karin and Marie Simon Vandenberg. 2011. "Pragmatic Markers." In *Discursive Pragmatics*, edited by Jan Zienkowski, Jan-Ola Östman, and Jef Verschueren, 223 – 247. Amsterdam: John Benjamins Pub.
- Andersen, Gisle. 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents*. Amsterdam: Benjamins.
- Archer, Dawn, Karin Aijmer, and Anne Wichmann. 2012. *Pragmatics: An Advanced Resource Book for Students*. Milton Park, Abingdon, Oxon: Routledge.
- Biber, Douglas. 1999. *Longman Grammar of Spoken and Written English*. Harlow, England: Longman.
- Blakemore, Diane. 1987. *Semantic Constraints on Relevance*. Oxford, UK: Blackwell.
- Brinton, Laurel J. 2008. *The Comment Clause in English. Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press.
- Brinton, Laurel. 2001. "Historical Discourse Analysis." In *The Handbook of Discourse Analysis* edited by Deborah Schiffrin, Deborah Tannen, and Heidi Ehnberger, Hamilton. Malden, 138-160. MA: Blackwell Publishers.

- Brinton, Laurel. 2007. "The Development of *I mean*: Implication for the Study of Historical Pragmatics". In *Methods in Historical Pragmatics*, edited by Fitzmaurice, Susan M., and Irma Taavitsainen, 37-80. Berlin: Mouton De Gruyter.
- Brinton, Laurel. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton De Gruyter.
- Brown, Geoffrey and Geoffrey Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge UP.
- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Chafe, Wallace L. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Crystal, David. 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- Cvrček, Václav. 2014. "Alignment (zarovnění)." Last modified November 24, 2014. Available at [http://wiki.korpus.cz/doku.php/pojmy:alignment?s\[\]=alignment](http://wiki.korpus.cz/doku.php/pojmy:alignment?s[]=alignment). Accessed September 16, 2016
- Del, Saz Rubio M. Milagros. 2007. *English Discourse Markers of Reformulation*. Bern: Peter Lang.
- Erman, Britt. 2001. "Pragmatic Markers Revisited with a Focus on You Know in Adult and Adolescent Talk." *Journal of Pragmatics*, 1337-1359.
- Fraser, Bruce. 1996. "Pragmatic Markers." *Pragmatics* 6(2).
- Fraser, Bruce. 1999. "What Are Discourse Markers?" *Journal of Pragmatics* 31: 931-952.
- Fraser, Bruce. 2009. "An Account of Discourse Markers." *International Review of Pragmatics* 1: 293-320.
- Fuller, Janet M. 2003a. "Discourse marker use across speech contexts: A comparison of native and non-native speaker performance." *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 22(2):185-208.

- Fuller, Janet M. 2003b. "The influence of speaker roles on discourse marker use." *Journal of Pragmatics* 35: 23-45.
- Furkó, Bálint Péter. 2014. "Perspectives on the Translation of Discourse Markers." *Acta Universitatis Scientiae, Philologica* 6: 181–196.
- Halliday, M. A. K., and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
- Hansen, Maj-Britt Mosegaard. 1998. *The Function of Discourse Particles: A Study with Special Reference to Spoken Standard French (Pragmatics & Beyond)*. John Benjamins Publishing Company.
- Havránek, B., Bělič, J., Helcl, M. and A. Jedlička. *Slovník spisovného jazyka českého*, s. v. "teda," Last modified 2011.  
<http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=teda&sti=EMPTY&where=hesla&hsubstr=no>. Accessed November 20, 2016.
- Holmes, Janet. 1986. "Functions of You Know in Women's and Men's Speech." *Language in Society* 15: 1-22.
- Huang, Lan Fen. 2011. "Discourse Markers in Spoken English a Corpus Study of Native Speakers and Chinese Non-native Speakers." Dissertation thesis. The University of Birmingham. <http://etheses.bham.ac.uk/2969/9/Huang11PhD.pdf>.
- Jahandarie, Khosrow. 1999. *Spoken and Written Discourse: A Multi-disciplinary Perspective*. Stamford, CT: Ablex Pub.
- Johansson, Stig. 2003. "Contrastive Linguistics and Corpora." In *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, edited by Granger, Sylviane, Jacques Lerot, and Stephanie Petch-Tyson., 31-44. Amsterdam: Rodopi.
- Johansson, Stig. 2007. *Seeing Through Multilingual Corpora*. In: R. Facchinetti (ed.), *Corpus Linguistics 25 Years On*. Amsterdam – New York: Rodopi, s. 51–71.
- Jucker, Andreas H. and Yael Ziv. 1998. *Discourse Markers: Description and Theory*. Amsterdam and Philadelphia: John Benjamins.
- Karamitroglou, Fotios. 2000. *Towards a Methodology for the Investigation of Norms in Audiovisual Translation: The Choice between Subtitling and Revoicing in Greece*. Amsterdam: Rodopi.



