

Towards discourse markers

A cross-linguistic study of
grammatical constructionalization

Yinchun Bai

Towards discourse markers:
a cross-linguistic study of grammatical
constructionalization

Yinchun Bai

**Towards discourse markers:
a cross-linguistic study of grammatical
constructionalization**

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs-Universität
Freiburg i. Br.

vorgelegt von

Yinchun Bai

aus Tianjin

WS 2021/22

Erstgutachter: Prof. Dr. Bernd Kortmann
Zweitgutachter: Prof. Dr. Johan van der Auwera
Drittgutachter: Prof. Dr. Frank Brisard

Vorsitzender des Promotionsausschusses
der Gemeinsamen Kommission
der Philologischen und
der Philosophischen Fakultät: Prof. Dr. Frank Bezner

Datum der Disputation: 04.07.2022



University
of Antwerp

Faculty of Arts
Department of Linguistics

Towards discourse markers: a cross-linguistic study of grammatical constructionalization

PhD thesis submitted for the degree of doctor in de
taalkunde at the University of Antwerp
by Yinchun Bai

Supervisors:

Prof. Dr. Bernd Kortmann,
Prof. Dr. Johan van der Auwera,
Prof. Dr. Frank Brisard

Antwerp, 2022

Summary

This thesis takes an interest in the emergence and development of discourse markers. It develops within the theoretical framework of Construction Grammar and treats discourse markers as conventionalized form-function units and their diachronic change as a process of grammatical constructionalization. It addresses the following questions: (1) Which incremental changes are involved during the process that leads a linguistic element towards a discourse marker? (2) What motivational factors are behind specific constructional changes? (3) Are there cross-linguistic generalizations to be made, both in terms of the semantic and/or syntactic sources and the development paths? In order to approach these questions, the thesis draws on existing studies of grammaticalization and diachronic Construction Grammar that account for the case of discourse markers. It unites different theories and examines the formal and functional representations and especially changes of a construction on its way of becoming a discourse marker. In the search of cross-linguistically universal processes and/or patterns of change, it further develops a comparative approach. It examines a pair of linguistically heterogeneous expressions that typically function in the same pragmatic domain: i.e. topic-introducing discourse markers *speaking of X* (SPOX) in English and *huashuo* in Chinese. The study is corpus-based and includes both functional and frequency distributional analyses both panchronically and diachronically. The results show that SPOX and *huashuo* share quite a few formal-semantic properties as discourse markers and there are many

overlaps in their functions and usage patterns. The major difference lies in their mechanisms in linking different topics together and establishing topical relevance. Diachronically, they share the semantic root of “speak” at the very beginning of their constructionalization and pragmaticalization processes, and both constructions were often used as clausal elements in a sentence prior to the emergence of discourse marking functions. But their individual development paths still have distinctive courses and specific motivating factors. On the basis of the observations made in this thesis, it appears that pragmatic strengthening, syntactic/prosodic detachment, scope expansion, and development of functional polysemy are universal processes during the constructionalization process of discourse markers, while formal reduction and semantic bleaching seem to be less categorical and more contingent upon the type of the discourse markers and differ from one case to another. The phenomena of *layering*, *persistence*, *divergence*, *deategorialization* and *paradigmaticization* seem to be universal common properties of emergent discourse markers as well.

Zusammenfassung

Die Doktorarbeit beschäftigt sich mit der Entstehung und Entwicklung von Diskursmarkern. Die Arbeit bewegt sich im theoretischen Rahmen der Konstruktionsgrammatik und behandelt Diskursmarker als konventionalisierte Form-Funktions-Paare und ihre diachrone Veränderung als Prozess der grammatischen Konstruktionalisierung. Sie befasst sich mit den folgenden Fragestellungen: (1) Welche inkrementellen Veränderungen sind an dem Prozess beteiligt, der zur Entwicklung eines linguistischen Elements zu einem Diskursmarker führt? (2) Welche Faktoren motivieren die konkreten konstruktionalen Veränderungen? (3) Gibt es sprachübergreifende Generalisierungen sowohl hinsichtlich der semantischen und/oder syntaktischen Ursprünge als auch der Entwicklungswege? Um diese Fragen zu beantworten, stützt sich die Arbeit auf bestehende Studien zur Grammatikalisierung und diachronen Konstruktionsgrammatik, die das Phänomen Diskursmarker behandeln. Sie vereint unterschiedliche Theorien und untersucht die formalen und funktionalen Repräsentationen und insbesondere die Veränderungen einer Konstruktion auf ihrem Weg zum Diskursmarker. Auf der Suche nach sprachübergreifenden universellen Prozessen und/oder Veränderungsmustern entwickelt die Arbeit einen kontrastiven Ansatz weiter. Sie untersucht zwei sprachlich heterogene Ausdrücke, die typischerweise in derselben pragmatischen Domäne eingesetzt werden, d.h. die Thema-einführenden Diskursmarker *speaking of X* (SPOX) im Englischen und *huashuo* im Chinesischen. Die Studie ist

korpus-basiert und beinhaltet sowohl funktionelle als auch Häufigkeitsverteilungsanalysen nicht nur panchronisch, sondern auch diachronisch. Die Ergebnisse zeigen, dass SPOX und *huashuo* einige formal-semantische Eigenschaften als Diskursmarker gemeinsam haben und es viele Überschneidungen in ihren Funktionen und Anwendungsmustern gibt. Der Hauptunterschied liegt in ihren Mechanismen, verschiedene Themen miteinander zu verknüpfen und thematische Relevanz herzustellen. Diachronisch teilen sie die semantische Wurzel „sprechen“ ganz am Anfang ihrer Konstruktionalisierungs- und Pragmatikalisierungsprozesse, und beide Konstruktionen wurden oft als Teilsatz verwendet, bevor sich ihre Funktionen als Diskursmarker entwickelten. Dennoch haben ihre jeweiligen Entwicklungen unterschiedliche Verläufe und spezifische Motivationsfaktoren. Auf Basis der in dieser Arbeit gemachten Beobachtungen scheint es, dass pragmatische Verstärkung, syntaktische/prosodische Abtrennung, Scope-Erweiterung, und Entwicklung funktionaler Polysemie universelle Prozesse während des Konstruktionalisierungsprozesses von Diskursmarkern sind, während formale Reduktion und semantisches Ausbleichen („bleaching“) weniger kategorisch und abhängig von der Art der Diskursmarker erscheinen und sich von Fall zu Fall unterscheiden. Auch die Phänomene des *Schichtung* („Layering“), *Persistenz*, *Divergenz*, *Dekategorisierung* und *Paradigmatisierung* scheinen universelle Gemeinsamkeiten in der Entstehung von Diskursmarkern zu sein.

Samenvatting

Dit proefschrift onderzoekt het ontstaan en de ontwikkeling van *discourse markers*. Het situeert zich binnen het theoretische kader van constructiegrammatica en behandelt *discourse markers* als geconventionaliseerde eenheden van vorm and functie en hun diachrone verandering als een proces van grammaticale constructionalisering. Het onderzoek stelt de volgende vragen: (1) Welke incrementele veranderingen zijn betrokken in het proces dat van een talig element naar een *discourse marker* leidt? (2) Welke motiverende factoren zitten er achter specifieke constructionele veranderingen? (3) Zijn er generalisaties te maken over talen heen wat betreft zowel de semantische en/of syntactische eigenschappen van de oorspronkelijke lexicale elementen als de ontwikkelingstrajecten? Om deze vragen te beantwoorden baseert het proefschrift zich op bestaande studies over grammaticalisatie en diachrone constructiegrammatica die de kwestie van *discourse markers* verklaren. Het verenigt verschillende theorieën en onderzoekt de formele en functionele representaties en vooral de veranderingen van een constructie die op weg is om een *discourse marker* te worden. In de zoektocht naar crosslinguïstisch universele processen en/of veranderingspatronen past het een vergelijkende benadering toe. Het onderzoekt een stel talig heterogene uitdrukkingen die typisch in hetzelfde pragmatische domein functioneren, met name de thema-introducerende *discourse markers* *speaking of X* (SPOX) in het Engels en *huashuo* in het Chinees. Het onderzoek is corpusgebaseerd

en omvat zowel functionele analyses als frequentieverdelingsanalyses, zowel panchroon als diachroon. De resultaten laten zien dat SPOX en *huashuo* ondanks hun verschillende oorsprong een aantal formeel-semantiche eigenschappen als discourse markers gemeenschappelijk hebben en dat er veel raakvlakken zijn in hun functies en gebruikspatronen. Het grote verschil ligt in de talige mechanismen om verschillende thema's aan elkaar te koppelen en om relevantie van het thema tot stand te brengen. Wat ze diachroon gemeenschappelijk hebben, is de semantische stam 'spreken' aan het begin van hun constructionaliserings- en pragmaticalisatieprocessen. Beide constructies werden ook vaak gebruikt als syntactisch geïntegreerde zinslementen voordat ze functies als discourse markers ontwikkelden. Hun individuele ontwikkelingstrajecten worden echter nog steeds gekenmerkt door aparte trajecten en specifieke motiverende factoren. Op basis van de observaties in dit proefschrift lijkt het erop dat pragmatische versterking, syntactische/prosodische onthechting, uitbreiding van *scope* en de ontwikkeling van functionele polysemie universele processen zijn tijdens het constructionaliseringsproces van discourse markers. Formele reductie en semantische bleking ('*bleaching*') daarentegen lijken minder categorisch en meer afhankelijk van het type discourse marker te zijn en van geval tot geval te verschillen. De verschijnselen van *gelaagdheid* ('*layering*'), *persistentie*, *divergentie*, *deategorisatie* en *paradigmatisering* lijken ook universele gemeenschappelijke eigenschappen te zijn van opkomende discourse markers.

Acknowledgements

This PhD project is conducted under a double-PhD agreement between the University of Freiburg, Germany and the University of Antwerp, Belgium. I thank all the professors and administrators involved at both universities to have made this cooperation happen.

A major part of the research conducted in this doctoral thesis was funded by the three-year LGFG doctoral scholarship “Empirical Linguistics” (Landesgraduiertenförderungsgesetzes Promotionsstipendium „Empirische Linguistik“) granted to me by the Hermann Paul School of Linguistics Freiburg-Basel. Additionally, I was also supported by the “educational credit” system at the Antwerp Doctoral School, which covered part of the costs of participating in classes and attending conferences.

I would like to thank Prof. Bernd Kortmann and Prof. Johan van der Auwera for their supervision and constructive criticisms over the years. They have given me freedom with guidance, trust with nudging pushes. I would also like to thank Prof. Frank Brisard for stepping in as a third supervisor towards the end of my PhD and for his generous advice that helped improve the thesis at a later stage.

A round of thanks goes to all my colleagues (old and new) at the University of Freiburg and the University of Antwerp, who have inspired and motivated me in different ways and at different stages of my PhD project. Special thanks go to Verena Schröter, who proofread the thesis on very short notice and corrected the summary in German, and to Daniël van Olmen, who translated the summary into Dutch.

Acknowledgements

I also thank my parents. They made me who I am today. And I'm grateful that they've always allowed me to follow my interest in academia and in life – a luxury that not every Chinese family can afford.

Last but not least, I owe my greatest gratitude to my boyfriend and fiancé Haifeng Chu, for his confidence in me and generous support in various ways through all these years.

Table of Contents

Summaryi

Zusammenfassung.....iii

Samenvatting.....v

Acknowledgements.....vii

Table of Contentsix

List of Tablesxiv

List of Figures.....xv

List of Abbreviations.....xix

Chapter 1 Introduction1

Chapter 2 Theoretical background.....13

 2.1 Discourse markers13

 2.1.1 Terminology13

 2.1.2 Characteristics16

Table of Contents

2.2	Construction Grammar.....	19
2.2.1	Discourse studies and Construction Grammar	23
2.2.2	Cross-linguistic Construction Grammar	25
2.2.3	Diachronic Construction Grammar.....	27
2.3	Grammatical constructionalization.....	30
2.3.1	Grammaticalization.....	32
2.3.2	Lexicalization and constructionalization	37
2.3.3	Pragmaticalization	41
2.3.4	Subjectification and intersubjectification	44
2.4	Essential concepts in discourse organization	46
2.4.1	Discourse coherence and topical relevance	47
2.4.2	Digression vs. topic introduction	53
2.4.3	Discursive and pragmatic pretexts.....	60
2.4.4	Periphery and functional asymmetry.....	62
2.4.5	Dialogual and monological contexts	64
Chapter 3 The <i>speaking of X</i> construction		67
3.1	Data and methods.....	67
3.1.1	Corpora and datasets	67
3.1.2	Classification and annotation.....	71
3.2	Formal and semantic properties	77
3.2.1	Semantic content.....	77

3.2.2	Formal fixedness and variation.....	82
3.3	Functional aspects.....	86
3.3.1	Topic-introducing functions.....	88
3.3.1.1	Topic change and topic expansion	88
3.3.1.2	Topic initiation.....	92
3.3.1.3	Frequency distribution	95
3.3.2	Speech act adverbial.....	97
3.3.3	Emphasized commenting	99
3.3.4	Speaking of the devil	100
3.3.5	Peripheral positions.....	102
3.3.6	Dialogual and monologual uses.....	109
3.4	Diachronic changes	114
3.4.1	Grammatical constructionalization	114
3.4.2	Pragmaticalization	117
3.4.3	Change in peripheral positions	120
3.4.4	Motivation for change	123
3.4.4.1	Sentence-initial position.....	123
3.4.4.2	Asyndetic subordination.....	126
3.4.4.3	Co-occurrence with speech-act verbs	127
3.5	Summary	129

Chapter 4 The *huashuo* construction..... 131

4.1	Data and methods.....	131
4.1.1	Corpora and datasets	131
4.1.2	Classification and annotation.....	137
4.2	Formal and semantic properties	142
4.2.1	Semantic content.....	142
4.2.2	Formal fixedness	148
4.2.3	Loose syntactic connection.....	150
4.3	Functional aspects	152
4.3.1	Topic-introducing functions.....	154
4.3.1.1	Topic expansion and topic change	155
4.3.1.2	Topic initiation.....	160
4.3.1.3	Frequency distribution.....	162
4.3.2	Interpersonal functions	164
4.3.2.1	Self-repair	165
4.3.2.2	Politeness management	167
4.3.3	Peripheral positions	169
4.3.4	Dialogual and monological uses.....	172
4.4	Diachronic changes	177
4.4.1	Grammatical constructionalization	177
4.4.2	Pragmaticalization	187
4.4.3	Change in peripheral positions	190
4.4.4	Motivation for change	194

4.5	Summary	196
Chapter 5 Conclusion		199
5.1	Summary of findings	199
5.1.1	Formal-semantic properties in contrast....	200
5.1.2	Functional aspects in contrast	201
5.1.3	Diachronic development in contrast.....	209
5.2	Implications of findings.....	212
5.3	Suggestions for further research	215
5.3.1	Dataset expansion.....	215
5.3.2	Comedic effect	216
References		219

List of Tables

Table 1	Dataset COHA composition
Table 2	Dataset COCA-SPOK composition
Table 3	Gerundial uses of SPOX
Table 4	Participial uses of SPOX
Table 5	Discourse marker uses of SPOX
Table 6	Datasets CNCORPUS and CCL composition
Table 7	Dataset Chat composition
Table 8	Different uses of <i>huashuo</i>
Table 9	Internal structural variations of <i>huashuo</i>
Table 10	Subjectivity and intersubjectivity of different discourse marking functions of SPOX and <i>huashuo</i>

List of Figures

- Figure 1 Type frequency distribution of discourse marker SPOX in the COCA-SPOK dataset
- Figure 2 Frequency distribution of the discourse marking functions of SPOX in the COCA-SPOK dataset
- Figure 3 Frequency distribution of the discourse marking functions of SPOX in the COHA dataset
- Figure 4 Frequency distribution of the topic-introducing functions of SPOX in the COCA-SPOK dataset
- Figure 5 Frequency distribution of the topic-introducing functions of different variants of SPOX in the COCA-SPOK dataset
- Figure 6 Frequency distribution of discourse marker SPOX in different peripheral positions in the COCA-SPOK dataset
- Figure 7 Frequency distribution of different discourse marking functions of SPOX in different peripheral positions in the COCA-SPOK dataset
- Figure 8 Frequency distribution of monologal and dialogal uses of discourse marker SPOX in the COCA-SPOK dataset
- Figure 9 Frequency distribution of monologal and dialogal uses of different variants of discourse marker SPOX in the COCA-SPOK dataset
- Figure 10 Frequency distribution of different discourse marking functions of SPOX in dialogal and monologal contexts in the COCA-SPOK dataset

List of Figures

- Figure 11 Frequency distribution of the topic-introducing functions of SPOX in dialogual and monologal contexts in the COCA-SPOK dataset
- Figure 12 Diachronic frequency distribution of different uses of SPOX in the COHA dataset
- Figure 13 Diachronic frequency distribution of different variants of topic-introducing SPOX in the COHA and COCA-SPOK datasets
- Figure 14 Diachronic frequency distribution of the discourse marking functions of SPOX in the COHA dataset
- Figure 15 Diachronic frequency distribution of the topic-introducing functions of SPOX in the COCA-SPOK dataset
- Figure 16 Cline of discourse units
- Figure 17 Diachronic frequency distribution of discourse marker SPOX in different peripheral positions in the COHA and COCA-SPOK datasets
- Figure 18 Diachronic frequency distribution of topic-introducing SPOX in different peripheral positions in the COHA and COCA-POK datasets
- Figure 19 Diachronic frequency distribution of speech act adverbial SPOX in different peripheral positions in the COHA and COCA-SPOK datasets
- Figure 20 Diachronic relative frequency distribution of adverbial SPOX in different syntactic positions in the COHA dataset
- Figure 21 Frequency distribution of the subordinators of the SPOX adverbial in the COHA dataset

- Figure 22 Frequency distribution of the discourse marking functions of *huashuo* in the Chat dataset
- Figure 23 Frequency distribution of the topic-introducing functions of *huashuo* in the Chat dataset
- Figure 24 Frequency distribution of the topic-introducing functions of *huashuo* with different linking mechanisms in the Chat dataset
- Figure 25 Frequency distribution of the discourse marking functions of *huashuo* in different peripheral positions in the Chat dataset
- Figure 26 Frequency distribution of the topic-introducing functions of *huashuo* in dialogual and monologual contexts in the Chat dataset
- Figure 27 Frequency distribution of the interpersonal functions of *huashuo* in dialogual and monologual contexts in the Chat dataset
- Figure 28 Diachronic frequency distribution of the different uses of *huashuo* in the CCL and Chat datasets
- Figure 29 Emergence of non-chapter-initial uses of the SV-clausal *huashuo*
- Figure 30 Diachronic relative frequency distribution of the different uses of *huashuo* in the Modern time period (1912 – 2019) in the CCL and Chat datasets
- Figure 31 Diachronic frequency distribution of the discourse marking functions of *huashuo* in the Chat dataset
- Figure 32 Diachronic frequency distribution of the topic-introducing functions of *huashuo* in the Chat dataset

List of Figures

- Figure 33 Diachronic frequency distribution of topic-introducing *huashuo* with different linking mechanisms in the Chat dataset
- Figure 34 Diachronic frequency distribution of discourse marker *huashuo* in different peripheral positions in the Chat dataset
- Figure 35 Diachronic frequency distribution of topic-introducing *huashuo* in different peripheral positions in the Chat dataset
- Figure 36 Diachronic frequency distribution of interpersonal *huashuo* in different peripheral positions in the Chat dataset
- Figure 37 Relative frequency distribution of the discourse marking functions of SPOX and *huashuo* in the COCA-SPOK dataset and the Chat dataset respectively
- Figure 38 Relative frequency distribution of the linking mechanisms of topic-introducing SPOX and *huashuo* in the COCA-SPOK dataset and the Chat dataset respectively
- Figure 39 Frequency distribution of monologual and dialogual uses of topic-introducing SPOX and *huashuo* in the COCA-SPOK dataset and the Chat dataset respectively
- Figure 40 Relative frequency distribution of SPOX and *huashuo* as discourse markers in different peripheral positions in the COCA-SPOK dataset and the Chat dataset respectively

List of Abbreviations

ADJ	->	adjective
ADV	->	adverb(ial)
CLF	->	classifier
COMPL	->	completive
COP	->	copula
CNJ	->	conjunction
INTS	->	intensifier
LOC	->	locative
MOD	->	modal particle
NEG	->	negation
POSS	->	possessive
PRF	->	perfect
PROG	->	progressive
PST	->	past
Q	->	question particle/marker

Chapter 1

Introduction

Discourse markers have long been an interesting subject in linguistic studies. They have been investigated from different perspectives, quite intensively in synchronic studies of their functions and morpho-syntactic features (e.g. Schifffrin 1987; Fraser 1990; 1996; 1999; Jucker & Ziv 1998; Fischer 2006a; Celle & Huart 2007; Aijmer 2002; 2013; Degand et al. 2013) and increasingly in diachronic studies of their development paths (e.g. Brinton 1996; 2008; 2017; Lewis 2011; 2018; Prevost 2011; Heine 2013; Pons Border á & Loureda Lamas 2018; Traugott 2020) as well as in contrastive studies of cross-linguistic differences and generalizations (e.g. Aijmer & Simon-Vandenberg 2006; Wang 2011; Auer & Maschler 2016; Fagard & Blumenthal 2020).

Joining the interests of pragmatics, historical linguistics and contrastive linguistics, this thesis investigates not only the formal-functional aspects of discourse markers but also the mechanisms and processes of change involved in their emergence and development, both within a language and across languages. It adopts a constructional approach in the analyses and views discourse markers as conventionalized form-function pairs, whose meanings are not fully predictable from their lexical components, and the diachronic development of discourse markers as the sum of various changes occurring in different parts of the form-function pairing (a detailed

introduction of the theoretical background of this thesis will be given in Chapter 2).

Overall, the thesis raises the following questions: (1) Which incremental changes are involved during the process that leads a linguistic element towards a discourse marker? (2) What motivational factors are behind specific constructional changes? (3) Are there cross-linguistic generalizations to be made, both in terms of the semantic and/or syntactic sources and the development paths? In order to address these questions, this thesis focuses on two constructions: the *speaking of X* construction in American English and the *huashuo* construction in Mandarin Chinese, which are comparable both in terms of their semantic sources and pragmatic functions.

The *speaking of X* construction (further referred to as SPOX) refers to the set phrase “*speaking of X*”, which is typically used as a discourse marker to introduce a new (sub)topic into the ongoing discourse events. As shown in (1), the speaker switches the topic about Andrea to the “Dancing with the Stars” show with the use of “*speaking of gorgeous*”. And in (2), Stubbs uses “*speaking of which*” to bring up another co-worker before Bernard leaves the conversation.

(1) *That's great news, congratulations, Andrea. You are gorgeous. And speaking of gorgeous, what a night it was on “Dancing With The Stars”, [...]* (2012, SPOK, ABC_GMA, COCA)

(2) Bernard: *If you don't mind, I should get back to work.*
Stubbs: *Of course. I know you're shorthanded. Speaking of which, have you heard from Elsie Hughes?*
(Westworld, Season 1 Episode 8, November 20, 2016)

The *huashuo* construction (further referred to as *huashuo*) refers to the linguistic element “*huàshuō* [话说] – ‘(it) speak(s) of/talk(s) about’”, which is considered to be the Chinese counterpart of SPOX and is typically used in the functional domain of topic introduction as well. As shown in (3), W uses *huashuo* to bring up another point when offering help to K. In (4), Y steers away from the original topic with the use of *huashuo* to start another seemingly unrelated topic. In (5), *huashuo* is even used at the very beginning of a new chat session.

(3) [W is helping out K who is new in town.]

W: 这回 我 在 市里, 有 问题 打 电话。
 zhèhuí wǒ zài shìlǐ, yǒu wèntí dǎ diànhuà
 this time I LOC city in have problem make phone call
 话说 你 得 买 个 德国 手机卡。
 Huàshuō nǐ děi mǎi ge déguó shǒujīkǎ.
 speaking of which you must buy CLF German cellphone card

“W: *I’ll be in town this time. If (you) have any problem, call me. Speaking of which, you need to buy a German SIM card.*”

(Chat History between K and W 2013-08-11.txt, 2013, Chat)

(4) [G and Y are talking about flight tickets and baggage allowance.]

G: 我 觉得 这 要 是 我 估计 能 用 满.....
 Wǒ juéde zhè yào shì wǒ gūjì néng yòng mǎn...
 I feel this if COP I estimate can use full
 看 我的 包 就 知道 了..... 哈哈
 Kàn wǒ de bāo jiù zhīdào le... haha
 look I-POSS purse CNJ know PRF haha

Y: 话说 你 牙 怎么 疼 了?
Huàshuō nǐ yá zěnmē tég le?
so you tooth how hurt PRF

“G: *I think if it were me I could probably use up (the entire quota)...
Just look at my purse and (you) will see... haha*

Y: So *how come you are having a toothache?*”

(2650361431(new).txt, 2012, Chat)

- (5) [X starts a new chat session in the group chat 4 days after the last one.]

X: 话说
Huàshuō
you know¹

X: 好 怀念 大家 一起 玩 三国杀 的 日子 啊
hǎo huánniàn dàjiā yìqǐ wán sānguóshā de rìzi a
very miss all together play Sanguosha ADJ day MOD

“X: You know, *(I) really miss the days when we were playing
Sanguosha together.*”

(Chat History of the WeChat-Group 2014-12-04.txt, 2014, Chat)

From these examples we get the first impressions of the intriguing formal-semantic properties and procedural status of SPOX and *huashuo*: their meanings do not seem to fully correspond to the lexical meanings of their components and they provide pragmatic

¹ See Schifffrin (1987: 285–290) for the speech-opening variant of the discourse marker *you know*.

information to the utterance instead of participating in the propositional content of the sentence.

However, neither SPOX nor *huashuo* has received enough attention yet. So far, the discourse marker usage of SPOX has only been mentioned in a few studies. It is mentioned in a few handbooks of the English grammar as an example for expressions that mark the upcoming of a new topic or a topic shift (Carter & McCarthy 2006: 211; Huddleston & Pullum 2002: 1371). Grosz and Sidner (1985: 836), in discussing discourse structure and thereby defining digression as forming a separate “intentional structure” in the global discourse structure, uses the *speaking of X* expression to illustrate this phenomenon. Fraser, in several of his studies on discourse markers, lists *speaking of X* among expressions which he labels “topic change markers” (Fraser 1988: 28; 1996: 187; 1999: 949–950; 2001: 1628–1629), i.e. discourse markers that indicate that the speaker sees the following utterance as a departure from the current topic and signal that s/he wishes to change the topic or shift from one subtopic to another under one hyper-theme, and “topic orientation markers” (Fraser 2009), i.e. discourse markers that convey the speaker’s intentions concerning the immediate future topic of the discourse, including both topic change and continuation. Gast (2010: 27), in discussing the linguistic devices that encode information structure of a discourse, mentions the *speaking of X* expression as an example of lexical indicators of sub-informativity. They refer to lexical devices which signal that the concerned utterance is only partially informative in the sense that it does not fully answer the question raised by the previous discourse.

It appears to be generally acknowledged that SPOX is used to indicate potentially incoherent topical movement in discourse and to link different topic units. However, the formal-semantic properties, discourse marking functions, and the underlying mechanisms of SPOX have not been fully explored yet. And even less attention has been paid to its usage patterns and changes over time.

In Chinese discourse marker studies, there have been a number of studies conducted on discourse markers comprising the lexical stem of *shuo* [说] – ‘speak/say’, but most of them focus on the epistemic meaning that arises from the subjectification of *shuo* and discuss the stance marking functions therewith, such as in *woshuo* [我说] – ‘I say’ and *nishuo* [你说] – ‘you say’ (e.g. Lin 2004; Liu [刘] 2008; Yao [姚] 2008; Liu [刘] 2009; Xian [鲜] 2012; Yin [尹] 2009a; Yin [殷] 2012: 248–296; Cao [曹] 2016: 128–145), as shown in (6) and (7), and the counter-expectation marker *bieshuo* [别说] – ‘don’t say’ (e.g. Yin [尹] 2009b; 2009c; Sun [孙] 2013; Zhou [周] 2013; Chen 2017), as shown in (8).

- (6) 孩子 的 内心世界 的 东西
H ázi de nèixīn shìjiè de dōngxi
child POSS heart world POSS thing

我说 你们 没有 和 他 沟通 好。

wǒshuō nimen méiyǒu hé tā gōutōng hǎo.

I say you-pl. NEG with he communicate good

“The things inside the child’s psychological world, ***I believe*** you didn’t communicate well with him.”

(Chéngzhǎng Zǎixiàn 2 [成长在线 2]) (Yin [殷] 2012: 256) (glossed and translated by the author)

- (7) 爸 你说 这 黄金 咋 掉色 呢?
 Bà nǐshuō zhè huángjīn zǎ diào sè ne?
 dad you say this gold why drop color Q

“Dad, why do you think is this gold losing color?”

(Xiāngcūn Àiqíng 2 [乡村爱情 2]) (Yin [殷] 2012: 276–277) (glossed and translated by the author)

- (8) 甲： 怎么样？ 比较 像 吧？
 Jia: Zěnmeyàng? Bǐjiào xiàng ba?
 how is it quite similar Q
 乙： 嘿嘿， 别说， 还 真 有点 像。
 Yi: Heihe, biéshuō, hái zhēn yǒu diǎn xiàng.
 hehe don’t say INTS really a bit similar

“Jia: What do you think? Quite similar, isn’t it?”

Yi: Hehe, what do you know, indeed somewhat similar.”

(Yin [尹] 2009c: 58) (glossed and translated by the author)

There are some *shuo*-type discourse marker studies which address the functional domain of topic introduction, but not nearly as many, e.g. on topic-change markers *shuodao* [说到] – ‘speaking of’ and *shuoqi* [说起] – ‘speaking of’ (e.g. Li [李] 2009; Li [李] & Zhang [张] 2012; Zhang [张] 2012), as shown in (9) and (10), the topic marker *yaoshuo* [要说] – ‘if must say’ (e.g. Dong [董] 2003: 53–54; Li [李] 2007; Li [李] 2010; Sun [孙] 2017: 160–162), as shown in (11),

and the narrative starter *shuo* [说] – ‘talk about’² (e.g. Yu [喻] & Yao [姚] 2018; Xuan [玄] 2011), shown in (12).

- (9) 说到 老李，那可真 是好帮手。
Shuōdào lǎo lǐ, nà kě zhēn shì hǎo bāngshǒu.
speaking of Old Li that INTS really COP good helper

“*Speaking of Mr. Li, he is really a good helper.*”

(Li [李] & Zhang [张] 2012: 29) (glossed and translated by the author)

- (10) 说起 开拓团，
Shuōqǐ kāitu òu án,
speaking of pioneer squad

也 是 韩家 发财 的 地方。
yě shì hán jiā fācái de dìfāng.
also COP Han family make a fortune ADJ place

“*Speaking of the pionner squad, (that) is also where the Han family made their fortune.*”

(B àofēng Zh èuyǔ [暴风骤雨]) (Zhang [张] 2012: 40) (glossed and translated by the author)

- (11) 要说 主意，人 主意 比 咱 大。
Yàoshuō zhúyì, rén zhúyì bǐ zán dà
if must say plan man plan than we big

² This narrative starter usage of *shuo* is considered to be the reduced form of the subject-verb clausal *huashuo* meaning ‘story talks about’ and is only found in northern Chinese dialects.

“As to plans, their plans are bigger than ours.”

(Biānjǐ bù de Gùshì [编辑部的故事]) (Dong [董] 2003: 53) (glossed and translated by the author)

- (12) 说 这一天， 小朋友们 玩 得 很 开 心……
Shuō zhè yī tiān, xiǎo péngyǒu men wán dé hěn kāixīn...
 talk about this one day little friend-pl. play ADV very happy

“So on this particular day, the children were playing happily...”

(Bǐ Bàba [比爸爸]) (Xuan [玄] 2011: 34) (glossed and translated by the author)

The discourse marker *huashuo* is mostly mentioned as an example among other *shuo*-type discourse markers with minimal and sporadic discussion (e.g. Li [李] 2017; Sun [孙] 2017: 77–83). There are only a handful of studies that are dedicated to *huashuo* in its own right (i.e. Zhou [周] 2012; Shi [施] 2013; Lu [鲁] 2016; Zhang [张] 2016; Chen [陈] & Huang [黄] 2017). In these studies, *huashuo* has been discussed in terms of its current discourse marking functions and its functional change through time in the light of lexicalization and subjectification. However, it has not yet been properly analyzed in terms of its constructionalized formal and semantic properties; a number of functional aspects are still left to be explored; and more importantly, even though some of the studies are more empirical than others, the findings and claims about the usage and development of the *huashuo* construction have not been well supported with empirical data in general. It is especially understudied from a frequency distributional perspective.

Therefore, this thesis aims to provide the first comprehensive accounts of SPOX and *huashuo* as constructions, exploring their formal and semantic properties, discourse marking functions and usage patterns, as well as their diachronic development paths (Chapter 3 and Chapter 4). On the basis of the observations made in this thesis and drawing on existing studies of grammaticalization and diachronic Construction Grammar, this thesis further explores cross-linguistic differences and generalizations in the development of discourse markers in general (Chapter 5)³.

Methodologically, this thesis adopts a corpus-based approach, which means that the linguistic inquiries and analyses made in this thesis are based on data derived from a large collection of authentic language use⁴. As this thesis focuses on both synchronic and diachronic behavior of discourse markers, the corpus data employed in this thesis cover a broad time span, including not only contemporary but also historical linguistic data (detailed descriptions of the corpus data are given in Section 3.1 and Section 4.1 respectively). The corpus-based approach supports the empirical nature of this thesis enormously, as it frees the analyses from personal intuition, perception and experience and reveals usage patterns and trends of development which would otherwise not be as apparent.

³ Some of the findings in this thesis have been published in Bai (2021).

⁴ See Biber et al. (1998), McEnery & Wilson (2001), and Stefanowitsch (2020) for introductions of the field of Corpus Linguistics.

Different computer programs are used for data processing and analytical purposes in this thesis. R⁵ (ver. 3.2.1) and AntConc⁶ (ver. 3.5.8) are primarily used to assist data collection and annotation in building up the datasets, while Microsoft Excel and Origin⁷ are mostly used to carry out frequency distributional analyses and draw up visual presentations of the findings.

⁵ R is a free software environment for statistical computing and graphics. More information can be found at <https://www.r-project.org/> (Accessed: 17 March 2022).

⁶ AntConc is a freeware corpus analysis toolkit for concordancing and text analysis, developed by Laurence Anthony. More information can be found at <https://www.laurenceanthony.net/software/antconc/> (Accessed: 17 March 2022).

⁷ Origin is a computer program produced by OriginLab Corporation for interactive scientific graphing and data analysis. More information can be found at <https://www.originlab.com/> (Accessed: 17 March 2022).

Chapter 2

Theoretical background

In this chapter, important notions and theories, on which this thesis is based, will be introduced and clarified. Section 2.1 focuses on the concept of discourse markers. Section 2.2 introduces the theoretical framework of the thesis, namely Construction Grammar. Section 2.3 presents the research field of language change in the light of grammatical Constructionalization with a special focus on the emergence of discourse markers. Section 2.4 discusses and distinguishes some important notions in discourse organization, which are of special relevance to the functional analyses of SPOX and *huashuo*, including the concepts of discourse coherence and topical relevance, different terminologies and categorizations in the functional domain of topic introduction, as well as definitions and distinctions of discursive and pragmatic pretexts, periphery, and monologual and dialogual contexts.

2.1 Discourse markers

2.1.1 Terminology

Discourse markers are known to be a fuzzy category. They can refer to a large spectrum of linguistic elements, ranging from set phrases and clauses to interjections and exclamations, and cover a vast variety of features, both formal and functional. A variety of terms have been

used in the studies of discourse markers, including, but not limited to, “discourse marker” (e.g. Jucker 1993; 1997; Lenk 1998a; Lewis 2011; Müller 2005; Schifffrin 1987; Schourup 1999; Urgelles-Coll 2010), “pragmatic marker” (e.g. Aijmer 2013; Aijmer & Simon-Vandenberg 2006; Andersen 2001; Beeching 2016; Brinton 1996; 2008; 2017; Fraser 1996; Lauwers et al. 2012; Redeker 1990), “discourse particle” (e.g. Abraham 1991; Aijmer 2002; Bayer & Struckmeier 2017; Diwald 2006a; Fischer 2006a; Hansen 1998; Schourup 1982), “pragmatic particle” (e.g. Östman 1981), “discourse connectives” (e.g. Blakemore 1987; Celle & Huart 2007; Warner 1985), “pragmatic connectives” (e.g. Van Dijk 1979), and “style disjuncts” (e.g. Quirk et al. 1985: 1072–1073). In some studies some terms are treated as synonyms and used interchangeably, e.g. Hansen (1998) for “discourse marker” and “discourse particle”, Lauwers et al. (2012) for “discourse marker” and “pragmatic marker”, Aijmer (2002; 2013) for “discourse particle” and “pragmatic marker”, and Andersen (2001) for “discourse marker”, “pragmatic marker” and “pragmatic particle”. However, in other studies these terms are distinguished from each other, e.g. Fischer (2006a) distinguishes “discourse marker” from “discourse particle”, and Fraser (1996) and Redeker (1990) both distinguish “pragmatic marker” from “discourse marker”. Different scholars seem to have their own preferences in choosing a certain term. And oftentimes they do not agree on the definitional and categorizational schema, even when they use the same terminology. It has been a well-known challenge in this field that so far there has been little consensus either on the definition of the term or on the members of the group.

There have been a number of studies addressing the terminological issue. Some intend to establish some sort of structure amidst different voices (e.g. Degand et al. 2013; D'É 2010; Schourup 1999), and others aim to clarify their understanding of a specific term (e.g. Brinton 2017: 2–11; Fraser 1990; Watts 1988: 236–250; Haselow 2015: 80–89). There have also been studies dedicated to sorting out what should qualify as a member of this category (e.g. Fischer 2006a; Fraser 1999; Heine 2013: 1206–1213; Jucker & Ziv 1998). In this thesis, it is not my goal or interest to settle these issues. I use “discourse marker” as an umbrella term for all the competing terms used with partially overlapping reference, as it seems to cover the widest range of instances and is less strictly syntactically defined. I follow the broad definition proposed in Heine (2013: 1206–1213) and treat any linguistic element that is used as a discourse management device as a member of the category of discourse marker. That is to say, expressions/constructions such as modal particles, speech act adverbials, conditionals, comment clauses, etc. can be understood as discourse markers with different syntactic features and specific functions. In addition, I adopt a constructional approach to the understanding of discourse markers: They always represent a pairing between a specific form and a specific functional role in a specific linguistic context, and therefore should always be understood in relation to usage (more discussion about Construction Grammar follows in Section 2.2).

2.1.2 Characteristics

In previous studies, different scholars have observed a wide variety of characteristics of discourse markers from different perspectives. Some characteristics are considered core, while others are considered peripheral, i.e. optional; some are categorical/distinctive, while others are rather descriptive/general; some concern the formal aspects, while others focus on the functional aspects. Integrating the different observations, I focus on the following characteristics that separate discourse markers from other grammatical categories.

First, **discourse markers are procedural elements rather than conceptual-propositional**⁸. They have non-referential meaning, i.e. their primary functions in discourse are to give instructions to the hearer for processing certain proposition, rather than to refer to certain entities or aspects of the world. In other words, they are not used to compose the propositional content of the utterance but to provide metalinguistic information with respect to the speech act and possibly the interlocutor relation, which in turn influences the processing and interpretation of the utterance. For example, the discourse marker *well* in (15) retains little, if any, of the propositional meaning of the adjectival/adverbial *well* as in (13) and (14). In (16), the post-verbal adverbial *seriously* is part of the proposition and it modifies the predicate verb “answer”, while the utterance-initial discourse marker

⁸ See Blakemore (1987; 2002) and Wilson and Sperber (1993) for the *conceptual-procedural* distinction in the Relevance Theory, and also Fraser’s (1988; 1996) *content/propositional meaning* vs. *pragmatic meaning* distinction in defining discourse markers.

seriously signals that the speaker's manner of speaking is serious and does not concern the content of his/her utterance.

- (13) *I hope this letter finds you well.*
- (14) *He drives well.*
- (15) *Well, if you put it that way, there is nothing more I can say.*
- (16) *Seriously, you should have answered seriously.*

Second, **discourse markers are syntactically non-integrated elements**. As their meaning is not part of the propositional content of the sentence, their presence or absence does not affect the truth conditions of the proposition contained in the utterance. Syntactically, this is often manifested in their detachability and mobility, i.e. they often have loose syntactic connections to the rest of the utterance and can be detached from the sentence without hurting its grammatical integrity, and they can be used in different positions of the utterance. For example, the comment clause *I think*, denoting the speaker's epistemic stance, can be used either as an embedded main clause or as a detached syntactic element, and in utterance-initial, -middle, and -final positions, as shown in (17) to (20).

- (17) *I think (that) it was last night that he finally came home.*
- (18) *I think, it was last night that he finally came home.*
- (19) *It was last night, I think, that he finally came home.*
- (20) *It was last night that he finally came home, I think.*

Prosodically, discourse markers are often set off from the rest of the utterance and form separate tone units. This also reflects their non-participation in the propositional content of the utterance.

Third, **discourse markers are “short” items**. This usually applies to the morpho-syntactic structures and/or phonological features. Discourse markers often exhibit structural and/or phonological reduction in comparison to their lexical counterparts, e.g. *y’know* vs. *you know* (Schiffrin 1987: 267–295), /nfækt/ or /fæk/ vs. *in fact*, /əmi:n/ or /mi:n/ vs. *I mean* (Brinton 2017: 4, 21).

More importantly, this characteristic should be interpreted with a constructional understanding. It alludes to the constructional status of discourse markers. Discourse markers are not necessarily short in the formal sense, such as in the case of multi-word phrasal or clausal discourse markers like *and stuff/things like that* (Overstreet 1999; Aijmer 2002: 211–249) and *Is it just me or ...* (Bai 2014; 2015). Rather, they are short in the sense that they are perceived, processed, stored and used as one unit of formulaic feature, instead of as a composition of smaller units. The whole structure, regardless of its physical length, is one construction with a specific form-function correspondence.

Lastly, **discourse markers are typically polyfunctional**. They often operate on different planes of the discourse (e.g. textual, illocutionary, interpersonal, etc.) and take effect in different functional domains (e.g. coherence, politeness, evidentiality, etc.). For example, Haselow (2011) observed that the utterance-final *then*, such as in (21) and (22), covers a full functional spectrum from ideational to textual and to metalinguistic levels. It provides the contextual coordinates of

the utterance, expresses subjectivity, modifies the illocutionary force, and attends to politeness management.

- (21) A: *oh he's fairly happy*
B: *why do you think he doesn't write then?*
(Haselow 2011: 3610)

- (22) A: *I haven't the faintest idea what you're talking about*
B: *well you have to listen to the tape then*
(Haselow 2011: 3611)

In sum, discourse markers are procedural elements that contribute to non-propositional aspects of the discourse. They are versatile discourse management devices, which are used to meet a wide range of communicative needs. Their quality and behavior can be better understood by adopting a constructional approach to language.

2.2 Construction Grammar

This thesis develops within the theoretical framework of Construction Grammar, in which our grammatical knowledge is understood to be made up of a dynamic taxonomic network of constructions (cf. Croft 2001; Diessel 2015; Fillmore 1988; Fillmore et al. 1988; Goldberg 1995; 2003; 2006; Hilpert 2014; Hoffmann & Trousdale 2013; Lakoff 1987).

The concept of a *construction* in Construction Grammar needs to be distinguished from the loosely used descriptive label that simply

refers to a linguistic expression consisting of several parts. Instead, *constructions* are defined as pairings between a particular structural pattern and a specific function or meaning, and they are considered to exist at all levels of linguistic representation. They can refer to idiomatic expressions that do not follow canonical syntactic rules or whose meanings are non-compositional (e.g. Bai 2014; Bender & Kathol 2001; Kay 2002; Kay & Fillmore 1999; Van Eynde 2007; Zwicky 2002); they can refer to grammatical units such as phrasal and clausal structures (e.g. Davidse 2000; Goldberg & Van der Auwera 2012; Hilpert 2009; Kay 1984; Michaelis & Lambrecht 1996); and they can refer to smaller units such as morphemes and lexemes as well (e.g. Boyd & Goldberg 2011; Dąbrowska 2009; Michaelis 2013). The formal structure of a construction can contain (partially) concrete and particular lexical items, e.g. *kick the bucket*, *all of a sudden*, the *let alone* construction (X A Y *let alone* X B Y) (Fillmore et al. 1988); or it can be fully abstract, e.g. the ditransitive construction [S V NP NP] (Croft 2003; Goldberg 1995: 141–151; Perek 2015) and the caused-motion construction [S V NP PP] (Goldberg 1995: 152–179). A short list of selected examples is given below in order to offer a glimpse of the broad spectrum of constructional representations in the linguistic system.

- (23) The JB-X DM-Y construction:

Just because we live in Berkeley doesn't mean we're left wing radicals.

(Bender & Kathol 2001: 13)

- (24) The *just me* construction:

Is it just me or is it hot in here? (Bai 2014: 128)

- (25) The *kind of/sort of* construction:
A mastodon is kind of an elephant. (Kay 1984: 157)
- (26) The *is-to* construction:
The match is to begin at 11 pm.
(Goldberg & Van der Auwera 2012: 112)
- (27) A-adjectives:
afloat, alive, asleep, etc. (Boyd & Goldberg 2011)
- (28) The *let alone* construction:
Fred will not eat shrimp, let alone squid. (Fillmore et al. 1988: 512)
- (29) The [N *waiting to happen*] construction:
Bands like that are accidents waiting to happen in a world where [...]
(Stefanowitsch & Gries 2003)
- (30) The ditransitive construction:
Sally baked her sister a cake. (Goldberg 1995: 141)
- (31) The caused-motion construction:
Frank sneezed the tissue off the table. (Goldberg 1995: 152)

Construction Grammar is an integral part of the usage-based approach in linguistics. It differs from the generativist tradition by its fundamental theory that linguistic knowledge is acquired in a bottom-up manner through use. Instead of postulating an innate language faculty in the human mind that governs a speaker's "Universal Grammar" (Chomsky 1957; 1965; 1995), the usage-based model treats grammar as an "emergent phenomenon" shaped by general cognitive mechanisms such as categorization, analogy, and entrenchment and derived from linguistic experience (cf. Auer & Pfänder 2011; Hopper 1987; 1998). In other words, the usage-based

approach rejects the presupposition of an innate inventory of atomic syntactic categories and rather proposes that grammatical structures emerge as generalizations over recurrent encounters and uses of concrete expressions with similar forms and meanings (see also Bybee 2006; 2010). Studies in Construction Grammar do not separate the “competence” from the “performance” part of language, nor do they postulate a rigid distinction between the “core” and the “periphery” of the linguistic system. They set out to account for the entirety of linguistic phenomena. Language-specific generalizations across constructions are understood in terms of an “inheritance network”, in which less general constructions inherit properties from other more general constructions (Goldberg 2003; Kay & Fillmore 1999). Cross-linguistic generalizations are explained at the level of cognitive capacities and constraints as well as general functional principles (Croft 2005; Goldberg 2003).