- Kolářová, Ivana. 1998. "Významy a funkce slova tedy (teda) v souvislých projevech." *Naše řeč* 81(2-3): 118-123. Available online at <http://nase-rec.ujc.cas.cz/archiv.php?art=7439>
- Leech, Geoffrey N., M. Deuchar, and Robert Hoogenraad. 1982. *English Grammar for Today: A New Introduction*. London: Macmillan Press in Conjunction with the English Association.
- Lenk, Uta. 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. Tübingen: Gunter Narr Verlag.
- Lutzky, Ursula. 2012. *Discourse Markers in Early Modern English*. John Benjamins Publishing Company.
- Macaulay, Ronald. 2002. "You Know, It Depends." *Journal of Pragmatics* 34: 749-767.
- Mullan, Kerry. 2010. *Expressing Opinions in French and Australian English Discourse: A Semantic and Interactional Analysis*. Amsterdam: John Benjamins Pub.
- Müller, Simone. 2005. *Discourse Markers in Native and Non-native English Discourse*. Amsterdam: John Benjamins.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Redeker, Gisela. 1990. "Ideational and Pragmatic Markers of Discourse Structure." *Journal of Pragmatics* 14: 367-381.
- Redeker, Gisela. 1991. "Linguistic Markers of Discourse Structure." *Linguistics* 29:1139-1172.
- Romero-Trillo, Jesús. 2015. "Prosodic Modeling and Position Analysis of Pragmatic Markers in English Conversation." *Corpus Linguistics and Linguistic Theory*: 1-27.
- Rosen, Alexander. 2016. "Korpus InterCorp verze 8." Last modified June 22, 2016. Available at <http://wiki.korpus.cz/doku.php/cnk:intercorp:verze8> Accessed September 16, 2016.
- Scheler, Gabriele and Fischer, Kerstin. 1997. "The Many Functions of Discourse Particles: A Computational Model of Pragmatic Interpretation." In *Proceedings*

- of the Nineteenth Annual Conference of the Cognitive Science Society, edited by Shafto, Michael G. and Pat Langley, 668-673. Stanford: Psychology Press.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Schiffrin, Deborah. 2008. "Discourse Markers: Language, Meaning, Context." In *The Handbook of Discourse Analysis*, edited by Schiffrin, Deborah, and Deborah Tannen, and Heidi Ehernberger Hamilton, 54-75. MA: John Wiley & Sons.
- Schourup, Lawrence. 1982. "Common Discourse Particles in English Conversation." Dissertation thesis. The Ohio State University.
- Schourup, Lawrence. 1999. "Discourse Markers." *Lingua* 107: 227-265.
- Sindoni, Maria G. 2014. *Spoken and Written Discourse in Online Interactions: A Multimodal Approach*. New York: Routledge.
- Sperber, Dan and Deirdre Wilson. 2012. *Meaning and Relevance*. Cambridge: Cambridge University Press.
- Stede, Manfred, and Birte Schmitz. 2000. "Discourse Particles and Discourse Functions." *Machine Translation*, 125-147.
- Szczyrbak, Magdalena. 2014. "Pragmatic Marker Use in Police Interviews: The Case of *I Mean* and *You Know*." *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 131: 371–379. Accessed September 23, 2015. Available at <http://www.ejournals.eu/pliki/art/3572/>
- Tree, Jean E. Fox, and Josef C. Schrock. 2002. "Basic Meanings of *You know* and *I mean*." *Journal of Pragmatics* 34: 727-747. Accessed December 15, 2015.
- Urgelles-Coll, Miriam. 2010. *The Syntax and Semantics of Discourse Markers: Continuum Studies in Theoretical Linguistics*. London: Continuum.

## CORPUS

- Czech National Corpus – InterCorp*. Prague: Institute of the Czech National Corpus. Available online at <http://www.korpus.cz>.

## LIST OF TABLES

<b>Table 1.</b> Functions of DMs according to Lutzky (2012, 39).....	23
<b>Table 2.</b> Brinton’s classification (2007, 39-53) of functions of <i>I mean</i> .....	32
<b>Table 3.</b> Erman’s classification (2001, 1341) of functions of <i>you know</i> .....	34
<b>Table 4.</b> Czech translation equivalents of <i>I mean</i> in the subcorpus of fiction .....	55
<b>Table 5.</b> Czech translation equivalents of <i>I mean</i> in the subcorpus of subtitles .....	58
<b>Table 6.</b> Czech translation equivalents of <i>you know</i> in the subcorpus of fiction.....	60
<b>Table 7.</b> Czech translation equivalents of <i>you know</i> in the subcorpus of subtitles .....	62
<b>Table 8.</b> Syntactic position of <i>I mean</i> in the subcorpora.....	73
<b>Table 9.</b> Changes concerning syntactic position of <i>I mean</i> in source texts.....	75
<b>Table 10.</b> Syntactic position of <i>you know</i> in the subcorpora.....	76
<b>Table 11.</b> Changes concerning syntactic position of <i>you know</i> in source texts .....	77
<b>Table 12.</b> Co-occurrence of <i>I mean</i> and <i>you know</i> with other discourse markers .....	78
<b>Table 13.</b> Co-occurrence of <i>I mean</i> and <i>you know</i> with expressions of agreement or disagreement .....	80

## LIST OF GRAPHS

<b>Graph 1.</b> Percentage of Czech translation equivalents of <i>I mean</i> in the subcorpus of subtitles .....	68
<b>Graph 2.</b> Percentage of Czech translation equivalents of <i>you know</i> in the subcorpus of subtitles .....	69
<b>Graph 3.</b> Percentage of Czech translation equivalents of <i>I mean</i> in the subcorpus of fiction.....	70
<b>Graph 4.</b> Percentage of Czech translation equivalents of <i>you know</i> in the subcorpus of fiction.....	70