The constructionist approach to language has attracted linguists with different backgrounds, as it offers exciting new insights into different fields of linguistic studies. A large body of constructionist research exists in various research areas such as psycholinguistics, first and second language acquisition, discourse analysis, variationist (socio)linguistics, contrastive linguistics, historical linguistics, and many more.

This thesis adopts the Construction Grammar approach to study the behavior and development of discourse markers from a cross-linguistic perspective. It is thus situated at the crossroads of several different research areas. The following subsections offer a brief introduction to each of the relevant research fields individually.

2.2.1 Discourse studies and Construction Grammar

With an increasing number of scholars acknowledging the close relationship between discourse and grammar and the importance of studying them together (e.g. Du Bois 1987; 2003; Fox & Thompson 1990; Givón 2018; Prince 1981; Heine et al. 2013; Hopper & Thompson 1980; Kaltenböck & Heine 2014; Kaltenböck et al. 2011), a growing body of studies has been carried out in the intersection of discourse and Construction Grammar. Some studies focus on the importance of integrating discourse phenomena and pragmatic aspects of constructions into Construction Grammar studies (e.g. Bai 2014; Bergs & Diewald 2009; Fischer 2010; Fried 2010; Günthner & Imo 2006; Imo 2006; Linell 2009; Östman 2005; Pirc 2013; Steen & Turner 2013; Zima & Bergs 2017). Other studies rather explore how discourse studies can profit from adopting a grammatical perspective in addressing specific discourse phenomena (e.g. Aijmer 2016; Alm et al. 2018; Antonopoulou & Nikiforidou 2011; Ariel 1998; 2009; Bai 2015; Fischer 2015; Fischer & Alm 2013; Fried & Östman 2005; Geluykens 1992; Imo 2014; Kay 2004).

The constructionist view of language, in which the formal, semantic and pragmatic aspects of language use are all treated equally in one unit, has proven to be particularly useful in the field of discourse studies. First of all, the understanding of grammatical representations as form-function pairings is highly compatible with the nature of real-world linguistic phenomena and especially with the production and perception of discourse patterns, which typically not only involve substantive (lexical) elements but also more schematic

(grammatical/structural or thematic) and conventionalized (idiomatic) features. Second, Construction Grammar takes all chunks and facets of language – large and small, verbal and non-verbal – as equally deserving of description and analysis. It aims to describe real-world linguistic behavior and to account for linguistic creativity and contextual felicitousness. Third, the discourse-analytic notion of “context”, which determines the functions and formal properties of a certain expression, is an integral part of Construction Grammar. Lastly, the notion of “meaning” in Construction Grammar comprises both lexical/semantic and pragmatic meanings as well as metalinguistic information relating to the discursive environment, interlocutor relation, information structure, encyclopedic knowledge, etc. Construction Grammar is therefore a well-suited and powerful framework that helps to enhance our understanding of the mechanisms of discourse and allows us to explicate discourse structures and processes in a more systematic manner (see also Croft & Cruse 2004: 258; Fischer & Alm 2013; Fried & Östman 2005: 1754–1755; Kay & Michaelis 2012).

The constructionist approach does not mean that expressions and utterances are interpreted as fixed or static chunks. On the contrary, our linguistic production is viewed as a dynamic network of schematic and specific units, which are interconnected both through an inheritance network (vertical links) and by semantic/structural analogy and metaphorical extensions (horizontal links) (Diessel 2019; Goldberg 1995: 72–81; Sommerer & Smirnova 2020). The constructional network is further influenced by factors such as usage frequency and is constantly being updated and always

context-dependent. In this thesis, discourse markers are treated and analyzed as part of the constructional network. Their meanings are not construed in stagnation or isolation but always in relation to other constructions and contingently under specific textual and interactional conditions at the time of speech.

2.2.2 Cross-linguistic Construction Grammar

Even though Construction Grammar has historically focused on studies of the English language and much constructional research nowadays still focuses on single languages, there are an increasing number of studies exploring the extent to which the Construction Grammar approach can be employed for cross-linguistic analysis. These studies can be categorized broadly into three groups according to the issues they address and the cross-linguistic perspective they take.

The first group of studies explores the validity of applying Construction Grammar to other languages than English. Some of them focus on single languages (e.g. Bisang 2008; Boas & Gonzálvez-García 2014; Bouveret & Legallois 2012; Fried & Östman 2004; Hilpert 2006; 2008: 49–87; 2009; Östman 2006; Peng 2017; Tsujimura 2005; Zhang 2018), while others are contrastive studies – with English as the reference language most of the time (e.g. Boas 2003: 285–312; 2010; Fischer & Alm 2013; Gilquin 2015; Gonzálvez-García 2017; Hilpert 2008: 89–123; Keizer 2016; Nođ & Coleman 2009; 2010; Pedersen 2009; 2013). These studies show that the analytic and representational tools of Construction Grammar can

indeed be successfully applied to the description, analysis, and explanation of diverse linguistic phenomena in a variety of languages.

The second group of studies is interested in the potential for Construction Grammar to capture cross-linguistic generalizations, i.e. a typologically relevant conception of language universals (e.g. Croft 2005; Croft et al. 2010; Pedersen 2008). As constructions can be language-specific, so can grammatical categories – the construction used as a diagnostic for a syntactic category in one language may be absent in another language. As a result, valid cross-linguistic generalizations can only be captured by the architecture of the representation system, i.e. by how meaning and function are encoded in linguistic forms, and by the sharing of abstract constructions across languages (cf. Boas 2010: 1–20; Croft 2001: 363; Kay & Fillmore 1999). The conception of constructions as the basic units of all languages is thus a viable descriptive and analytical tool for cross-linguistic comparisons that can capture both language-specific (idiosyncratic) properties as well as cross-linguistic generalizations.

The third group of studies takes language contact and bi/multilingualism into consideration (e.g. Boas & Höder 2018; Hilpert & Östman 2016; Höder 2012; Ziegeler 2015). They investigate whether Construction Grammar can be used to model the linguistic knowledge of speakers with competence in multiple grammars and explore how such speakers make use of constructional generalizations. In this respect, they underline the necessity to recognize Construction Grammar as a dynamic model of linguistic competence, in which constructional knowledge includes both intra-linguistic variation, i.e. speakers' ability to use different varieties and their ability to

accommodate to different interlocutors, and cross-linguistic competence, i.e. speakers' ability to form meta-generalizations that connect corresponding constructions from different languages in their linguistic repertoires.

This thesis develops a cross-linguistic perspective in the study of discourse markers. It compares the form, function, and usage patterns of the English discourse marker *speaking of X* with those of the Chinese counterpart *huashuo*. It thus adds to the body of work in cross-linguistic constructional studies that supports the use of the constructional approach in analyzing and comparing typologically different languages.

2.2.3 Diachronic Construction Grammar

Diachronic Construction Grammar refers to the research field that joins Construction Grammar and historical linguistics, in which various facets of language change are addressed with a constructionist approach (cf. Barðdal et al. 2015; Bergs & Diewald 2008; Coussée et al. 2018; Fried 2013; Hilpert 2013; 2018; Traugott & Trousdale 2013). In this field, given the conception of constructions as the basic units of a linguistic system, a wide range of phenomena of language change are identified and analyzed as constructional changes, and the gradient process of language change is comprehended as the sum of many individual transformations of constructions. In comparison to the traditional rule-based approach to explaining language change, the construction-based approach integrates form and meaning, concrete expressions and abstract schemata, and therefore lends itself

particularly well to addressing the often unpredictable and idiosyncratic phenomena in semantic and pragmatic development, which often involve concurrent changes in different components and on different levels of a particular linguistic representation. Furthermore, the nature of constructions invites the understanding that changes may happen in any part – or the combination of several parts – of the constructional architecture, including the form, the meaning, and the mapping between them. This assumption helps to systematize different types/stages of constructional change and to explain different diachronic phenomena within a language and how the processes of change may differ between languages. It is therefore a highly useful and especially well-suited framework in comparing the development paths of expressions from linguistically heterogeneous languages, e.g. the English *speaking of X* construction and the Chinese *huashuo* construction in this thesis.

There is a growing body of studies that addresses diachronic change in the framework of Construction Grammar. Quite a few focus on identifying and/or distinguishing different processes of change by way of analyzing individual cases of language change, while emphasizing the importance of adopting a constructional approach (e.g. Diewald 2006b; Fried 2008; 2009a; 2013; Hilpert 2008; Hundt 2014; Israel 1996; Norde & Trousdale 2016; Sommerer 2015; Traugott 2008a; 2008b; Trousdale 2008a; 2008b; Trousdale 2010; Van de Velde 2014). Other studies rather explore general factors that induce constructional change and/or influence the type of constructional change, including inter alia cognitive mechanisms such as analogy and blending (Bai 2014; Fischer 2018), contextual factors

such as syntagmatic patterns and discourse-pragmatic information (e.g. Fried 2009b; Smirnova 2015; Waltereit 2012), frequency and saliency (e.g. Bybee 2003; Lewis 2011; Petr é & Cuyckens 2008), and social factors such as language contact (e.g. Bisang 1998; Leino & Östman 2008; Mithun 2008). This thesis combines the interests of both groups of studies and discusses not only the processes of change but also the motivating factors involved in the emergence and development of discourse markers (Section 3.4 and Section 4.4).

As the constructionist approach in diachronic studies gives rise and prominence to new concepts such as *constructionalization* and *constructional change*, a few studies set out to look into issues concerning their relationship to other existing theories of language change, including grammaticalization theory (e.g. Gisborne & Patten 2011; Nođ 2007; Trousdale 2012; 2014). Depending on how grammar and construction are defined respectively in these theories, different scholars hold different opinions on which phenomena should be categorized under which label and consequently on the purview of these terms and their relations. Nonetheless, vigorous effort has been made in integrating the theoretical frameworks of diachronic Construction Grammar and grammaticalization, i.e. identifying the roles of constructions in grammaticalization on one hand, and relating different processes in grammaticalization to different types of constructional change on the other (e.g. Hilpert 2013; Langacker 2005; Traugott 2015a; Traugott & Trousdale 2013). More discussion and clarification of the different processes of language change will follow in Section 2.3.

2.3 Grammatical constructionalization

Combining Construction Grammar with grammaticalization theory, Traugott and Trousdale (2013) address a range of theoretical topics concerning language change. They first distinguish between *grammatical constructions* and *lexical constructions* as the basic units of language. Acknowledging the gradient nature between these categories, *grammatical constructions* are defined as elements that have primarily procedural, linguistically relational, and non-referential semantics, while *lexical constructions* are defined as those that have primarily contentful and referential semantics. Then, on the basis of that distinction, they further introduce the notions of *grammatical constructionalization* and *lexical constructionalization* to refer to the types of change that result in the respective types of constructions.

This approach has several advantages in addressing the issue of language change. First, it acknowledges the gradient relationship between grammar and lexicon and thus integrates the traditional notions of *grammaticalization* and *lexicalization* into one framework. Second, it unites two contrasting views of grammaticalization: one that views grammaticalization as reduction in form and independency (e.g. Lehmann 1995), and the other rather as expansion of contexts and functional scopes (e.g. Himmelmann 2004: 31–34). The opposition between the reductionist view and the expansionist view rises from the dichotomy between a modular model of grammar and a holistic model of grammar. As reduction processes often pertain to phonological and morpho-syntactic changes, and expansion

phenomena often characterize semantic and pragmatic changes, the term *grammaticalization* is in fact used to refer to quite different types of language change in these two views. A modular definition of grammar, which treats syntax and semantics as separate components, limits the purview of grammaticalization, so the use of the term is problematic for certain cases. In a holistic model like Construction Grammar, by contrast, both formal and semantic changes are simultaneously accounted for by the notions of *constructional change* and *constructionalization*. The terminological issue of grammaticalization hence becomes irrelevant. Lastly, the constructionist approach defuses the directionality debate, i.e. whether or not unidirectionality should be considered a universal and intrinsic property of grammaticalization (e.g. Campbell 2001: 124–141; Haspelmath 2004; Janda 2001; Newmeyer 1998: 260–278; Norde 2001; 2009: 48–105). This issue is not criterial within the framework of grammatical constructionalization, as it provides a more encompassing model for language change than as defined under traditional views of grammaticalization.

This thesis follows the definitions proposed in Traugott and Trousdale (2013) and uses *grammatical constructionalization* as the cover term for processes involved in the development of discourse markers. In existing diachronic studies of discourse markers, a range of different terminologies has been used, which either account for different types of change or represent different understandings of the status of discourse markers in language. The following subsections aim to clarify these notions, including *grammaticalization*, *lexicalization*, *constructionalization*, *pragmaticalization*, and

(inter)subjectification, especially in light of their relation to the emergence and development of discourse markers.

2.3.1 Grammaticalization

According to Lehmann (1995: 11; 2015: 11), grammaticalization is “a process which may not only change a lexical into a grammatical item, but may also shift an item from a less grammatical to a more grammatical status”. On that basis, Hopper and Traugott (2003: 18) define grammaticalization as “the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions.”

Discourse markers have long been a controversial case for grammaticalization theory due to their ambiguous grammatical status and obscurity as a linguistic category (see Section 2.1 for relevant discussion). Scholars have different views on whether or not their development should be accounted for in terms of grammaticalization. And the answer to that question, as pointed out by Degand and Evers-Vermeul (2015: 74), largely depends on the conception of grammar and the set of (sub)processes that are considered to be criterial for grammaticalization.

Scholars who are in favor of including discourse markers in grammaticalization studies hold the expansionist view of grammar, so that grammaticalization includes movement towards discourse functions as well (e.g. Boye & Harder 2012; Brinton 1996: 50–65; 2007; 2008; Diwald 2006a; 2011; Schwenter & Traugott 1995;

Traugott 1995a). Opponents to this approach rather view the development of discourse markers as distinct from grammaticalization and resort to other descriptive notions such as *lexicalization* (Section 2.3.2) and *pragmaticalization* (Section 2.3.3). In short, “if grammar is restricted to semantics, phonology, morphology and syntax, DMs will not be viewed as the result of a grammaticalization process; if, in contrast, grammar is viewed as including discourse-pragmatic phenomena, then DMs will most probably find a satisfying account in terms of grammaticalization” (Degand & Evers-Vermeul 2015: 61). In an attempt to resolve this issue, some scholars have proposed that perhaps there is no need to dwell on the different definitions of grammaticalization, if discourse markers are discussed in a different framework entirely, e.g. in terms of Thetical Grammar and Discourse Grammar (Heine 2013; 2018; Heine et al. 2013; Kaltenböck et al. 2011) and of Construction Grammar (Hilpert 2013; Traugott & Trousdale 2013).

As to the (sub)processes involved, the development of discourse markers exhibits many characteristics that are typically associated with grammaticalization. For example, the following processes and/or phenomena (a) to (f), as detailed in Lehmann (1995; 2015) and Hopper (1991), apply to the evolution of discourse markers:

- (a) *phonological attrition*: gradual loss of phonological substance (which often goes hand in hand with *desemanticization* – semantic bleaching), e.g. Latin *hāc hōrā* ‘at this hour’ -> Portuguese *agora* and Spanish *ahora* ‘now’ (Lehmann 2015: 4), the French verbal complement *à (ce) propos (de)* ‘at (this/the)

subject (of)' -> topic-shift marker *à propos* 'by the way' (Prevost 2011);

- (b) *paradigmaticization*: integration into a new paradigm or grammatical category, e.g. *during* from participle to preposition (Lehmann 2015: 4), *then* from temporal adverb to utterance-final discourse marker (Haselow 2011);
- (c) *layering*: newly developed grammatical forms may carry similar or even identical functions as the older ones, hence adding new layers to the existing functional domain, e.g. *be going to* layered with *will* in the functional domain of future time reference (Hopper 1991: 23), the adversative marker *only* layered with older adversative conjunctions and adverbials, such as *but*, *yet*, *nevertheless*, *however* (Brinton 2008: 54);
- (d) *divergence*: the development of the new grammatical item may be a “split” phenomenon as the original item continues to exist, e.g. the aspectual auxiliary *have* diverging from *have* as a lexical verb of possession (Hopper 1991: 25); the comment clause *I think* diverging from its origin as a matrix clause (Thompson & Mulac 1991: 324–325);
- (e) *persistence*⁹: some traces of the original lexical meaning persist in the grammaticalized forms, e.g. the volitional meaning of *will* may persist in the future auxiliary *will* (Olga Fischer 2007a:

⁹ This parameter has later been extended to include not only lexical persistence, but also structural and procedural persistence (Breban 2009; Hancil 2018).

119), the parenthetical *I mean* has both its original meaning of intention and the newly developed pragmatic meanings of explicitness, reformulation, etc. (Brinton 2007);

- (f) *decategorialization*: items undergoing grammaticalization tend to lose their (typically nominal or verbal) categoriality and take on secondary/ancillary roles in forming an utterance, e.g. *considering* from verb/participle to preposition/conjunction (Hopper & Traugott 2003: 108; Kortmann & König 1992), French *alors* from a clause-internal adverbial with temporal meaning to a connective marking temporal, causal or conditional relations and further on to a conversation-structuring discourse marker (Degand & Fagard 2011).

Other processes (g) to (j), however, are not necessarily involved in the development of discourse markers. Especially the processes of *condensation*, *coalescence*, and *fixation* are very often opposite to what can be observed in the case of discourse markers.

- (g) *obligatorification/specialization*: transition from an optional or variable item to an obligatory/specialized one in its functional domain, resulting in the decrease of possible choices within that functional domain, e.g. *ne...pas* becoming the standard negation form in French;

- (h) *condensation*: reduction of syntactic scope and degradation to a lower level of grammatical structure, e.g. demonstrative pronoun -> definite article;
- (i) *coalescence*: increase of syntagmatic bondedness, e.g. the development of clitics and affixes;
- (j) *fixation*: decrease in syntactic freedom, e.g. English noun *while* -> conjunction *while*.

It has been shown in different studies that the development of discourse markers only conforms to some tenets of the classical grammaticalization theory, but not to all (e.g. Barth-Weingarten & Couper-Kuhlen 2002; Beijering 2015; Brinton 2007; Günthner & Mutz 2004). Simon-Vandenberghe & Willems (2011) further observe that some criteria of grammaticalization apply to some types of discourse markers, but not to others. In particular decategorialization, bondedness and syntagmatic fixation vary from one discourse marker to the next. Alternatively, Traugott (1995a) adds pragmatic strengthening (Section 2.3.3) and subjectification (Section 2.3.4) to her expansionist model of grammaticalization and disregards the decrease in syntactic freedom and scope to account for the diachronic change of discourse markers.

Within the Construction Grammar framework, this thesis treats grammaticalization as constructionalization towards the procedural pole of the functional spectrum, i.e. *grammatical constructionalization* in Traugott and Trousdale's (2013) term. At the same time, it draws on Boye and Harder's (2012) distinction between *primary/central* and

secondary/ancillary functional roles and views the development of discourse markers in terms of a shift from being independent linguistic elements that convey (part of) the main/propositional information of an utterance to elements that are dependent upon other expressions and convey supplementary/metalinguistic information about those expressions.

2.3.2 Lexicalization and constructionalization

Lexicalization and *constructionalization* are important notions in understanding the emergence of discourse markers, as they concern processes that lead to the addition of new form-function pairs to the inventory of existing linguistic knowledge.

The term *lexicalization* is used differently in different lines of linguistic studies. In synchronic studies, and very often from a typological perspective, it refers to the lexical representation of conceptual categories, i.e. how conceptual structures are expressed through words/lexemes (e.g. Proost 2007; Talmy 2000: 21–212; 2007). From a historical perspective, it carries the general meaning of “adoption into the lexicon/inventory” (see Brinton & Traugott 2005; 2007; Lehmann 2002), which then accommodates different further interpretations.

Firstly, *lexicalization* is used quite often as an alternative term to refer to ordinary word formation processes, covering instances like compounding, clipping, blending, etc. (e.g. Blank 2001; Van der Auwera 2002). But Brinton and Traugott (2005: 89–99) argue that word formation and lexicalization should rather be treated as separate

phenomena. According to their definition, lexicalization is an unconscious process of change in usage that leads to “new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the [original] construction or word formation pattern”. That is to say, deliberate and conscious word formation practices that lead to “one-off” coinages are excluded from the notion of lexicalization; typical word formation processes such as borrowing, conversion, and clipping do not qualify as instances of lexicalization either, as they produce new items with predictable forms and meanings. Word formation may precede or even motivate lexicalization, but these processes need to be viewed as independent of each other.

Secondly, the term *lexicalization* is also used in grammaticalization studies. It is often discussed alongside the notion of *degrammaticalization* as an instance of the opposite, reverse, or mirror image of grammaticalization, i.e. more grammatical -> less grammatical -> lexical, or syntax -> lexicon (e.g. Campbell 2001; Janda 2001; Moreno Cabrera 1998; Ramat 1992). This conception is usually based on a strict distinction between lexicon and grammar/syntax. Using Himmelmann’s (2004) box metaphor, *lexicalization* under such an understanding refers to the movement from a box of grammatical items into a box of lexical items. But this conception brings about more complications than clarifications, as not all linguistic elements have a clear-cut lexical or grammatical status, as in cases of derivation and inflection (Himmelmann 2004: 21–25). Studies that categorize the emergence of discourse markers as lexicalization are mostly motivated by the difficulty of fitting

discourse markers into traditional grammatical categories. And they focus more on elementary processes such as univerbation, fossilization, idiomatization, etc. to explain the development of discourse markers (e.g. Olga Fischer 2007b; Krug 1998).

In contrast, a construction-based model of language, which recognizes the gradient nature of lexicality and grammaticality, is well rid of such problems. It dissociates lexicalization from degrammaticalization and treats the former as a descriptive term summarizing different processes involved in the creation of new entries in the inventory of linguistic units. It is also noted in numerous studies that lexicalization can overlap with processes involved in grammaticalization, but they occur on different levels of linguistic representation or in different domains, e.g. changes on lexical vs. morpho-syntactic levels, enrichment in semantic vs. pragmatic domains (see Brinton 2002; Fischer 2008; Lehmann 2002; Lightfoot 2011; Trousdale 2008c; Wischer 2000).

Constructionalization, as its morphology suggests, concerns the process in which constructions are coined (cf. Traugott & Trousdale 2013). Its scope encompasses lexicalization: as lexical items are fundamentally constructions themselves, the formation of lexical items is essentially a subtype of formation of constructions¹⁰. It is a broader concept than *grammaticalization* as well: as the conception of constructions does not only apply to grammatical categories but to lexical representations as well, grammaticalization is subsumed under

¹⁰ This insight is also hinted at in Lehmann's (2002), Rostila's (2004; 2006), and Brinton and Traugott's (2005) definitions of lexicalization.

constructionalization as a subtype, namely *grammatical constructionalization* (Hilpert 2013; Trousdale 2014). The notion of *constructionalization* represents a way of reflecting on phenomena and the nature of language change in terms of constructions. In doing so, Traugott and Trousdale (2013: 20–21) point out the importance of taking both the holistic nature of constructions and their internal dimensions (e.g. semantics, morphophonology, collocational constraints, etc.) into account. More specifically speaking, they distinguish between *constructionalization* and *constructional change*: the former pertains to the creation of a new construction and therewith of a new node in the constructional network among a population of speakers, while the latter pertains to the incremental steps affecting the internal dimensions of an existing construction, typically involving expansion of pragmatics, semanticization of that pragmatics, mismatch between form and meaning, and distributional changes, which may gradually amount to the creation of a new construction.

This thesis does not posit a categorical distinction between *lexicalization* and *constructionalization*. Both terms apply to the emergence of discourse markers. I opt for the term *constructionalization* in this thesis in order to focus on discussing the development of discourse markers as the creation of new form-function pairings and the outcome of gradual constructional changes.

2.3.3 Pragmaticalization

Another important notion related to the emergence and development of discourse markers is *pragmaticalization* – the development of pragmatic meaning and a functional shift from truth-conditional/propositional elements to procedural elements.

Some scholars associate pragmaticalization especially with the development of discourse markers and consider it distinct from grammaticalization (e.g. Aijmer 1997; Claridge & Arnovick 2010; Erman & Kotsinas 1993; Frank-Job 2006; Günthner & Mutz 2004; Mosegaard Hansen 2008). This view does not repudiate the similarities between the mechanisms of change involved in these two processes, but it focuses more on the “output” in defining and distinguishing the notions: as discourse markers belong to the domain of discourse instead of the grammatical system of a language, their development should not be considered a case of grammaticalization. The distinction between pragmaticalization and grammaticalization reflects the traditional modular view of grammar, and particularly the separation of pragmatics from the grammatical system (cf. Ariel 2008; 2010). In spite of this, Wiemer and Bisang (2004: 11) still propose that the development of discourse markers belongs to “the fringes” of grammaticalization, as their pragmatic functions are often built in expressions with clear morpho-syntactic structures.

In contrast, with an expansionist or holistic view of grammar, i.e. when grammar encompasses discourse functions and discourse markers are thus grammatical items, pragmaticalization is considered a subtype of grammaticalization or even a superfluous notion. For

example, Diewald (2011) defines pragmaticalization as “grammaticalization of discursive functions”; Wischer (2000: 357) sees pragmaticalization as “grammaticalization on the text or discourse level”; and Degand and Evers-Vermeul (2015), following Traugott (1995a), insist that no clear line can be drawn between pragmatics and grammar and that a separate concept of pragmaticalization lacks justification (see also Barth-Weingarten & Couper-Kuhlen 2002; Brinton 2007; Lewis 2011; 2018; Prevost 2011).

In this thesis I do not use *pragmaticalization* as a categorical term that defines the nature of change that leads linguistic items to discourse markers. Nor do I discard the notion completely. Instead, I use it as a descriptive term to identify the specific processes that feature *pragmatic strengthening* – the enrichment of pragmatic meaning.

Drawing on Claridge and Arnovick (2010), I identify the following mechanisms and processes of change that are often involved in the development of discourse markers, from (a) to (h):

- (a) the development of textual/discourse-oriented meaning out of propositional meaning, e.g. the rise of the causal meaning of *after all* justifying a previous speech act out of the original temporal meaning associated with this prepositional phrase (Lewis 2007; 2018);
- (b) the development of (inter)personal meaning, i.e. (inter)subjectification (see Section 2.3.4);

- (c) semantic bleaching, e.g. the discourse-topic shift marker *by the way* losing its lexical meaning of ‘along the way/beside the road’ (Traugott 2020: 123–128);
- (d) the development of functional polysemy, e.g. the temporal adverb *then* developing new textual, interpersonal and metalinguistic functions which are contextually dependent (Haselow 2011);
- (e) scope expansion, e.g. the development of *in fact* from a clause-internal adverbial to an additive discourse marker that has a scope over the entire proposition (Schwenter & Traugott 2000);
- (f) decategorialization, e.g. *is it just me* losing its clausal status and being recognized as a conventionalized discourse marker (Bai 2014; 2015);
- (g) syntactic and/or prosodic detachment (reflecting propositional non-integration), e.g. the German discourse markers *obwohl* ‘although’ and *wobei* ‘whereby’ indicating reassessment and correction become independent of the subordinate clausal structure and have their own intonation contour (Günthner & Mutz 2004);
- (h) formal reduction (often result of the frequency effect), e.g. the Italian discourse marker *va be* ‘fine/okay’ being the reduced form of *va bene* (Frank-Job 2006: 365–366).

As understood in this thesis, the term *pragmaticalization* thus only refers to the first two processes (a) and (b) – the enrichment of pragmatic meaning on different illocutionary planes. The development of discourse markers involves pragmaticalization, but it is a far more complex enterprise that also involves processes common to other types of constructional change or grammaticalization.

2.3.4 Subjectification and intersubjectification

Subjectification and intersubjectification are important processes of semantic change and are highly relevant to the development of discourse markers. In most straightforward terms, they refer to processes whereby linguistic elements develop (stronger) *subjectivity* and *intersubjectivity*. The notions of *subjectivity* and *intersubjectivity* have been theorized differently by different scholars (cf. Narrog 2017), e.g. from the perspective of “construal” (Langacker 1985; 1990; 1997; 2002; 2006), in relation to evidentiality (Nuyts 2001; 2014), and in terms of speaker/hearer-orientation (Traugott 1982; 1989; 1995b; 1997; 1999; 2003; 2010; see also Traugott & Dasher 2002; Traugott & König 1991).

This thesis follows the approach represented by Traugott. Within this approach, *subjectivity* concerns speaker-orientation. It involves the expression of “self” and the representation of the speaker’s attitudes, perspectives, or viewpoints in discourse. Typical examples of linguistic items that index subjectivity include epistemic modal verbs, speech act adverbials, concessive conjunctions, focus particles, etc. *Intersubjectivity*, on the other hand, concerns

hearer-orientation. It reflects the speaker's attention to the hearer and their assumption and assessment of the hearer's "here and now" in the speech event. It is often manifested in attention-getting and politeness management strategies. Simply said, *subjectivity* concerns how the speech act is related to the speaker's beliefs and attitudes, and *intersubjectivity* concerns how the speech act is designed for the intended hearer. Historically, intersubjectivity tends to develop on the basis of subjectivity, i.e. meanings are typically first recruited by the speaker to encode and regulate their own attitudes and beliefs (subjectification), and, once subjectified, may be recruited to encode information centered on the addressee (intersubjectification). Consequently, intersubjectification can be viewed as a special subtype that represents a further stage of subjectification.

Subjectification is observed to be an overarching tendency of semantic change, i.e. meanings tend to become increasingly based on the speaker's subjective belief state or attitude toward the proposition, and move from weaker subjectivity toward stronger subjectivity. It accounts for a great variety of linguistic phenomena, ranging from the rise of epistemic meaning (e.g. deontic *must* "*You must leave.*" -> epistemic *must* "*You must be hungry.*"), the semanticization of social deictics (e.g. the development of the honorifics systems), performative uses of locutionary verbs (e.g. locutionary *promise* "*He promised that he would protect us.*" -> performative *promise* "*Everything will be fine, I promise you.*"), to the development of discourse markers (e.g. prepositional phrase *in fact* -> self-reparatory / elaborative discourse marker *in fact* "*They were pretty impressed by your presentation this morning. In fact, they loved it.*").

In this thesis *subjectification* and *intersubjectification* are discussed in relation to the emergence and further development of discourse markers, namely how subjective and intersubjective meanings arise during the development of *speaking of X* and *huashuo* as discourse markers. They represent pragmatic strengthening at the (inter)personal level, manifested in the development of self-expressiveness and interpersonal functions.

2.4 Essential concepts in discourse organization

In order to analyze the discourse marking functions and usage patterns of the SPOX construction and the *huashuo* construction, it is important to first clarify a few notions in the context of discourse organization.

In the following sections, Section 2.4.1 introduces the notions of *discourse coherence* and *topical relevance*, which are essential in understanding the need and motivation of using discourse markers like SPOX and *huashuo*. In Section 2.4.2, different types of topic-introducing events are distinguished based on their relation to the global topic structure, including *topic expansion*, *topic change*, and *topic initiation*. After that, the concepts of *discursive pretext* and *pragmatic pretext* are introduced in Section 2.4.3, which are important in understanding the different linking mechanisms of SPOX and *huashuo* as topic-introducing devices. Section 2.4.4 clarifies the definition of *periphery* and introduces the “functional asymmetry” and “left-to-right movement” hypotheses associated with the peripheral behavior of discourse markers. In Section 2.4.5, the notions of

dialogual and *monologual* contexts are defined and distinguished, which are useful in analyzing the usage patterns of SPOX and *huashuo* in turn-taking operations.

2.4.1 Discourse coherence and topical relevance

The need to establish coherence is a central facet of discourse understanding. The establishment of a meaningful discourse representation in the human mind involves distinguishing coherently-linked utterances from those that are not. The execution of such an abstract yet categorical task has long been recognized to be largely relying on general human cognitive instruments such as logic and reason (Grice 1975; Labov 1972: 252–253). It is generally considered that the perception of coherent and incoherent discourse is conditional upon both the semantic interpretation of the utterances and extra-linguistic factors such as the world knowledge of the interlocutors and the common ground shared by them at the time of speech (cf. Dascal & Katriel 1979: 203–205). Semantic relatedness alone does not render discourse coherence, even though it can contribute to it; and vice versa, discourse coherence does not necessarily involve any semantic relations between the utterances.

The following constructed conversation (32) with four alternative answers (A1 – A4) to the question (Q) illustrates this point.

(32) Q: *Hey John, did you pass the exam?*

A1: *I sure did!*

A2: *Karen owes me 5 bucks now.*

2 Theoretical background

A3: *The weather is lovely today.*

A4: **The exam was about the history of America.*

The Q-A1 pairing represents the most natural flow of conversation among the four alternatives. Q points at a problem and A1 solves it directly. A1 is both semantically related to Q and pragmatically effective.

A2, however, does not seem to be semantically related in any way to Q. Still, coherence can be established effortlessly, if both interlocutors have the common ground knowledge about the bet between Karen and John – if John passes the exam, Karen will have to pay 5 bucks over to him. Or better yet, if the question is actually raised with the purpose of finding out the winner of the bet, A2 would provide a natural and direct piece of information that solves the problem put forward by Q.

Similarly, the Q-A3 pairing does not show any propositional connection between the utterances either. To make sense of this transition requires the activation of the world knowledge that talking about the weather could be a strategy of deflecting an unwanted topic. Whether the transition from Q to A3 is accepted as coherent or not depends on whether A3 is thought to have met the communicative demand raised by Q or not. If the person raising the question accepts the deflection of topic as a negative answer, the discourse will be perceived as coherent, even though their question remains unsolved; if the person lacks the needed world knowledge and cannot recognize the pragmatic function of A3, A3 will definitely be taken as an infelicitous response and incoherent to the previous speech.

Lastly, in the Q-A4 pairing, even though A4 shares a semantic component with Q – “the exam”, it is still an incoherent contribution, because the information provided by A4 is not relevant to the topic proposed by Q and therefore it does not meet the communicative need at the time of speech.

Grice (1975: 45–47) stresses the importance of relevance when proposing the Cooperative Principle, which governs a natural and functioning conversation. He uses a cake-baking metaphor to illustrate the importance of a contribution being “appropriate to the immediate needs” for the succeeding utterance to be coherent to the previous one: “[I]f I’m mixing ingredients for a cake, I do not expect to be handed a good book, or even an oven cloth (though this might be an appropriate contribution at a later stage).”

However, the notion of relevance has been recognized as a vague concept by many scholars and it has been defined and classified from different perspectives. For example, Grice (1975: 46) defines the term implicitly in terms of the fulfillment of the communicative need at the corresponding time of speech and he points out the difficulty in classifying relevance. In an attempt to classify this abstract notion, Dascal (1977) proposes a distinction between *pragmatic relevance* and *semantic relevance*: the former concerns goal-directed speech acts and the latter concerns propositions contained in the utterances. Sperber and Wilson (1986) called *relevance* a “fuzzy” term and defined it in terms of the “contextual effect” of a speech act, i.e. the influence of a speech act on updating the common ground and assumptions shared by the interlocutors, and proposed that the notion of relevance is a gradient concept.

As far as this thesis is concerned, the notion of relevance will be understood as *topical relevance*, which inherits the Gricean definition of relevance, emphasizing the linguistic and social contexts at the time of speech, and incorporates the notion of *topicality*.

The notion of *topic* in discourse should not be confused with that defined in terms of the syntactic/semantic features of sentences, i.e. in terms of the theme/rheme or topic/comment distinction (see Firbas 1964; 1975; Halliday 1985). Instead, a conversational topic is a salient piece of information centered by a set of semantically, pragmatically or situationally relevant information, which serves as the background of the ongoing conversation (Dascal 2003: 159–161, 213–223; Dascal & Katriel 1979: 209; Hilpert 2014: 109; Lambrecht 1994: 117–127). That is to say, a topically relevant utterance serves the direct interest of the topic, which could be semantically, pragmatically, or situationally motivated. Semantic relatedness may constitute topical relevance, but it is neither a necessary nor a sufficient condition. This point is illustrated and supported by the following constructed conversation (33) which deals with the topic of “the location of the chocolate”.

(33) X: *Where is the last piece of chocolate?*

Y1: *It's right behind your teacup.*

Y2: *Have you lost your glasses?*

Y3: *You should really try to lose weight.*

Y4: **I like chocolate.*

Despite the great distinction between the answers from Y1 to Y3, all of them are in fact topically relevant. In the X-Y1 pairing, the answer is both semantically related to the question and pragmatically effective: X asks for the location of the chocolate, and Y provides the information about it directly.

In the X-Y2 pairing, none of the semantic components of the answer are related to the question. The connection is rather pragmatically established. As the question implies that X has difficulties finding the chocolate, which could be due to their limited eyesight, the answer mentioning “glasses”, either seriously or jokingly implying “how can you not see it?”, thus still addresses the chocolate-finding topic.

In the X-Y3 pairing, there is also no apparent semantic relation between the sentences. However, the answer is situationally conditioned, involving not only the world knowledge of the effect of chocolate on body weight but also Y’s knowledge about X’s health condition. It conveys the idea of “You should not be looking for that chocolate.” implicitly and is still on topic.

By contrast, Y4 represents an off-topic reply to X, even though it has a semantic relation to it.

As topical relevance goes hand in hand with discourse coherence, when there is a breach of topical relevance, the new utterance would cause an incoherent perception of the flow of discourse. Such potential breaches thus motivate the use of extra linguistic signals to mitigate the effect, by hinting at the shift of topicality, linking the gap between the original and the shifted topics,

and/or compelling the hearer to actively assume unstated information necessary to achieve a coherent construal of the utterances.

Very often, the extra linguistic signals used for such needs are conventionalized discourse markers of various kinds. To name just one example, if Q-A4 and X-Y4 were accompanied by the use of discourse marker *you know*, they would have been perceived as much more coherent. Since *you know* implies that the hearer should know what the speaker is talking about (Müller 2005: 177–181; Östman 1981; Schiffrin 1987: 267–295), it motivates the hearer to search for a logical connection between the seemingly unrelated topics, which thereby establishes topical relevance for the new utterance. In the adapted examples below, *you know* in (32') implies that the person raising the question is supposed to know if the subject of American history is considered difficult or easy to John, which should make the answer to Q pretty obvious even without stating; in (33'), *you know* implies that X should understand what Y means by saying “I like chocolate”, which could very likely be “The reason you can't find the last piece of chocolate is that I ate it.”

(32') Q: *Hey John, did you pass the exam?*

A4': *The exam was about the history of America, you know.*

(33') X: *Where is the last piece of chocolate?*

Y4': *I like chocolate, you know.*

Such operations in the hearer's mind facilitate the establishment of topical relevance and the perception of discourse coherence. But the need for such operations would not be as obvious without the use of

specific discourse markers. In this thesis, *speaking of X* and *huashuo* represent discourse markers of such kind. Their functions will be discussed in detail in Chapter 3 and Chapter 4 respectively.

2.4.2 Digression vs. topic introduction

The term *digression* is often used to describe and define divergent and non-topical discourse events. However, the definition of digression itself has not yet reached a consensus among scholars. In a nutshell, the approaches to define and classify digression differ mostly in two respects: the point of recognition of digression and the conception of topical relevance and its gradience.

Firstly, the term *digression* is used by most scholars in referring to a temporary deviation from the current conversational topic (e.g. Dascal & Katriel 1979; Lenk 1998a; 1998b; Norrick 2001; Pons Border á & Estellés Arguedas 2009; Reig Alamillo 2007). Based on this understanding, a digressive speech act can only qualify as one by the end of the digression when the speaker returns to the original topic. If there is no return, the speech event should not be recognized as a *digression* but a *topic change*¹¹, which means that digressions and topic changes are considered two distinct discourse phenomena.

By contrast, there are some scholars who use the term *digression* in the same sense as *topic change* (e.g. Fraser 2009; Prevost 2011; Traugott 2020). Within this approach, a speech act

¹¹ Other used terms include *conversational switch* by Dascal and Katriel (1979) and *topic shift* by Reig Alamillo (2007) and Pons Border á and Estellés Arguedas (2009).

qualifies as a digression as soon as the new topic deviates from the current one, regardless of the future orientation of the new utterances. This approach has its merits in that it conforms to the fact that the hearer usually can sense digression before knowing if the speaker will ever come back to the original topic and that it is not uncommon for the speaker to digress unintentionally.

Secondly, digression is also defined differently among scholars due to the different conceptions of topical relevance and topical movement. Some scholars stand by a linear conception of topical relevance (e.g. Prevost 2011; Reig Alamillo 2007) and view digression as deviation from the original topic along a topic relevance continuum. Other scholars associate topical relevance rather with the conception of a mental field of consciousness and define digression in terms of the position of an utterance in this field (e.g. Schutz 1970). Some scholars combine the linear understanding of topical movement with the field-like conception of topical relevance and define digression both according to the linear movement of an utterance through the field of consciousness and in terms of its positional relation to the original topic in this field (e.g. Dascal & Katriel 1979; Schank 1977). Furthermore, these different models identify digression with different degrees of topical movement: In Reig Alamillo's (2007) model, digression is understood as a bigger departure from the original topic than in Prevost's (2011) model; Schutz's (1970) point of recognizing digression is further away from the original topic than in Dascal and Katriel's (1979) definition; and Schank's (1977) definition of digression requires the furthest topical movement away from the central stage of the field of consciousness.

In short, according to different definitions, the term *digression* can refer to quite different linguistic phenomena, ranging from topic expansion/continuation to temporary or permanent topic shift/change.

In this thesis, as far as topicality is concerned, I view digressions as equivalent to topic changes, as both events are off-topic in relation to the ongoing topic. When the future orientation of the discourse is taken into consideration, a difference can emerge between whether the new utterance is only a temporary off-topic episode in the global discourse structure or it is going to end the previous topic for good. But this aspect is not relevant in this thesis. The discourse markers *speaking of X* and *huashuo* can be used to mark both temporary and permanent deviations from the original topic. The future orientation of the new topic does not play a role in their usage.

For example, in (34) we can see that, after using “speaking of which” to mention Pete Domenici, with a digression-closing marker *anyway* (see Urgelles-Coll 2010: 107–109), the speaker returns to talking about the events related to the President’s speech, which makes the introduced utterance a temporarily interpolated digressive discourse event in the global topic structure. In example (35), however, the speaker uses “speaking of which” to start a new topic about Molly and ends the previous one about the fisher. The topical movement in this case represents a permanent departure from the original topic.

- (34) *The stock market took its worst dive in 15 months. And we have sent camera crews all around the country to gauge the reaction of taxpayers to the President’s speech last night. Speaking of which, Pete Domenici is the Republican senator from New Mexico. Anyway, his*

2 Theoretical background

press secretary called us today about President Clinton's speech, specifically about the President's contention that the deficit has increased so much beyond his earlier estimates and beyond even the worst official government estimates from last year.

(1993, SPOK, ABC_Nightline, COCA)

- (35) *"I'm a biologist with the D.E.P. We're trying to track down a fisher that's been spotted in these woods." # "What's a fisher?" # "It's a carnivorous predator. Sort of like a bobcat. Lets out a god-awful shriek. Eats small dogs, cats..." # "Next you're going to tell me it eats little girls," Molly scoffed. # "I'm perfectly serious. They wander down from Canada. Speaking of which, where did you wander from?" # "Prunus Cerebus." (2008, FIC, SourCherrySurprise, COHA)*

Similarly, in example (36) we see that P uses *huashuo* to bring up a baseball-watching suggestion in the middle of a conversation about a past event that happened during a baseball game. After receiving a short affirmative reply from N, P returns to the original topic immediately. Therefore, the topic inserted by *huashuo* is only a temporary deviation from the original topic. In example (37), however, *huashuo* is used to shift the conversational topic from package delivery to movie watching. The original topic is changed for good therewith.

- (36) [N and P are talking about a past event that happened during a baseball game.]

N: 你 看棒球 看着 挺好的,
 Nǐ kàn bàngqiú kàn zhe tǐng hǎo de,
 you watch baseball watch-PROG quite fine ADJ

结果 我 告诉 你 你 就 气的 不得了
 jiéguǒ wǒ gàosù nǐ nǐ jiù qì de bù dé liǎo
 in the end I tell you you CNJ angry ADJ extremely

P: 话说 咱 以后 也 要 看棒球 去
Huàshuō zán yǐhòu yě yào kàn bàngqiú qù
speaking of that we later also shall watch baseball go

N: 恩恩
 Ēn ěn
 hm hm

P: 不是 气的 不行 [...]
 Búshì qì de bùxíng [...]
 NEG COP angry ADJ extremely

“N: *You were all fine watching the baseball game. But as soon as I told you about it, you became extremely angry.*

P: *Speaking of that, we should go watch baseball games together sometime.*

N: *Hm hm.*

P: *[I] wasn't extremely angry [...]*”

(3332142874 - Archive (2)(new).txt, 2010, Chat)

2 Theoretical background

(37) [T and W are talking about the logistics of receiving a package delivery.]

W: 需要 我 在 家 等 吗?
Xūyào wǒ zài jiā děng ma?
need I LOC home wait Q

T: 除非 你 知道 他 什么时候 过来 啊~
Chúfēi nǐ zhīdào tā shénme shíhòu guòlái a
unless you know he what time come MOD

T: 不然 就 不用 等 了~
Bùrán jiù bù yòng děng le
otherwise CNJ NEG need wait PRF

W: 好 吧
Hǎo ba
good MOD

T: 麻烦 你 了
Máfan nǐ le
bother you PRF

T: 话说 我 昨天 看的 那个 电影
Huàshuō wǒ zuótiān kàn de nàge diànyǐng
by the way I yesterday watch ADJ that CLF movie
很好看 呢
hěn hǎokàn ne
very good watch MOD

T: 你 有时间 的话 也 看看
Nǐ yǒu shíjiān de huà yě kàn kan
you have time if also have a look

“W: *Do I need to wait for it at home?*

T: *Well only if you know when he comes. Otherwise [you] don't have to.*

W: *Ok.*

T: *Sorry for the troubles. By the way, the movie I watched yesterday was very good. You should check it out too if you have time.”*

(Chat History between T and W 2015-06-25, 2015, Chat)

Therefore, there is little point to distinguish *digression* from *topic change* when analyzing the topic-introducing functions of SPOX and *huashuo*. Instead, I draw the distinction between **topic expansion**, **topic change** and **topic initiation** events, according to their relations to the global topic structure. *Topic expansion* refers to minor topical movements such as movements from subtopic to subtopic under the same hyper-theme. *Topic change*, by contrast, refers to a prominent deviation from the ongoing topic, i.e. the introduction of a new topic that has little to no topical relevance to the previous discourse. A *topic initiation* event refers to the introduction of a brand new topic, and often the start of a new discourse event therewith, without the existence of any immediate discursive pretext – when there is no topical deviation involved. More detailed discussion and exemplification of these different topic-introducing events will be given in the functional analyses of SPOX and *huashuo* in Section 3.3.1 and Section 4.3.1 respectively.

2.4.3 Discursive and pragmatic pretexts

The notion of *discursive pretext* is defined in this thesis as propositional information that is linguistically expressed in the preceding discourse, i.e. verbalized propositional content of the existing discourse, which has thereby become given knowledge in the communicative situation. The notion of *pragmatic pretext*, following the definition given by Diewald and Fischer (1998), by contrast, refers to a proposition that is pragmatically given in the communicative situation, but which has not been linguistically encoded in the preceding discourse.

The concept of pragmatic pretext is often mentioned in studies of German modal particles (e.g. Fischer 2000; 2006b; 2007; Diewald 2013). For example, Diewald (2006a) observes that the use of *eben*, as shown in (38), points at the existence of a pragmatic pretext that the speaker regards the proposition “German is difficult” as given and as communicatively understood.

- (38) *Deutsch ist eben schwer.* (Diewald 2006a: 416)
“German is, as we all agree, difficult.” (translated by the author)

Haselow (2011) explains the notion of pragmatic pretext with his study of the utterance-final *then*, as in example (39). He shows that the use of *then* indicates a contrast between the speaker’s expected information supply and the actual information s/he draws or rather fails to draw from the previous conversation. It is this pragmatic

pretext that motivates the utterance, instead of any relation to a specific proposition located in the preceding discourse.

- (39) *So what did you do today then?* [ICE-GB s1a-025]
(Haselow 2011: 3613)

The indexical relation to a pragmatic pretext is also implied in Bolden's (2009) study of the incipient discourse marker *so*, as shown in (40). She notes that the incipient uses of *so* indicate the status of the upcoming action is not contingent on the immediately preceding talk, but rather motivated by the "pending" status of the activity launched, which the speaker constitutes as part of the common ground between the interlocutors.

- (40) *So, what do you think of the school?* (Longman Dictionary of American English, 1983: 766) (Bolden 2009: 974)

In the functional analyses of *speaking of X* and *huashuo* as discourse markers, i.e. in Section 3.3 and Section 4.3, the distinction between discursive pretext and pragmatic pretext is highly relevant and useful in distinguishing their different topic-introducing functions and linking mechanisms, which helps to explain how they facilitate the establishment of topical relevance and discourse coherence.

2.4.4 Periphery and functional asymmetry

The term *periphery* can be understood in different ways. In this thesis, it refers to the margin of an argument structure in a sentence, i.e. positions outside the propositional core of the matrix clause, following Traugott (2015b). The **left periphery (LP)** and the **right periphery (RP)** refer to the positions preceding and following an argument structure, while the **medial periphery (MP)** refers to the contingent position for parenthetical elements in the middle of an argument structure, as exemplified in (41) to (43) respectively.

(41) *Seriously, that man is a poet.*

(42) *So that's not gonna work. What do you suggest we do then?*

(43) *Would say 100 bucks be enough?*

In recent decades there has been increasing interest in comparing the pragmatic functions and discursive behavior in different peripheral positions. Despite the sometimes different definitions of the term *periphery*, it has been noted in numerous studies that the (meta)linguistic functions of elements appearing in the left periphery (LP) and the right periphery (RP) are asymmetrical. For instance, Ferrara (1997) shows that *anyway* functions as an additive or dismissive adverbial in clause-final position but as a resumptive discourse marker in clause-initial position, which signals a resumption of the train of thought of the speaker. Haselow (2011) distinguishes an utterance-linking *then* in utterance-initial position from a modal particle *then* in utterance-final position, which links the utterance to a

non-verbalized pragmatic pretext. Andersen (2015) observes that the English and Norwegian discourse markers *regarding* (*angående*) and *when it comes to* (*når det kommer til*) serve topic transitional functions in the LP position but proposition specificational functions in the RP position, specifying the scope or respect of the proposition. Chen (2018) identifies markedly different pragmatic functions of *bushi* ‘not true’ in Chinese in the left and right peripheries, i.e. negative evaluation, other-correction and self-clarification in the LP position and stance marking, confirmation seeking, downtoning and hearer involvement in the RP position.

In an attempt to establish cross-linguistic generalizations, Beeching and Detges (2014) develop the hypothesis that the LP is the expected locus for subjective functions, while the RP is typically associated with intersubjective functions. From a historical perspective, this correlation thus leads to the expectation that expressions recruited to the LP undergo subjectification, while those recruited to the RP undergo intersubjectification. As intersubjectification is considered a further stage of subjectification (Traugott 1982; 2003; Traugott & Dasher 2002), what this hypothesis means for the development of discourse markers is practically a left-to-right movement: when discourse markers develop intersubjective meanings out of subjective meanings, they are assumed to exhibit a change of peripheral preference from the LP to the RP.

The functional asymmetry hypothesis has been supported by a number of studies, such as Aijmer (2002: 257–259) on *actually*, Izutsu and Izutsu (2013) on Japanese *sosite* ‘and’, *hoide* ‘and’ and *sikasi*

‘but’, and Degand (2014) on French *alors* ‘then’ and *donc* ‘so’. Likewise, the left-to-right movement hypothesis has been supported by various studies, such as Degand and Fagard (2011) on French *alors* ‘then’, Sohn and Kim (2014) on Korean *kuntey* ‘but’, and Pons Bordería (2018: 350–354) on Spanish *o sea* ‘or (it might) be’. But at the same time, there are also numerous counterexamples that challenge the hypothesized correlation between the peripheral positions and the (inter)subjective functions and thereby question the validity of the left-to-right movement hypothesis, such as Traugott (2012) on *no doubt* and *surely*, Onodera (2014) on Japanese *dakara* ‘so’ and *dakedo* ‘although’, Tanno (2017) on Japanese *toiuka* ‘how to say’, and Chen (2018) on Chinese *bushi* ‘not true’. The studies of *speaking of X* and *huashuo* in this thesis will thus contribute to the vibrant discussion and further test the validity of these hypotheses.

2.4.5 Dialogual and monological contexts

The distinction between dialogual and monological contexts is made in regard to the number of participants involved in the discourse. That means, speech acts that are produced consecutively by one speaker are defined as **monological**, while those realized by different speakers with turn-taking operations are considered **dialogual** (see also Detges & Waltereit 2011; Roulet 1984; Sacks et al. 1974; Schwenter 2000).

Despite the similarity in the terminology, the *dialogual-monological* distinction should not be confused with the *dialogic-monologic* distinction, which is identified by the number of perspectives contained in the utterances. That means, speech acts that

contain heterogeneous orientation and multiple perspectives are defined as *dialogic*, while those that are closely aligned towards some argumentative conclusion are understood as *monologic* (see also Nøke 2006; Roulet 1984; Schwenter 2000).

In the following examples, sentence (44) represents a monologal utterance with dialogicity, as it involves only one speaker but entails more than one viewpoint – the speaker’s viewpoint “the rain did not affect whether the baseball game was played or not” and a general/default viewpoint “the rain would have been a sufficient cause for the baseball game not to be played”; example (45), on the other hand, is a monologic dialogue, as both speakers are contributing to the same viewpoint.

(44) *Even though it rained, they played the baseball game.*
(Schwenter 2000: 257)

(45) A: *John sure is smart.*
B: *He certainly is!* (Schwenter 2000: 278)

As *speaking of X* and *huashuo* are typically used at the locus of topical movement, they are closely associated with dialogic contexts. Therefore, the distinction between dialogic and monologic contexts would not be beneficial and it will not be part of the discussion in this thesis. Instead, the distinction between dialogal and monologal contexts will be useful in examining the usage patterns in turn-taking operations (Section 3.3.6 and Section 4.3.4).

Chapter 3

The *speaking of X* construction

This chapter aims to provide a comprehensive analysis of the SPOX construction. It starts with a brief introduction of the corpus data supporting the empirical study (Section 3.1). Then in Section 3.2, the formal-semantic properties of SPOX and the frequency representation of the different uses of SPOX in the dataset will be presented. Section 3.3 is dedicated to the empirical findings concerning the functional aspects of SPOX as a discourse marker, including the different discourse marking functions, peripheral behavior, and monologual-dialogual uses. After that, Section 3.4 takes on a diachronic perspective. It not only addresses the recent changes of SPOX in terms of its functionalities and usage patterns as a discourse marker, but also traces its development path in becoming a discourse marker and discusses the motivational factors. Section 3.5 offers a brief summary of the findings in this chapter.

3.1 Data and methods

3.1.1 Corpora and datasets

The data sources for the SPOX study are two corpora of American English. The first one is the **Corpus of Historical American English**¹² (the COHA corpus). It contains 400 million words in total

¹² <https://www.english-corpora.org/coha/> (Accessed: 18 March 2022)

and 1-30 million words for each decade from 1810 to 2009. It contains four different text genres: *fiction*, *magazine*, *newspaper*, and *non-fiction*, which are balanced in each decade. Being the largest structured corpus of historical English to-date, it informs the diachronic analyses of the SPOX construction, i.e. it facilitates the study of the development path of *speaking of X* in becoming a discourse marker.

The second data source is the **Corpus of Contemporary American English**¹³ (the **COCA** corpus), which contains more than 1 billion words of text in total and 25 million words each year from 1990 onwards. It is updated on a biannual basis¹⁴. The cut-off point of data collection for this thesis is December 2015. It is a genre-balanced corpus with five different types of data: *spoken*, *fiction*, *popular magazines*, *newspapers*, and *academic texts*¹⁵.

The data collected are restricted to *spoken* data only, i.e. from the subset **COCA-SPOK**, which contains 127 million words of transcripts of conversations from more than 150 different TV and radio programs. This type of spoken data has both strengths and limitations. On the one hand, the semi-spontaneous and semi-structured format and the semi-natural and semi-performative

¹³ <https://www.english-corpora.org/coca/> (Accessed: 18 March 2022)

¹⁴ The most recent text addition was made in December 2019 besides an exceptional one in March 2020 (which did not extend the time span of the corpus beyond December 2019).

¹⁵ The March 2020 update added *blogs*, *other web pages*, and *TV and movie subtitles* into the mix, amounting to a total of eight different genres.

nature of media language underlie potential bias in the functional analyses of the SPOX construction: certain less typical discourse marking functions and usage patterns might be genre specific (the relevant discussion follows in Section 3.3). On the other hand, TV and radio programs are bound by a preset structure and agenda, which means that the hosts always have a list of topics or questions to cover in the program or the interviews. As the SPOX construction is strongly associated with topical movement in discourse, its usage is particularly well represented in this discourse genre. On top of that, except for a small number of "formulaic/scripted" utterances, usually related to the progression of the show, such as "coming up next" and "after the (commercial) break", about 95% of the data are unscripted natural speeches¹⁶. The transcripts represent the actual conversations well, i.e. including interruptions, false starts, and so on. Generally speaking, the COCA-SPOK data are a reliable and fruitful source for the study of SPOX as a discourse marker.

My initial datasets were obtained through a KWIC search of the text string "speaking of" in both corpora powered by their respective websites. The search returned 5715 tokens from COHA and 2405 tokens from COCA-SPOK. Each token is a cluster of texts with a length of 20–30 words, containing the keyword "speaking of" in the middle position and an approximately 10-word context both before

¹⁶ See "Notes on the naturalness and authenticity of the language from these transcripts" at <https://www.english-corpora.org/coca/help/spoken.asp> (Accessed: 14 January 2022).

3 The speaking of X construction

and after the keyword. Metalinguistic information such as *year*, *genre* and *source* was also provided alongside the tokens.

In the next step, invalid tokens due to transcription errors or special uses (e.g. book title *Speaking of Operations*), duplicate tokens, and cases of accidental juxtaposition of “speaking” and “of” (e.g. “*I’m speaking of course about...*”) were excluded from the datasets. Repetitive occurrences of the same SPOX construction due to various reasons (e.g. interruptions by the hearer, self-corrections) were counted only once.

After the screening process, my final datasets consist of 5669 tokens from COHA and 2351 tokens from COCA-SPOK. The distribution of the data in the datasets by decade/year is shown in Table 1 and Table 2 below.

Table 1 Dataset COHA composition (number of tokens per decade)

Decade	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s
Tokens	19	158	301	338	376	393	376	447
Decade	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s
Tokens	439	342	320	342	272	209	220	217
Decade	1970s	1980s	1990s	2000s				
Tokens	197	221	247	235				
Sum	5669							

Table 2 Dataset COCA-SPOK composition (number of tokens per year)

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998
Tokens	53	57	54	51	65	70	73	66	96
Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
Tokens	91	84	48	72	70	161	126	123	96
Year	2008	2009	2010	2011	2012	2013	2014	2015	
Tokens	76	105	107	104	122	116	114	151	
Sum	2351								

The COCA data serve as the foundation for the synchronic functional analyses of the SPOX construction and at the same time provide insight into the functional changes and further development of SPOX as a discourse marker. The COHA data assist the diachronic analyses of the constructionalization process of SPOX, i.e. its development from a lexical item towards a discourse marker.

Due to the different sizes of the corpora in different time periods, the diachronic frequency distributional analyses will be done and presented with normalized and/or relative token frequency. The examples presented in this study mainly come from these two datasets and marginally from other media (e.g. TV shows).

3.1.2 Classification and annotation

The “speaking of” strings in my datasets are used in various ways. They are found in different grammatical structures, ranging from syntactically integrated verbal predicates to subordinate adverbial clauses and to procedural discourse markers. A functional classification is thus essential in the study of the SPOX construction.

3 The speaking of X construction

In Tables 3, 4 and 5 below, different uses of SPOX are sorted in terms of the word class and the functional role in the sentence. Each type of usage is illustrated with at least one example. The raw token frequencies (*f*) of each type of usage in the COHA and COCA datasets are also provided.

Table 3 Gerundial uses of SPOX

Type	Example		
Subject	(46) <i>And it shows you that, you know, <u>speaking of</u> the government is often not a helpful thing to do [...]</i> (2013, SPOK, NPR, COCA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		120	5
Subject complement	(47) <i>I think that the way that he can state his case is <u>speaking of</u> the millions of Soviets which have been killed in the World Wars.</i> (1990, SPOK, PBS_Newshour, COCA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		7	3
Verbal complement (Object)	(48) <i>But what we need to do is get to that situation so we can stop <u>speaking of</u> television news as an ideal, which I don't think ever exists.</i> (1994, SPOK, ABC_Nightline, COCA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		168	11

Prepositional complement	(49) [...] but he was convicted at a court martial for treating President Johnson contemptuously or <u>speaking of him</u> contemptuously. (2008, SPOK, NPR_TalkNation, COCA)		
	(50) You're not going to be able to dismiss Hanssen by <u>speaking of his</u> psychiatric categories. (2001, SPOK, CBS_Sixty, COCA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		588	10

Table 4 Participial uses of SPOX

Type	Example	
Verbal predicate	(51) Obviously, everybody today is <u>speaking of</u> the United Nations vowing to continue its mission and its role in Iraq. (2003, SPOK, NPR_TalkNation, COCA)	
		<i>f</i> (COHA)
		1314
Adjectival complement	(52) [...] her work will not be worth <u>speaking of</u> . (1837, FIC, LiveLetLiveDomestic, COHA)	
		<i>f</i> (COHA)
		73
		<i>f</i> (COCA)
		0

3 The speaking of X construction

Adnominal modifier	(53) <i>The voices of free men <u>speaking of</u> the fight for freedom in their many tongues were to be heard no more from the traditional capital of liberty.</i> (1992, SPOK, CBS_Special, COCA)		
	(54) <i>To me, the message was one of infinite pathos and rebuke, <u>speaking of</u> a heroism beyond my poor conception,[...] (1904, FIC, CapeCodFolks, COHA)</i>		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		95	10
Adverbial modifier ¹⁷	(55) <i>I think “zip it” is a good metaphor when <u>speaking of</u> Bill Clinton.</i> (2006, SPOK, Fox_Saturday, COCA)		
	(56) <i><u>Speaking of</u> additions like a pool and a home theater, Young confessed, “We lost our sense of perspective and the house got more and more extravagant.”</i> (2012, SPOK, CBS_ThisMorning, COCA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		2404	25

¹⁷ The adnominal and adverbial modifier uses of SPOX are distinguished solely on the semantic level. The adnominal modifier SPOX provides descriptive or identificational information about an antecedent, which may or may not be the subject of the matrix clause; the adverbial modifier SPOX provides situational information such as time, cause, manner, etc. and modifies how the event in the matrix clause, as depicted by the subject and the predicate verb, is carried out.

Absolute construction	(57) <i>Somebody <u>speaking of</u> having oysters for supper, much surprise was excited by Mrs. Inchbald's saying that she had never eaten one.</i> (1876, MAG, Atlantic, COHA)		
		<i>f</i> (COHA)	<i>f</i> (COCA)
		13	0

Table 5 Discourse marker uses of SPOX

Type	Example	
Discourse marker	(58) <i>The lady and I are here to get the bad guys off the streets. Make the world a safer place for you kids. And <u>speaking of</u> safety, you on the bike, why aren't you wearing a helmet?</i> (2007, FIC, FantasySciFi, COHA)	
	(59) <i><u>Speaking of</u> the devil -- here is the old boy himself.</i> (1909, FIC, Play:MansWorld, COHA)	
	(60) <i>In Baltimore, Philadelphia, and New York, I met with some exceptions to this; but <u>speaking of</u> the country generally, it is unquestionably true.</i> (1832, NF, DomesticManners, COHA)	
	(61) <i>60 years of Gourmet magazine, <u>speaking of</u> bookilicious.</i> (2002, SPOK, NBC_Today, COCA)	
	<i>f</i> (COHA)	<i>f</i> (COCA)
	887	2144

In both datasets, the tokens are annotated with “functional roles (subject, adverbial, discourse marker, etc.)”, “variant (*speaking of X*, *speaking of which*, *speaking of*, etc.)”, and “position (sentence-initial/-medial/-final)”. A number of parameters that are relevant to specific synchronic functional analyses and diachronic analyses are added to the COCA and COHA datasets separately. In the COCA dataset, the additional parameters include “X-element (the lexical item occupying the X-slot)”, “X-class (noun phrase, adjective, etc.)”, “discourse marking function (topic introduction, speech act adverbial, etc.)”, “discourse event (topic change, expansion, initiation)”, and “speech context (dialogual, monologual)”. In the COHA dataset, the additional parameters center on the adverbial use of SPOX, which I consider have motivated the constructionalization of the discourse marker SPOX (detailed discussion follows in Section 3.4.3). These parameters include “subordination (syndetic, asyndetic)”, “subordinator (*when*, *by*, *after*, etc.)”, “predicate verb (of the matrix clause)”, and “predicate verb type (speech-act, mental, action)”. During the annotation process, unclear cases due to limited context size were searched again in the corpora in order to be clarified through expanded contexts.

The classification and annotation processes were assisted with automated methods using the software R. Specific text-processing commands were written and implemented using the R language and environment to get the rough structure. The final results of data classification and annotation, as well as further functional and diachronic analyses of the data, were completed manually by examining every token.

3.2 Formal and semantic properties

3.2.1 Semantic content

The SPOX construction has the formal structure of an *ing*-form participle clause. This form, also called a present participle clause or a progressive nonfinite clause, refers to a subjectless subordinate clause which solely consists of a nonfinite verb phrase in present participle form. This type of clause is well-known for its functional versatility, as it can assume different syntactic and semantic roles in a sentence, such as adnominal modifier, adverbial modifier, and dangling participle. In this section, I will show that the discourse marker usage of the SPOX construction is fundamentally different from the other clausal uses, and that its semantic content cannot be predicted from the lexical meanings of its components or through analogy to the other clausal uses.

First of all, a discourse marker SPOX cannot be understood as an adnominal modifier SPOX. As shown in (62) and (63), the SPOX clauses appear to be both dependent clauses that are separated from the matrix clause by comma punctuation, and they both directly follow a noun phrase in terms of the syntactic position.

- (62) *Roberta disliked this new weatherman, whose delivery, speaking of the flood, tilted alternately toward the jaunty and the grave.*
(2002, FIC, SouthernRev, COHA)

- (63) TERRY-GARCIA# [...] *And we're embarking upon an exciting new era in exploration, and a lot of that is due to technology, [...]*
NEAL-CONAN# *Bob, you told me once a story, speaking of technology, of going down in Alvin with a couple of scientists, [...]*
(2013, SPOK, NPR, COCA)

However, it is clear that these SPOX clauses have very different semantic content and syntactic scopes. The SPOX clause in (62) can be understood as a reduced non-restrictive relative clause, which modifies the antecedent noun “delivery” by describing its content, i.e. “whose delivery, which was speaking of the flood”. It has a narrow scope over the element that it modifies. By contrast, the SPOX clause in (63) does not specify that the content of the “story” is “technology”. A narrow-scoped relative clause interpretation “a story, which is speaking of technology” is not pertinent. Instead, it has a scope over the entire proposition “you told me once a story of going down in Alvin with a couple of scientists” and carries the semantic content “the reason for me to bring up this story is that we just spoke of technology”.

Secondly, a discourse marker SPOX should not be confused with an adverbial SPOX. As shown in (64) and (65), the SPOX clauses are not only identical from a purely syntactic perspective, but they also both have a scope over the entire matrix clause to which they are attached.

- (64) Speaking of her sister-in-law, Valeria used to say, “Just a bit theatrical looking [...] (1971, FIC, Other, COHA)

- (65) *Speaking of Missus Mora, she said you didn't come to the Catholic ladies meeting yesterday.* (2007, FIC, FantasySciFi, COHA)

But evidently, these SPOX clauses have very different semantic meanings and relations to the matrix clauses. As an adverbial modifier, an *ing*-form participle clause without an overt subject is supposed to have the same subject as the matrix clause, which was defined as the “attachment rule” by Quirk et al. (1985: 1120–1121) and the “equi-subject constraint” by Givón (1993: 165–168) (see also König & Van der Auwera 1990: 346). In (64), the SPOX clause is indeed controlled by the subject of the matrix clause: “Valeria” is the subject of both the predicate verb “say” in the matrix clause and the verb phrase “speak of” in the SPOX clause. In (65), however, the SPOX clause has a different subject from the matrix clause. The subject of the matrix clause “she” refers to “Missus Mora”, while the subject of “speaking of Missus Mora” is an implicit “I/we” that refers to the interlocutor/s.

In terms of the semantic meaning, an adverbial clause is itself a conceptual part of the propositional content of the whole sentence. The SPOX clause in (64) modifies the matrix clause and expresses a concomitant relationship between the events depicted in these clauses. The discourse marker SPOX in (65), on the other hand, is not a conceptual part of the propositional content of the sentence but rather a procedural element, i.e. it provides metalinguistic information about the speech act – “it is motivated by them previously speaking of Missus Mora”, instead of affecting the truth condition of the proposition in the matrix clause.

Thirdly, a discourse marker SPOX should be distinguished from a dangling participle SPOX. Interestingly enough, just like the discourse marker SPOX, a dangling participle SPOX also has a different subject from the matrix clause and it is semantically neither an adnominal nor an adverbial modifier for the matrix clause: As shown in (66), “reading the evening paper” is neither an attribute of the noun “dog”, nor does it modify how the event “a dog started barking” is carried out.

- (66) *Then he fetched some newspaper from the kitchen table, went into the study, and settled down in his favorite armchair, looking forward to a quiet and undisturbed evening. Reading the evening paper, a dog started barking.* (Kortmann 1991: 46)

According to Hayase (2011), the relationship between a dangling participle and its matrix clause is one between “ground” and “figure event”: i.e. the dangling participle describes an unbounded (atemporal) background situation and the matrix clause describes a bounded (temporal) situation of cognition or perception. In example (66), “reading the evening paper” serves as the background scenario for the event “a dog started barking” to be perceived.

Since there are no empirical data of the dangling participle usage of SPOX in COHA or COCA-SPOK, a constructed example is given in (67). A comparison between (67) and (68) clearly shows that, despite formal similarities, a discourse marker SPOX and a dangling participle SPOX have very different semantic content.

- (67) *So one night she invited everybody over, gathered them in the living room, and started telling stories from her trip, from befriending a local chef to learning how to dive and to getting lost in the woods. Speaking of one supernatural experience, the TV in the far corner turned itself on.*
- (68) *And speaking of diet busting foods, Taco Bell is getting into the breakfast business. (2012, SPOK, CBS_NewsMorn, COCA)*

Without an overt subject in the structure and not sharing the subject with the matrix clause, both SPOX clauses in (67) and (68) require the identification of an implicit controller subject in order to construe the intended meaning. But the discourse marker SPOX and the dangling participle SPOX are controlled by subjects of different sources. In (67), the subject of “speaking of one supernatural experience” is “she”, which can be recovered from the context; in (68), however, the subject of “speaking of diet busting foods” is the interlocutor(s), which only exists in the extra-linguistic context.

Furthermore, even though a dangling participle clause does not modify the matrix clause, they belong to the same textual plane: i.e. the events “[she was] speaking of one supernatural experience” and “the TV in the far corner turned itself on” are both propositional components of the same narrative. The discourse marker SPOX, however, functions on the illocutionary plane instead. It refers to the speech act situation and establishes a metalinguistic ground for the production of the matrix clause: i.e. the idea and the utterance of “Taco Bell is getting into the breakfast business” are based on a previous speech act, which is “speaking of diet busting foods”.

To sum up, the meaning of a discourse marker SPOX cannot be construed with an isolated understanding of the lexical meanings of its components or through analogies to formally similar usage. It needs to be recognized as a symbolic unit with a conventionalized form-function pairing. Its discourse marking functions and usage patterns will be discussed in detail in Section 3.3.

3.2.2 Formal fixedness and variation

The SPOX construction has a semi-fixed form. It typically consists of a fixed head “speaking of” and an open X-slot. Because of the variety of things that can be spoken of, the elements that can fit in the X-slot exhibit a high degree of diversity.

According to the COCA-SPOK dataset, out of 2134 valid tokens of the discourse marker SPOX¹⁸, there are 1709 different types, as shown in Figure 1. On the upper and lower X-axes, the X-types are ranked according to their token frequencies and sorted in alphabetical order within the same rank group. Due to the huge number of the different X-types and hence the length of the data points, the X-axes are only partly shown. The break between rank 20 and rank 1701 suppresses parts of the X-types that appear less than 5 times throughout the dataset.

¹⁸ The invalid tokens refer to incomplete uses of SPOX due to interruptions, bleeps, false starts, etc.

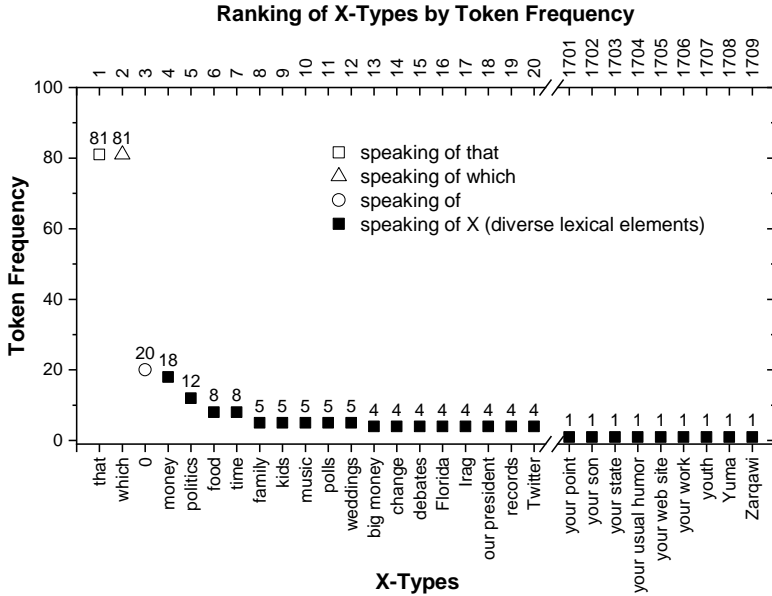


Figure 1 Type frequency distribution of discourse marker SPOX in the COCA-SPOK dataset (sum = 2134 excluding invalid tokens) (the X-axis has a break between rank 20 and rank 1701 for better visual presentation)

In the top-left corner of Figure 1, we can see two types of SPOX, i.e. *speaking of that* and *speaking of which*, which are significantly more frequent than the other types. It shows that, despite the high productivity and diversity of SPOX as a discourse marker, there are certain variants that are more formulaic and conventionalized. Ranking third in the array is *speaking of*, in which only the head of the construction is expressed and the X-slot is left unfilled. It represents an even more formulaic usage of the expression. Its lower frequency in comparison to *speaking of that* and *speaking of which* could be

attributed to the fact that it is still a relatively new form. The first occurrence of this usage is found towards the end of the 1990s in the dataset. More discussion about the diachronic development of the SPOX construction will be given in Section 3.4.

Based on these findings, I distinguish the following formal variants of discourse marker SPOX, besides the standard *speaking of X* form: i.e. *speaking of that* (SPOT), *speaking of which* (SPOW), and *speaking of* (SPO), as exemplified from (69) to (71).

- (69) KATHIE-LEE-GIFFORD# [...] *The country music fans are the most loyal. You can have the longest, longest career in country music.*
DARIUS-RUCKER# *Absolutely.*
KATHIE-LEE-GIFFORD# *Speaking of that, we have to ask you about our friend Randy Travis. How is he doing? Have you heard?*
(2013, SPOK, NBC, COCA)
- (70) DAVID-EDELSTEIN: *But everyone wants the racially-conscious, The Help, to get something. Speaking of which, the only contest is best actress.* (2012, SPOK, CBS_SunMorn, COCA)
- (71) JOHN-LEGEND# *They get to come hang out with us at my show in Los Angeles.*
SAVANNAH-GUTHRIE# *Speaking of, you're on tour right now. Did I hear a couple got engaged the other day at your show?*
(2014, SPOK, NBC, COCA)

Besides the formal invariability of the head “speaking of” across the variants, the fixedness of the SPOX construction also refers to the fact that it has a formulaic nature and does not always follow standard grammatical rules or the categorial behavior of its

components. In principle, with the preposition “of” in the structure, the X-slot is supposed to accommodate a nominal item, i.e. a noun phrase, a gerund or a pronoun, as exemplified from (72) to (74). But in reality, my data show that the X-slot can just as well be filled with non-nominal items, such as an adjective, a verb phrase, a prepositional phrase, or even a whole sentence, as shown from (75) to (78).

- (72) *And speaking of money, John Kerry is spending some of his own money on his current vacation in Idaho.*
(2004, SPOK, CNN_Politics, COCA)
- (73) *Speaking of disapproving, art is so subjective as we all know.*
(2005, SPOK, PBS_Tavis, COCA)
- (74) *Oh, hey. Speaking of this. Have you heard the -- the story floating around -- the gossip about why Oprah's book has been cancelled?*
(1993, SPOK, Ind_Limbaugh, COCA)
- (75) *Speaking of bloody, Iraq seems to be worse than it has been in years.*
(2013, SPOK, CBS, COCA)
- (76) *KING: Tomorrow night on LARRY KING LIVE, Heather Mills will be our special guest, the wife of Paul McCartney has written an extraordinary tell-all, and she will tell it all tomorrow night. Speaking of tell it all, the man who covers it all, and does it so superbly, night after night after night, the host of NEWSNIGHT in New York, the one and only Aaron Brown. (2002, SPOK, CNN_King, COCA)*
- (77) *Well, speaking of on budget, I do want to talk in our couple minutes left here about the cost issue.* (2012, SPOK, PBS_NewsHour, COCA)
- (78) *OLBERMANN: [...] Speaking of anything goes, there is Paris Hilton.*
(2006, SPOK, MSNBC_Olbermann, COCA)

These examples show that the SPOX construction presumes a nominal treatment of the lexical elements that occupy the X-slot, regardless of their actual word classes. The uncanonical syntactic property of the construction reflects its procedural nature as a formally conventionalized unit.

3.3 Functional aspects

As a common trait of discourse markers, SPOX appears to be polyfunctional. My data show that it can be used as a topic-introducing device (79), a speech act adverbial (80), a commenting device (81), and a variant of the idiomatic expression “speak of the devil” (82). The frequency distribution of these functions in the COCA-SPOK and COHA datasets are shown in Figure 2 and 3 respectively.

- (79) *The lady and I are here to get the bad guys off the streets. Make the world a safer place for you kids. And speaking of safety, you on the bike, why aren't you wearing a helmet?*
(2007, FIC, FantasySciFi, COHA)
- (80) *In Baltimore, Philadelphia, and New York, I met with some exceptions to this; but speaking of the country generally, it is unquestionably true.*
(1832, NF, DomesticManners, COHA)
- (81) *60 years of Gourmet magazine, speaking of bookilicious.*
(2002, SPOK, NBC_Today, COCA)
- (82) *Speaking of the devil -- here is the old boy himself.*
(1909, FIC, Play:MansWorld, COHA)

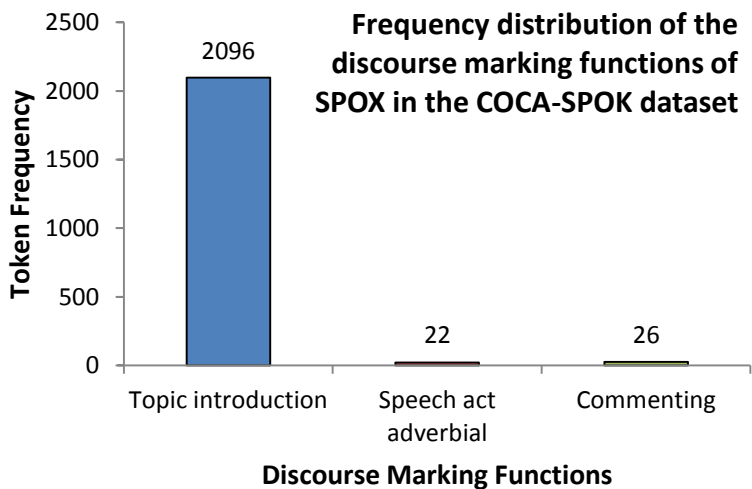


Figure 2 Frequency distribution of the discourse marking functions of SPOX in the COCA-SPOK dataset (sum = 2144)

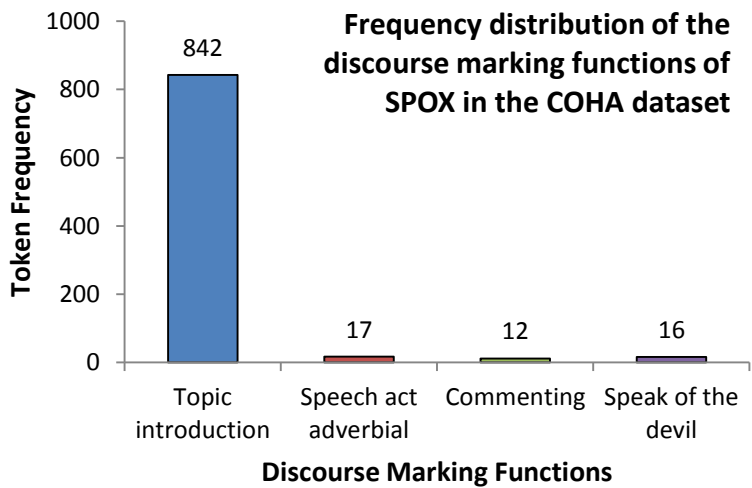


Figure 3 Frequency distribution of the discourse marking functions of SPOX in the COHA dataset (sum = 887)

It is clear that the topic introduction function is the most typical and conventional for SPOX as a discourse marker (97.76% in COCA-SPOK and 94.93% in COHA). The use of SPOX to modify the speech act or to make a comment is possible but much rarer. The sporadic usage of SPOX as a variant of the “speak of the devil” expression is only observed in the COHA corpus and therefore will only be discussed as part of the functional analyses of SPOX, but it will not be included in the functional analyses of the usage patterns or the frequency distributional analyses alongside the other discourse marking functions of SPOX, which are based on the COCA-SPOK dataset.

In the following parts of this section, the different discourse marking functions of SPOX will be introduced and distinguished in Section 3.3.1 to Section 3.3.4. The usage patterns of SPOX, including peripheral positions and dialogual-monologal uses will be analyzed in Section 3.3.5 and Section 3.3.6 respectively.

3.3.1 Topic-introducing functions

As a topic-introducing device, SPOX marks discourse events that feature different types and/or degrees of topical movement in relation to the global topic structure, including *topic change*, *topic expansion* and *topic initiation* (see Section 2.4.2 for the definition of the terms).

3.3.1.1 Topic change and topic expansion

Typically, SPOX marks discourse events that contain some kind of topical movement and prepares the hearer for the potentially

unexpected change of topicality. It can mark both major and minor topical movements, i.e. *topic change* and *topic expansion* events.

To illustrate, (83) and (84) represent examples of topic change events that are facilitated by the use of SPOX. In (83), by using “speaking of hockey”, Paula Zahn turns to talk about herself in a conversation that was originally about the U.S. team in the Winter Olympics. In (84), with the use of “speaking of which”, Lee Cowan shifts the topic of his conversation with Bob Cousy from Bob’s basketball career to the ball-playing hobby of the President. In both cases, the SPOX-introduced utterances are no longer addressing the original topics in the previous discourse. The transitions represent rather striking changes of topicality.

- (83) SMITH: *Yeah. And that hockey team he talked about just a second ago skated to a 2-2 tie with Sweden, which is a great, great, great feat going into these Games.*

ZAHN: *Hey, speaking of hockey, do you know that I worked out with the Austrian team this morning?*

(1994, SPOK, CBS_Morning, COCA)

- (84) BOB-COUSY# *It led to all sorts of things. I’ve been to every place imaginable. I’ve been invited to the White House by six sitting Presidents, because of my relationship with the ball.*

LEE-COWAN# *Speaking of which, even the current sitting President still plays with a ball; it relieves the way to the office, he says. ...*

(2012, SPOK, CBS, COCA)

In comparison, in (85) and (86), the topics are expanded rather than changed with the use of SPOX. In (85), following Bo Bice's short comment on Simon's smiling, Karyn Bryant uses "speaking of Simon" to zoom in on Bo Bice's interaction with the judges. The overall topic about Bo Bice's experience on American Idol stays unchanged. In (86), with the use of "speaking of that", Katie Couric does not show her interest in changing the topic about the tornado attack, but rather in continuing the topic with new aspects that contribute to the topic construction as a whole.

(85) BICE: *Does it get any better? I keep saying that. Every day single day, I say, it can not, can not, get any better than this, and then each day it continues to. So, I'm blessed.*

BRYANT: *Yes, we've got a shot of you right there with Skynyrd. I mean, that -- the smile, the joy, that says it all. Look at that*

BICE: *Even Simon is smiling, yes*

BRYANT: *So, speaking of Simon, what's the best advice the judge gave you ...*

(2005, SPOK, CNN_Showbiz, COCA)

(86) Reporter: *We got used to these things touching down in rural areas, ranches, or farmsteads in the Midwest, but this one, this series of storms was different. It seemed to sort of hug all the inner states between Mississippi and Georgia, and hitting all those populated areas in between. And by doing that, it has two results. One is, that the damage is much more significant, and secondly, that the recovery will be that much longer. Katie.*

COURIC: *And, speaking of that, Dean, will the search for survivors*

continue through the night?

(2011, SPOK, CBS_NewsEve, COCA)

In these topic change and topic expansion operations, the use of SPOX establishes a link between the different topics and subtopics. The link can be explicit or implicit, on both textual and conceptual levels, depending on the type of the X-element in SPOX. In (83) and (85), for instance, the X-slots are filled with the concrete lexical items “hockey” and “Simon” extracted from the discursive pretext, which are conceptually related to the “Austrian team” and “judges” in the new utterances. The links are thus made explicit both on the textual level and the conceptual level. By contrast, in (84) and (86), the X-elements are not concrete textual elements from the discursive pretext but semantically unspecific “which” and “that”, which point at certain concepts or propositions contained in the discursive pretext. The suggested links are therefore implicit both on the textual level and the conceptual level. Their exact references are subject to the hearer’s own interpretation. The same implicit linking mechanism can also be observed in cases where the *speaking of* variant of SPOX is used, i.e. when there is no X-element of any kind given in the structure.

Sometimes, the textual and conceptual links put forward by SPOX represent a manipulation of word meaning. For example, in (87), the use of “speaking of behind” to link up “let’s support her” and “her butt food diet” takes advantage of the double entendre of “behind”. While SPOX does point at the lexical item “behind” in the discursive pretext, the meaning of “behind” that is recruited to create the conceptual link is different from the original meaning.

(87) RIVERA: *Well, let's all get behind her and support her in her efforts, ladies and gentlemen.*

Mr-WALKER: *Well -- and speaking -- and speaking of behind, this is the most incredible one of all -- to me, anyway. Heather is now -- you know, she's got a terrific slim figure. She's eating a lot of junk food: burritos, hamburgers, french fries. She calls it butt food because she feels the one area where she is deficient is her rear end, and she believes -- she believes that this kind of fat goes to her rear end.*

(1994, SPOK, Ind_Geraldo, COCA)

Overall, no matter which exact kind of link is to be created, SPOX signals to the hearer that the upcoming utterance is related to the preceding discourse in some way. The X-element, when present, serves as a signpost to help the hearer find the conceptual connection and make sense of the transition, which in turn facilitates the perception of discourse coherence.

3.3.1.2 Topic initiation

Besides marking topic change and topic expansion events, SPOX is also used for marking *topic initiation* events, i.e. the introduction of a brand-new topic into the discourse without any previous topic to depart from, as shown in (88) and (89).

In (88), at the beginning of a new segment after the commercial break, Soledad O'Brien uses SPOX to start her next topic about flying, which is not related to anything from her earlier conversation with Karen Tumulty. In (89), Rush Limbaugh uses SPOX to start his very first topic of the show about the "Epidemic Awards", which is not

linguistically encoded or conceptually evoked anywhere in his previous utterances.

- (88) TUMULTY: [...] *But the concerns about immigration, as you said, are on two levels, one on security and one on the incredible financial burden that these immigrants are putting on a lot of cities and states across the country.*

S-O'BRIEN: *Karen Tumulty of "Time" magazine. Karen, thank you very much.*

TUMULTY: *Thank you, Soledad.*

S-O'BRIEN: *My pleasure.*

(COMMERCIAL-BREAK)

Speaking of flying, the summer travel season is just a couple months away, and a lot of folks are finding out it's harder than ever to use those frequent flyer miles. [...]

(2006, SPOK, CNN_AM, COCA)

- (89) Announcer: *Ladies and gentlemen, Mr. Rush Limbaugh.*

LIMBAUGH: *Hey, hey, hey. Hey! Thank you. Thank you very much. Thank you. Thank you. Thank you. Thank you so much. Thank you. Thank you. That's -- thank you. Thank you.*

Members of Audience: (In unison) *Whoo! Whoo!*

LIMBAUGH: *Whoo! Whoo! All right. All right. OK. All right. Looks like I have to stop the applause by sitting down, otherwise this crowd would go on all night. You can see it in their faces. Hey, speaking of the Epidemic Awards last night, it got off to the typical start, here comes a has-been director named Arthur Hiller, and within 10 seconds, here's a big plea for continued federal funding, National Endowment for the Arts.*

(1995, SPOK, Ind_Limbaugh, COCA)

In this usage, SPOX does not create any textual or conceptual links between the new utterance and the discursive pretext. Instead, it points at the existence of a pragmatic pretext, which is part of the host-audience relationship. Firstly, the unstated pragmatic pretext can be understood as the audience's awareness of the characteristics of the show and their readiness to accept the succession of different topics. SPOX is therefore used as a topicalization device that introduces the next item on a readily-made topic list for the show. Secondly, SPOX points at the existence of the pragmatic pretext that it is generally acknowledged that the upcoming topic is a popular and current one and therefore the new utterance should not be perceived as completely ungrounded.

This kind of pragmatic pretext, however, is constrained by the specific characteristics of the media language and the specific host-audience relationship. Therefore, the topic initiation function of SPOX supported by such pragmatic pretext should be understood as genre-specific at this point. It calls for further research with a different type of dataset in order to better understand the usage of SPOX in topic initiation events.

To summarize, the topic-introducing SPOX facilitates the perception of discourse coherence by linking the new utterance to pre-existing information, which is either linguistically encoded in the previous discourse or pragmatically established in the interlocutor relationship.

3.3.1.3 Frequency distribution

According to the results of the frequency distributional analyses of the topic-introducing SPOX in the COCA-SPOK dataset (Figure 4), it is most typically used in marking topic change and topic expansion events. And it appears to be equally functional in marking both major and minor topical movements. In comparison, the topic initiation uses of SPOX are by far the least common. This could be attributed to the genre-specific characteristics of this usage. It thus calls for further research to find out how this usage is represented, if at all, in other kinds of spoken data. The “unidentifiable” tokens refer to cases in which the use of SPOX is interrupted by another speaker, and therefore the matrix clause to which SPOX is supposed to be attached is missing from the utterance. In such cases, the exact type of the topic-introducing function of SPOX as intended by the speaker remains unclear.

The same frequency distribution of topic-introducing functions can be found across all the variants of SPOX. In Figure 5 we can see that, all of the variants of SPOX are used more often in marking (sub)topical movements and rather scarcely and selectively in marking topic initiation events. The “not applicable (N/A)” tokens refer to cases, in which the type of the variant cannot be definitely identified as the X-element remains unclear in the usage due to interruptions, bleeps, false starts, etc.

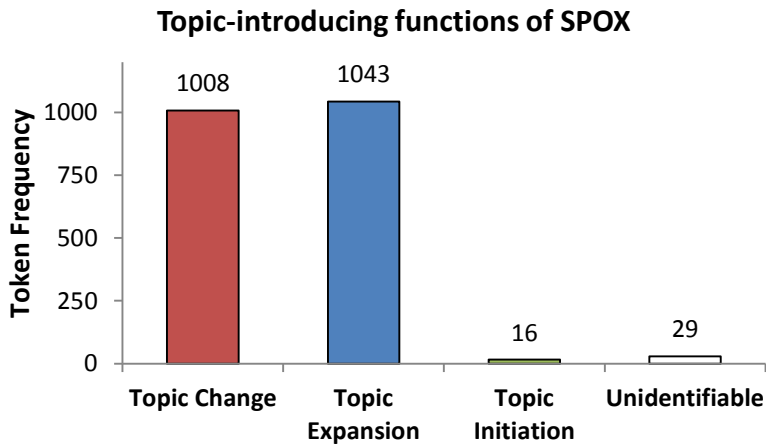


Figure 4 Frequency distribution of the topic-introducing functions of SPOX in the COCA-SPOK dataset (sum = 2096)

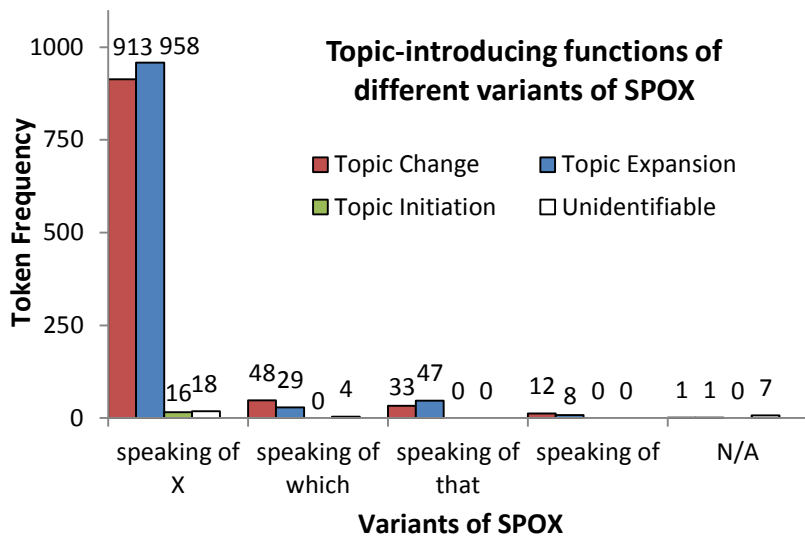


Figure 5 Frequency distribution of the topic-introducing functions of different variants of SPOX in the COCA-SPOK dataset (sum = 2096)

In Figure 5, only subtle differences can be seen in the use of the variants: *speaking of X* and *speaking of that* seem to be used more often in topic expansion events, while *speaking of which* and *speaking of* seems to favor topic change events. But these differences are too small to be statistically significant. The topic initiation function is only observed with the *speaking of X* variant, as the X-slot needs to be filled with the topicalized item.

3.3.2 Speech act adverbial

Following Leech's (1974: 356–359) definition, a *speech act adverbial* is an adverbial that modifies the speech act instead of the sentence to which it is attached. It identifies the modality, manner, or respect with which the speaker is going to perform the speech act. The linguistic elements that carry this function are sometimes also categorized as “style adverbials” (Biber et al. 1999: 857), “style disjuncts” (Greenbaum 1969: 81–93; Quirk et al. 1985: 615–618), or “manner-of-speaking markers” (Fraser 1996: 181; Kong & Qin 2017), as they are peripheral rather than integral to the clause structure and they draw attention not only to WHAT is said but also to HOW it is said. Some typical examples are shown in (90) to (92).

(90) *Quite frankly, he looked terrible.* (Biber et al. 1999: 857)

(91) *Metaphorically (speaking), he is a camel.* (Fraser 1996: 181)

(92) *Personally, I find the music too arid.* (Quirk et al. 1985: 616)

As a speech act adverbial, SPOX does not create any link between the new topic and the discursive or pragmatic pretext, but it makes explicit in which manner the new utterance is produced and/or in which respect the proposition contained in the matrix clause should be understood. In (93), for example, SPOX specifies that the speech act is produced in a “general” manner and that the proposition contained in the matrix clause only obtains when the topic about the West is understood in a “general” manner. In (94), while the clause-initial “speaking of people” is used to indicate that the matrix clause brings about a change of topic, “speaking of people and men” following the matrix clause specifies that the strangeness of the story should be understood in respect of “people and men”.

- (93) *Oh, I think we, speaking of the West generally, need to be far more engaged, more deeply engaged than we have been.*
(1992, SPOK, PBS_Newshour, COCA)

- (94) *Our dad, from day one, was like, make sure you buy your own house, make sure you buy property, make sure you finish school, make sure you do this and that. And people used to say, why are you bringing your daughters up like men and he goes, I'm not. I'm bringing them up as people. ... So speaking of people, this is kind of a strange story, speaking of people and men. Rapper Vanilla Ice, who is also apparently a reality star, was arrested for allegedly robbing the house next door to the one he was renovating for his show.*
(2015, SPOK, ABC, COCA)

In brief, even though the speech act adverbial SPOX has the same sentential scope over the same matrix clause as the topic-introducing SPOX, it provides very different pragmatic information.

3.3.3 Emphasized commenting

SPOX can also be used to emphasize that something is very noticeable in a stated way, being an extreme or striking example of a particular characteristic, state or situation. It is comparable to the expression “talk about X!”, as shown in (95) and (96) respectively. In (95), we see that Andrea Buchanan uses “speaking of legends” to add emphasis to the statement that Kris Kristofferson is a striking example of being a legend.

- (95) WINFREY: *We're here with Jennifer Aniston and her best friend, Andrea Buchanan. They co-directed Kris Kristofferson in a short film called "Room 10." [...] Were you intimidated at all, were you guys?*
Ms-BUCHANAN: *Speaking of legends...*
Ms-ANISTON: *It's Kris Kristofferson.*
Ms-BUCHANAN: *Yeah. I mean, he's... [...] He makes you – he's so – he's just who he is, there's nothing... [...] there's no other shades there, he's just an amazing...*
Ms-ANISTON: *Elegant.*
Ms-BUCHANAN: *Elegant.*
WINFREY: *OK, but I'm just asking, no intimidation? This is your first time ever doing it, you just walked in, 'Hi, I'm Jen. I'm your director'?*
Ms-BUCHANAN: *Oh, that.*

Ms-ANISTON: *Well, here's the funny...*

(2006, SPOK, Ind_Oprah, COCA)

(96) *All his plays have such ridiculous plots — talk about good drama!*

(Collins English Dictionary)

The commenting function is fundamentally different from the topic-introducing functions. Firstly, the commenting SPOX does not imply any shift in topicality or create any link between different topics but rather comments on an existing proposition in an emphatic way. Secondly, the X-element in the topic-introducing SPOX refers to an old piece of information derived from the discursive or pragmatic pretext, while the commenting SPOX brings forth a new proposition with an X-element that is not linguistically encoded or pragmatically given in the ongoing discourse. And thirdly, the topic-introducing SPOX is attached to a matrix clause, while the commenting SPOX suffices alone as a turn.

3.3.4 Speaking of the devil

This usage of SPOX represents a variation of the proverb “Speak of the devil (and he shall appear)”. This expression is originally rooted in the Christian belief that mentioning the devil by name would incite him to appear instantly and cause misfortune to befall you. But in modern-day English it is being used more lightheartedly to comment on the sudden or coincidental appearance of someone or something that has just been spoken of, such as in (97) and (98).

- (97) “*Yep, Angus buys Three Musketeers regular,*” Mr. Larsen continued, oblivious. “*Them and the Junior Mints, you know.*” Then, noticing Angus, he said, “*Well, well, speak of the devil, and doesn’t he just show up. Hello, Angus.*” (1993, FIC, BkJuv:LightHogback, COHA)
- (98) *Do you know – it’s a curious coincidence -- I was just talking about you with a friend of yours before you came in. Speak of the devil, you know.* (1920, FIC, Play:AnnaChristie, COHA)

The SPOX construction, in different forms of realization, is observed in the COHA dataset to be occasionally used for this function as well, as exemplified in (99) to (102).

- (99) WARE [...] *Yes, he is. That’s just what Gaskell is. Whatever his faults may be at least they’re honest, right out from the shoulder!*
 BAHN *I am not -- so sure. [...] Gute nacht.* (Giving her his hand)
 WARE *Good night, Fritz.*
 (FRITZ goes up to table by piano and picks up his violin case and overcoat. There is a knock at hall door.)
 WARE *Open the door.*
 (FRITZ opens door and MALCOLM GASKELL stands in the doorway. [...])
 BAHN *Speaking of the devil -- here is the old boy himself.*
 (1909, FIC, Play:MansWorld, COHA)
- (100) [...] *she murmured, half aloud. “ After tomorrow I can defy Lester Stanwick to bring one charge against me. [...]” “ Speaking of angels, you often hear ‘ the rustle of their wings.’ [...] ” said a deep voice beside her; and turning around with a low cry she saw Lester Stanwick himself standing before her.* (1889, FIC, DaisyBrooks, COHA)

(101) KHALIFAH [...] *Have you seen Sun? She hasn't return any of my phone calls. I'oun't know what's up with that.*

FERGUSON *I've seen her but I can't speak for her. She'll be down here in a little while.*

KHALIFAH *I ain't got much time. What time you expecting her?*

FERGUSON *Now. She's late. Probably got caught up in traffic.*

[...]

KHALIFAH *Man, it's cold in here. (Sees SUN.) Here she comes. (TO SUN.) Speaking of the queen.*

(2001, FIC, Play:SunRisingOnHill, COHA)

(102) “[...] *Crisp, light, and not too fruity, according to the waiter who was. Fruity, that is. Speaking of, here he comes.*” Gillian sat down across from her. *The waiter served her glass of Pinot Grigio, [...]*

(2001, FIC, Switch, COHA)

The *speaking of the devil* usage needs to be distinguished from the topic-introducing functions of SPOX. Instead of introducing a new topic into the discourse, it comments on the entrance of a real-world entity into the physical environment where the discourse takes place. As this usage is only observed in the COHA dataset, it will not be part of the further functional analyses, which are based on the COCA-SPOK dataset.

3.3.5 Peripheral positions

Based on the COCA-SPOK dataset, discourse marker SPOX can be used in **clause-initial (LP)**, **clause-medial (MP)** and **clause-final (RP)** positions, as shown in (103) to (107).

(103) ADRIAN-COVERT-CNN# *I was really impressed with Sony's Playstation Now service, which is essentially gamestreaming, which lets anyone connect to the Internet. [...] And it offers a promising glimpse at the future in gaming.*

BURKE# Speaking of the future, one of the best ones I saw were the curved televisions. When you're actually sitting there you kind of feel like you're in the picture.

(2014, SPOK, CNN, COCA)

(104) GAYLE-KING: [...] *Variety.com says Fonda will play First Lady Nancy Reagan in the movie The Butler. [...] It's a about a White House servant who worked alongside Presidents over thirty-four years, ending with Ronald Reagan. Liam Neeson may play Lyndon Johnson. John Cusack may play Richard Nixon.*

ERICA-HILL: *It could shape up to be quite a cast and speaking of first ladies, Michelle Obama arranged quite a school trip.*

(2012, SPOK, CBS_ThisMorning, COCA)

(105) *I think Kris knows exactly who he is. I think Adam obviously knows who he is. I just -- I have some new news. I just saw those two over at the set earlier today. And in the finale Tuesday, during the finale, round one -- speaking of songs -- will be the contestant favorite. So one of their favorite songs from the season. ...*

(2009, SPOK, CNN Newsroom, COCA)

(106) HAMMER: [...] *A scathing report released today by the all-Republican special House committee that was investigating the government's reaction to Hurricane Katrina. And I want to read you a quote from this report that came out today. It said that "If 9/11 was a failure of imagination, then Katrina was a failure of initiative." So*

Ben, I've got to ask you. How do you see the administration digging itself out of this one, speaking of P.R. messes?

FERGUSON: *Well, I don't know there's much you can do, because ...*

(2006, SPOK, CNN_Showbiz, COCA)

- (107) *And the family ties don't stop at the governor's office. House Speaker Michael Madigan has been criticized for waiting until today to launch the impeachment process against the governor. Critics say the speaker didn't want to steal the thunder from his daughter. His daughter is Attorney General Lisa Madigan, who Friday called for Blagojevich to be removed from office. Lisa Madigan is reportedly considering a run for governor herself in 2010. There could be another Kennedy in the Senate, speaking of family ties. CNN has learned that Caroline Kennedy told New York Democrats she wants to be the state's next senator, replacing Hillary Clinton.*

(2008, SPOK, CNN_Dobbs, COCA)

These peripheries are identified in relation to the matrix clause to which SPOX is attached, regardless of their positions in the utterance or the turn. For example, in both (103) and (104), SPOX is considered to occupy the left periphery, even though, in (104), SPOX and its matrix clause are embedded in a larger sentence structure. Similarly, in both (106) and (107), SPOX is considered to occupy the right periphery, even though, in (107), it does not occur at the end of the utterance or the turn.

In terms of the overall token frequency, Figure 6 shows that the left periphery is the most dominant position for the use of SPOX as a discourse marker. Uses in the medial and right peripheries do occur, but they are much rarer in comparison. The “not applicable (N/A)” cases correspond to the commenting uses of SPOX, as they are not attached to any matrix clause.

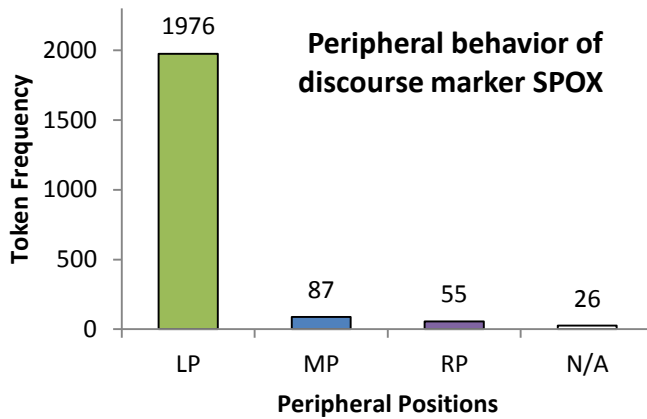


Figure 6 Frequency distribution of discourse marker SPOX in different peripheral positions in the COCA-SPOK dataset (sum = 2144)

With respect to different discourse marking functions, Figure 7 shows that the topic-introducing functions and the speech act adverbial function can be carried out in all of the peripheral positions but different functions seem to favor different peripheral positions: the topic-introducing functions prevail in the LP (94%), while the speech act adverbial function mostly occur in the RP (73%).

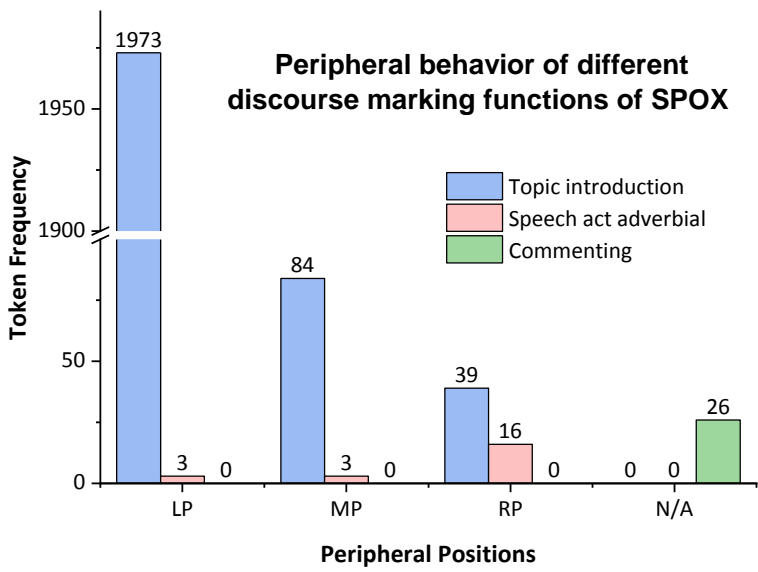


Figure 7 Frequency distribution of different discourse marking functions of SPOX in different peripheral positions in the COCA-SPOK dataset (sum = 2144)

The LP position is known to host discourse markers that typically assist the interpretation and processing of the following utterance and mark its relation to the preceding part, as supported by evidence from both corpus-based analyses and cognitive studies (e.g. Biber et al. 1999: 891; Virtanen 2004). It is thus natural for the topic-introducing SPOX to favor the LP position.

The low occurrences in the LP position, or rather, the seeming preference for the RP of the speech act adverbial SPOX, could be in fact a differentiation strategy. Firstly, the speech act adverbial function is by itself syntactically mobile. A speech act adverbial can be used in

all peripheral positions without compromising its function, which means its dependency on the LP is minimal. Secondly, the LP is the pivotal position for linking the new utterance to the previous utterance. It is thus a much more favorable and crucial position for the topic-introducing functions that concern the link between the successive utterances than for the speech act adverbial function that does not. Thirdly, the speech act adverbial function represents a much rarer way of using SPOX than the topic-introducing functions. Its association with the LP position is thus weaker than the more prominent topic-introducing functions.

Due to these reasons, the observed peripheral behavior of the speech act adverbial SPOX –preference for the RP and low occurrence in the LP – should be recognized as a SPOX-specific case only and it should not be viewed as a general property of speech act adverbials. Consequently, no categorical conclusion should be drawn about the functional asymmetry between the topic-introducing functions and the speech act adverbial function.

Neither the topic-introducing SPOX nor the speech act adverbial SPOX appears to be productive in the MP position. It has been pointed out in a few studies that the MP position lends itself particularly well to accommodating elements with epistemic modality and it often reflects online construction in unplanned oral discourse (see Beeching & Wang 2014; Pons Bordería 2018). It is thus a less relevant position for the particular discourse marking functions of SPOX.

In terms of subjectivity and intersubjectivity, SPOX represents a counterexample of the functional asymmetry hypothesis as well.

Generally speaking, SPOX expresses both subjectivity (speaker-orientation) and intersubjectivity (hearer-orientation) when used as a topic-introducing device and as a speech act adverbial regardless of the peripheral positions. It always indicates how the utterance is related to the speaker's perspectives or viewpoints and at the same time how the utterance is designed for the intended hearer.

First of all, the use of SPOX as a topic-introducing device reveals the speaker's evaluation of the relevance of the new utterance and their intention in the organization of their speech. As a speech act adverbial, it indicates the speaker's evaluation of the quality and quantity of the information contained in their utterance and their position toward the proposition contained in their utterance.

Secondly, the use of SPOX reflects the speaker's attention to the hearer, i.e. their assumption and assessment of the hearer's "here and now" in the speech event, which influences the organization of their utterance. The topic-introducing functions warn the hearer of the divergent topicality and the speech act adverbial function provides the hearer with metalinguistic information that facilitates the interpretation of the new utterance in the intended way.

The only difference arising from the peripheral variation is that, in non-initial positions, the functions of SPOX exhibit reparative characteristics: i.e. the later SPOX appears in the utterance, the more it invites a "repair" or "afterthought" interpretation to it. That is to say, instead of well-organized ahead of time, the speaker may only see the need to provide more information (e.g. the relation to the discursive pretext, the respect and manner of the speech act) to assist the hearer's interpretation as the new utterance proceeds. This kind of operation

would be much less likely to happen without the presence of a hearer. The non-initial uses thus can be interpreted as showing stronger intersubjectivity than the LP uses, as they represent additional effort in the hearer design and are therefore more expressive of the speaker's attention to the hearer. But overall speaking, the subjectivity and intersubjectivity of the discourse marking functions of the SPOX construction are not restricted to specific peripheral positions.

3.3.6 Dialogual and monological uses

The discourse marker SPOX can be introduced either by the speaker in their own course of talk, i.e. in a *monological* context, or by the hearer picking up a turn of speech, i.e. in a *dialogual* context, as exemplified in (108) and (109) respectively.

- (108) ROSIE O'DONNELL *When in doubt, you know, get a book. Or hire a professional, Dad. Yes. Well, you know, speaking of a professional, a wonderful woman, the first lady, Michelle Obama, says that people worried about the government getting into their lives too much can do a simple thing to keep them out, eat healthy.*
(2015, SPOK, ABC, COCA)

- (109) NINA-TERRERO# *She had bronchitis. [...] rumors are she didn't even make the rehearsals for the BBMAs because she was trying to rest her voice until the very last minute.*
HODA-KOT# *Speaking of vocal issue, Sam Smith had to use cue cards to make his acceptance speech.*
(2015, SPOK, NBC, COCA)

Generally speaking, SPOX is used more often in dialogual contexts than in monological contexts. In over two thirds of the occurrences in the COCA-SPOK dataset, SPOX is used by a different speaker to start a new turn, while less than one third of the uses are initiated by the same speaker (see Figure 8).

A similar distribution, i.e. a clear preference for the dialogual context, can be found in all of the variants of SPOX (see Figure 9). The “not applicable (N/A)” tokens refer to a few topic initiation events which are at the same time the initiation of brand new discourse events as well. Turn-taking operations are therefore not relevant in those cases.

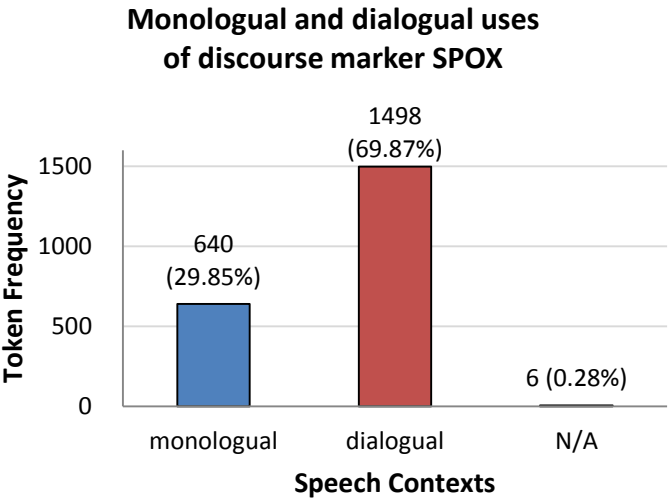


Figure 8 Frequency distribution of monological and dialogual uses of discourse marker SPOX in the COCA-SPOK dataset (sum = 2144)

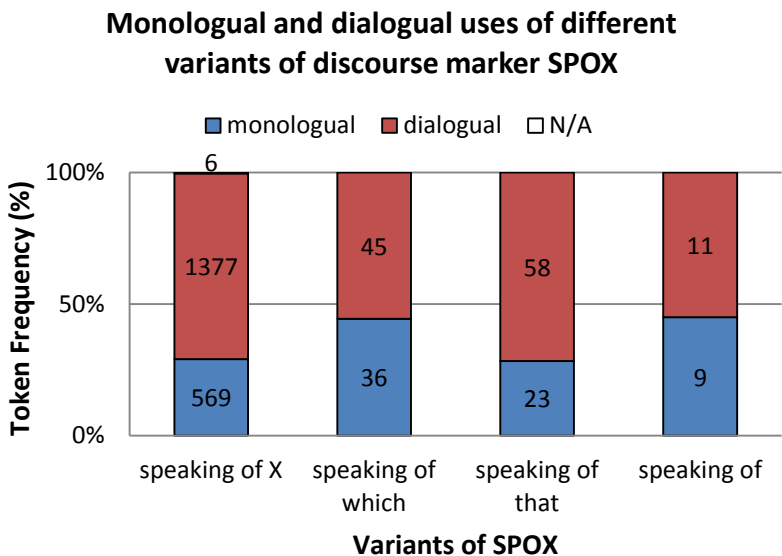


Figure 9 Frequency distribution of monologual and dialogual uses of different variants of discourse marker SPOX in the COCA-SPOK dataset (sum = 2134 excluding invalid tokens)

Regarding different discourse marking functions, as shown in Figure 10, SPOX is used more often in dialogual contexts than in monologual contexts as a topic-introducing device and a commenting device. That is to say, SPOX is used more often to mark (sub)topical movements across speakers than in one’s own course of talk and to comment on another speaker’s remarks than on their own. The speech act adverbial function, by contrast, is almost exclusively associated with monologual contexts. That means, in most of the cases, the speech act adverbial SPOX provides additional metalinguistic

information about the speaker’s own utterances and only in rare occasions is it used to support another speaker’s statements.

Within the topic-introducing functional domain, as shown in Figure 11, there are generally more dialogual uses than monological uses, regardless of the type of the topical movement. Only in topic initiation events is the dialogual-monological distinction sometimes irrelevant, i.e. when turn-taking operations are not involved.

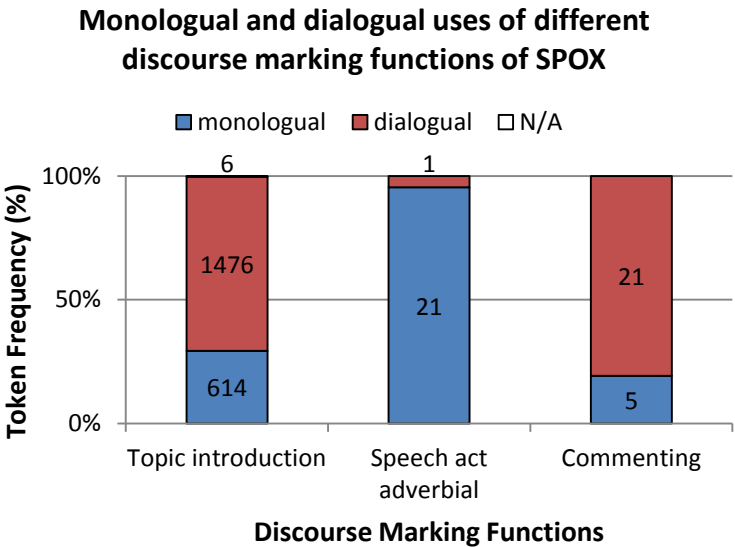


Figure 10 Frequency distribution of different discourse marking functions of SPOX in dialogual and monological contexts in the COCA-SPOK dataset (sum = 2144)

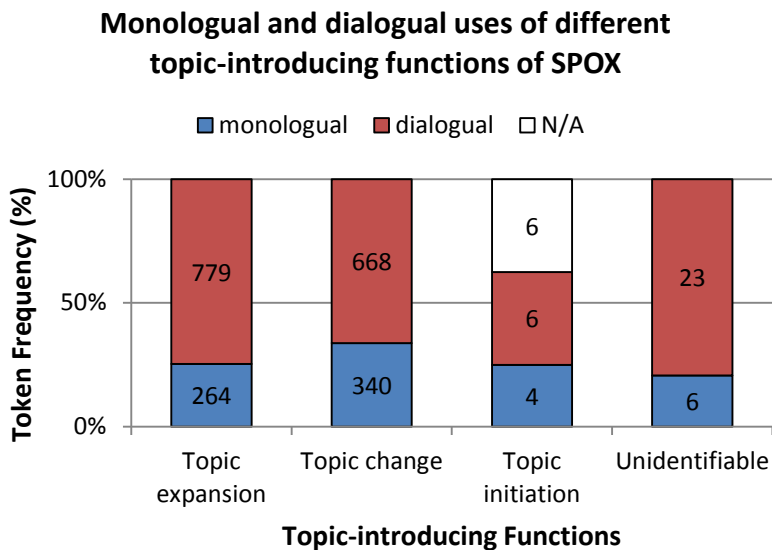


Figure 11 Frequency distribution of the topic-introducing functions of SPOX in dialogual and monologual contexts in the COCA-SPOK dataset (sum = 2096)

In sum, SPOX appears to be used more often in turn-taking operations when introducing a new (sub)topic into the discourse. As a commenting device, SPOX is mostly used to comment on another speaker’s speech, whereas as a speech act adverbial, it is more closely connected to the speaker’s own utterances.

3.4 Diachronic changes

3.4.1 Grammatical constructionalization

As presented in Section 3.1.2, SPOX can assume quite different functional roles in a sentence. Among them, the discourse marker usage appears to have developed from the adverbial modifier usage. In Figure 12 we can see that, while most of the functions of SPOX stayed relatively stable within a low frequency range throughout history, the adverbial modifier usage was once distinctively the most frequent. But it has been losing frequency drastically over time in stark contrast to the rise of the discourse marker usage in the more recent years.

The diachronic functional change of SPOX reflects a process of grammatical constructionalization, i.e. the meaning of SPOX has become no longer fully predictable from its lexical components but conventionalized in correspondence to its form. As an adverbial modifier, the meaning of SPOX is composed by its lexical components and refers to the action of speaking of a certain topic. It is used to modify the matrix clause and expresses a concomitant relationship between the speaking event and the event depicted in the matrix clause. As a discourse marker, however, the meaning of SPOX requires a constructional understanding of the expression: it is no longer entirely subject to the semantic content of the lexical components.

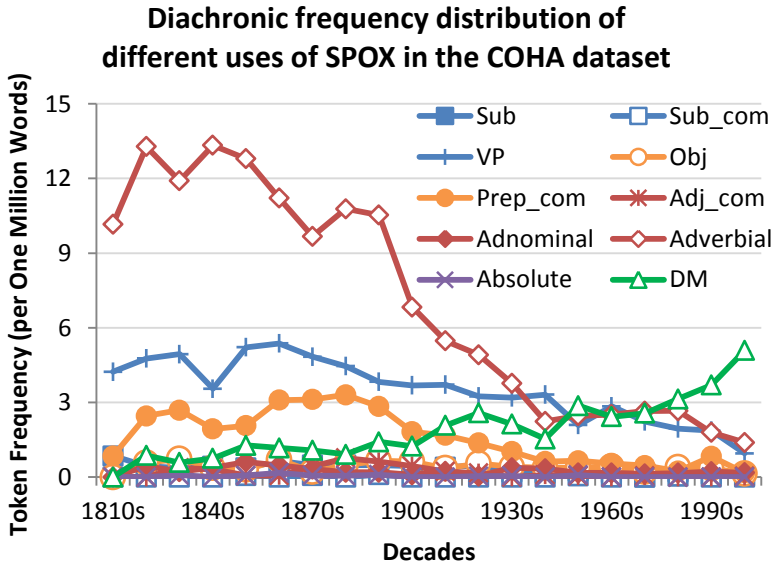


Figure 12 Diachronic frequency distribution of different uses of SPOX in the COHA dataset (sum = 5669)

A more recent change in variant usage also manifests further development of SPOX in the grammatical constructionalization process, namely the emergence and rapid growth of the *speaking of* variant (SPO). As shown in Figure 13, while the topic-introducing SPOX has been gaining frequency in all of its variants, with the *speaking of X* variant (SPOX) being the most productive one throughout the datasets, during the time period from 1990 to 2015, in which we observe an exponential growth rate of SPOX, there is also an evident rise of the *speaking of* variant.

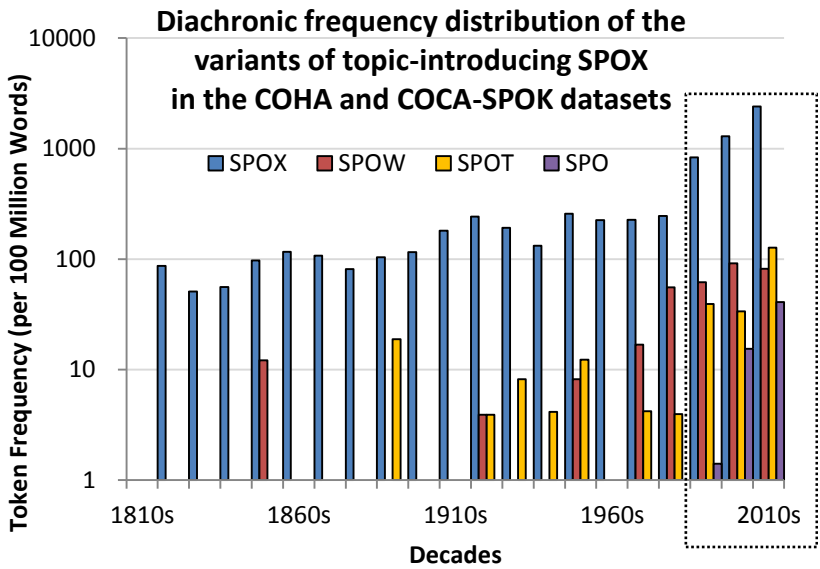


Figure 13 Diachronic frequency distribution of the variants of topic-introducing SPOX in the COHA and COCA-SPOK datasets (the Y-axis uses a logarithmic scale for better visual presentation)

In comparison to the *speaking of which* (SPOW) and *speaking of that* (SPOT) variants, which also create an implicit connection between the utterances, the SPO variant appears to be a more speaker-efficient option. It confirms the Pragmatic Principle of Least Effort “say no more than you must” (Zipf 1949) as long as the contribution is sufficient enough for the hearer to understand. And it shows that SPOX is becoming increasingly established as a construction: even the highly reduced and opaque form is paired up with the discourse marking functions sufficiently.

3.4.2 Pragmaticalization

Pragmaticalization in this thesis refers to the processes of pragmatic strengthening (see Section 2.3.3 for the definition and distinction of the term). The grammatical constructionalization process of SPOX is characterized by a process of pragmaticalization, i.e. the gradual establishment of the specific form-function pairing involves the development of pragmatic meanings and discourse marking functions out of the adverbial interpretation and usage of the *ing*-form participle clause “speaking of X”.

Within the time span of the COHA dataset, the discourse marking functions of SPOX do not show significant changes in their trends of development.

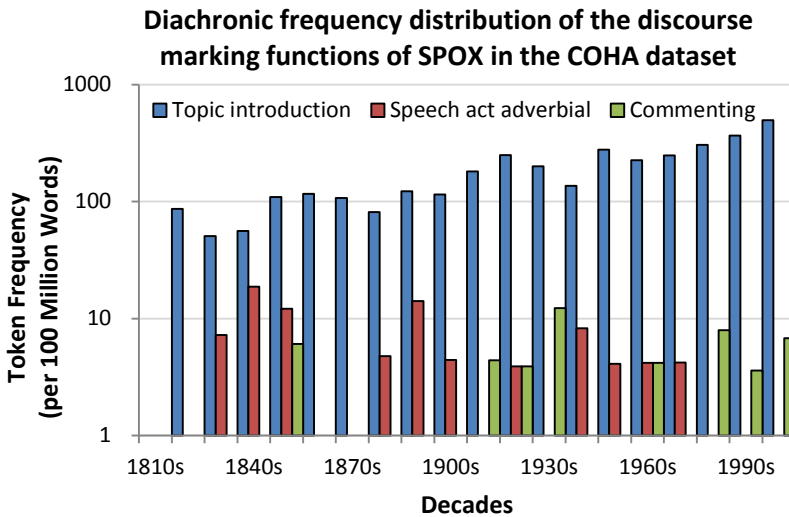


Figure 14 Diachronic frequency distribution of the discourse marking functions of SPOX in the COHA dataset (the Y-axis uses a logarithmic scale for better visual presentation)

As shown in Figure 14, the less typical speech act adverbial and commenting functions stay relatively stable with rather low frequencies, while the prevailing topic introduction function remains dominant and is even gaining frequency over time.

In more recent years, SPOX exhibits further development in the pragmatization process as a topic-introducing device.

First of all, in Figure 15 we see that the topic-introducing SPOX is becoming more frequently used in topic change events than in topic expansion events. It shows that the pragmatic meaning of SPOX in the functional domain of topic introduction is becoming more conventionalized and therefore its function in linking up unrelated topics starts to become increasingly accepted and even appreciated.

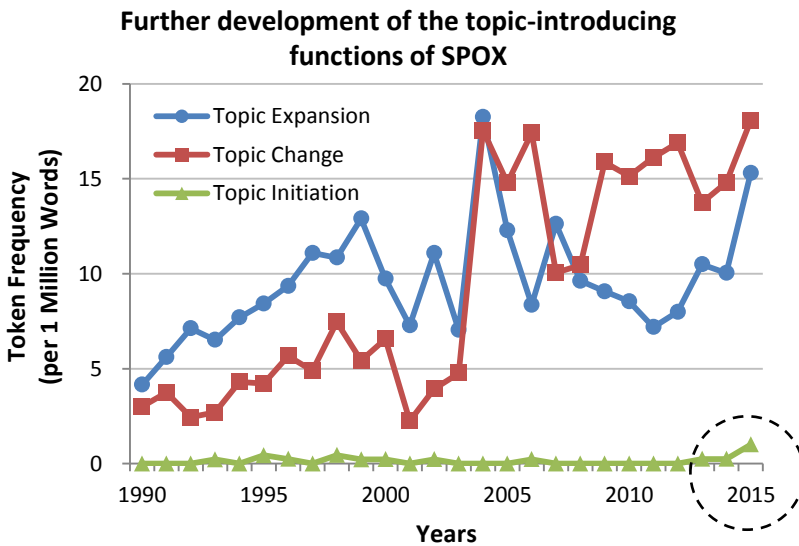


Figure 15 Diachronic frequency distribution of the topic-introducing functions of SPOX in the COCA-SPOK dataset

Secondly, Figure 15 also shows that, despite the small total number, there is a noticeable upward trend of the topic initiation function towards the end of the timeline. According to Pons Bordería and Estellés Arguedas (2009), who incorporated the Val.Es.Co Model of classifying discourse units (Briz Gómez et al. 2003) into the study of “New Relevant Information Markers (NRIMs)”, the function of topic initiation (assumed by discourse units of “absolute initiative interventions [aiI]”) is a further stage of the development of topic-introducing discourse markers, as illustrated in Figure 16.

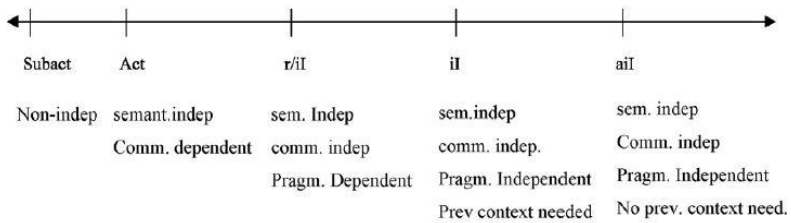


Figure 16 Cline of discourse units (Pons Bordería & Estellés Arguedas 2009: 934)

Within the Val.Es.Co model, the cline of discourse units corresponds to a process of gaining independence that leads from a totally dependent discourse unit to a totally independent one. The rising trend of the topic initiation function thus shows that SPOX is furthering along its development path as a discourse marker from being pretext-conditioned to pretext-independent. The linking mechanism of the topic initiation function, which is motivated by a pragmatic pretext instead of by a discursive pretext, also supports the

observation that the topic initiation function represents a step further along the pragmaticalization process in comparison to the topic expansion and topic change functions.

3.4.3 Change in peripheral positions

According to the results of the frequency distributional analyses of the COHA and COCA-SPOK datasets, the discourse marking functions of the SPOX construction seem to have first developed in the LP position and then gradually expanded to the MP and RP positions (see Figure 17).

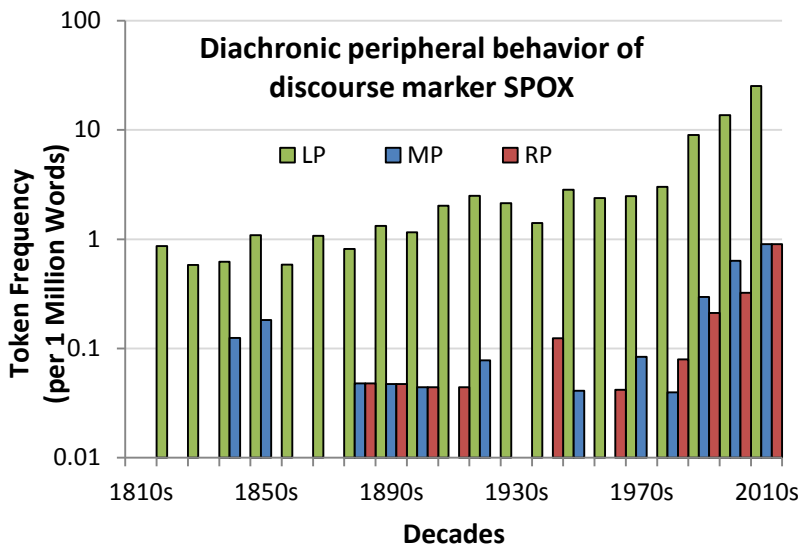


Figure 17 Diachronic frequency distribution of discourse marker SPOX in different peripheral positions in the COHA and COCA-SPOK datasets (the Y-axis uses a logarithmic scale for better visual presentation)

The expansion into non-initial positions reflects an increasing degree of syntactic non-integration and mobility of SPOX in its development as a discourse marker. It does not, however, represent positive evidence for the left-to-right movement hypothesis, which refers to the expected change in peripheral preferences from left to right during the intersubjectification process of a discourse marker (see Section 2.4.4 for more detailed discussion of the hypothesis).

In the case of SPOX, there is neither a change in peripheral preferences nor a correlated intersubjectification process. Firstly, as already discussed in Section 3.3.5, there is no strict correlation between the peripheral positions and the (inter)subjectivity of the discourse marking functions of SPOX. Secondly, in Figure 18 and 19 we can see that there are no significant diachronic changes in the peripheral behavior of the different discourse marking functions either: the LP is always the most dominant position for the topic-introducing functions, while the MP and the RP consistently accommodate more speech act adverbial uses than the LP. And lastly, there is also no sign of the speech act adverbial uses overpowering the topic-introducing functions.

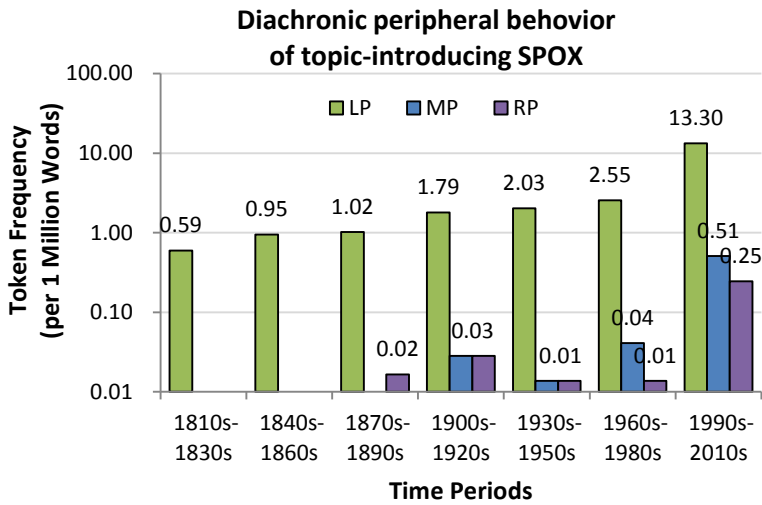


Figure 18 Diachronic frequency distribution of topic-introducing SPOX in different peripheral positions in the COHA and COCA-SPOK datasets (the Y-axis uses a logarithmic scale for better visual presentation)

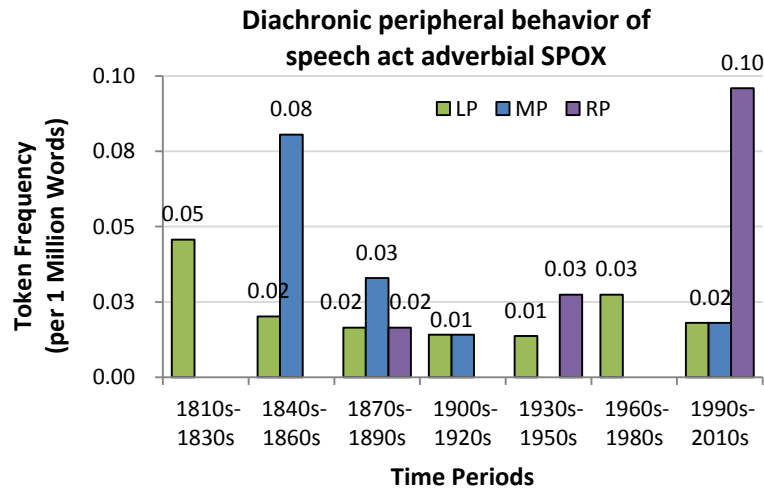


Figure 19 Diachronic frequency distribution of speech act adverbial SPOX in different peripheral positions in the COHA and COCA-SPOK datasets

The development of discourse marker SPOX so far thus does not support the left-to-right movement hypothesis. Its later expansion into the MP and RP positions from the original LP position, when viewed as a left-to-right movement at all, is not a manifestation of developing functional asymmetry and it does not coincide with intersubjectification.

3.4.4 Motivation for change

In the following subsections I present three dominant usage patterns associated with the adverbial usage of SPOX, which I believe have motivated the emergence of the current topic-introducing functions.

3.4.4.1 Sentence-initial position

According to the COHA data, adverbial SPOX can appear in *initial*, *quasi-initial*, *pre-verbal*, *post-verbal* and *final* positions of a sentence. The *initial* position refers to the slot at the very beginning of a sentence, e.g. in (110). The *quasi-initial* position is also situated at the front of a sentence before the main clause starts, but it is preceded by other elements, such as discourse markers or other adverbials, e.g. in (111). The *pre-verbal* and *post-verbal* positions both refer to the middle ground of a sentence, after the subject of the main clause has been introduced, but respectively before and after the predicate verb, e.g. in (112) and (113). The *final* position refers to the end slot in a sentence, which is as well post-verbal, but not followed by further elements, e.g. in (114).

3 The speaking of X construction

- (110) *Speaking of the Indians, he said: “Though they did abuse me mightily, [...]* (1849, FIC, SketchesCharacter, COHA)
- (111) *A month earlier, speaking of the same book, he had told a New Jersey audience: The idea [...]* (1990, MAG, RollingStone, COHA)
- (112) *A young college instructor, speaking of his freshman students, notes that “it’s not just the ignorance that [...]* (1988, MAG, Nation, COHA)
- (113) *He wrote, speaking of the bisexual problem: “In recent times we have begun to learn a [...]* (1970, NF, BuddhistLeader, COHA)
- (114) *“We could now play the game of war,” says Sherman, speaking of the plans for his Atlanta campaign. (1884, MAG, Century, COHA)*

Among these positions, the initial position appears to be the most common for adverbial SPOX, taking up 44.93% of all incidents (followed by post-verbal 23.13%, pre-verbal 20.22%, final 8.49%, and quasi-initial 3.24%). From a diachronic perspective, as shown in Figure 20, the initial position had been the primary position for adverbial SPOX for a very long time. Its decrease since the first half of the 20th century corresponds to the increase of the discourse marker usage: as the initial position became increasingly customary for discourse marker SPOX, adverbial SPOX started to appear more frequently in other syntactic positions.

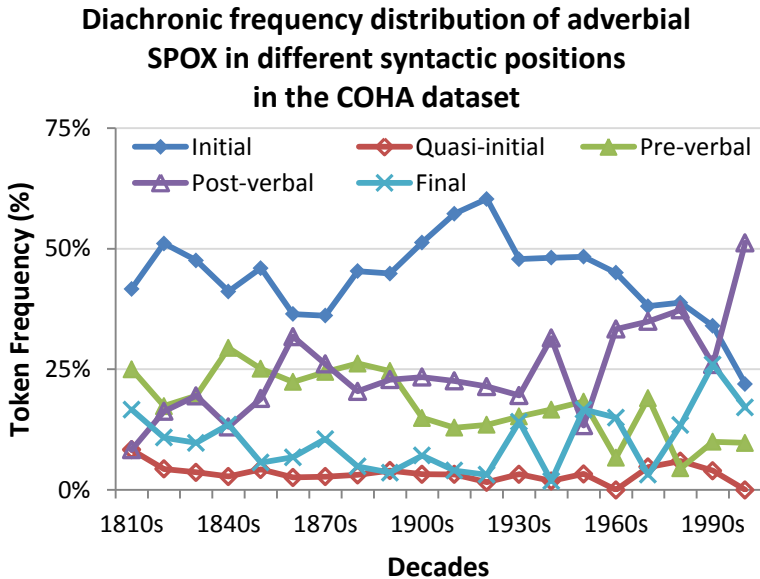


Figure 20 Diachronic relative frequency distribution of adverbial SPOX in different syntactic positions in the COHA dataset

It has been observed in a number of studies that, as the sentence-initial position has pragmatically ambiguous wide (sentential) as well as narrow (clause-internal) scope, it has been the locus where local reanalysis takes place and pushes adverbials towards becoming attitude markers and discourse organizers (e.g. Swan & Breivik 2011; Traugott 1995a; Virtanen 2004). The highly frequent usage of adverbial SPOX in the sentence-initial position has thus motivated the interpretation of SPOX as a sentence adverbial that highlights the “aboutness” of the whole statement and consequentially as a topic-introducing discourse marker that facilitates topical relevance and discourse coherence.

3.4.4.2 Asyndetic subordination

According to the COHA data, adverbial SPOX can be used either syndetically or asyndetically, i.e. with or without subordinators (e.g. *when, after, by*, etc.) when connected to the matrix clause. In Figure 21, it appears that asyndetic coordination is the most frequent usage pattern for adverbial SPOX, taking up almost 50% of all the incidents.

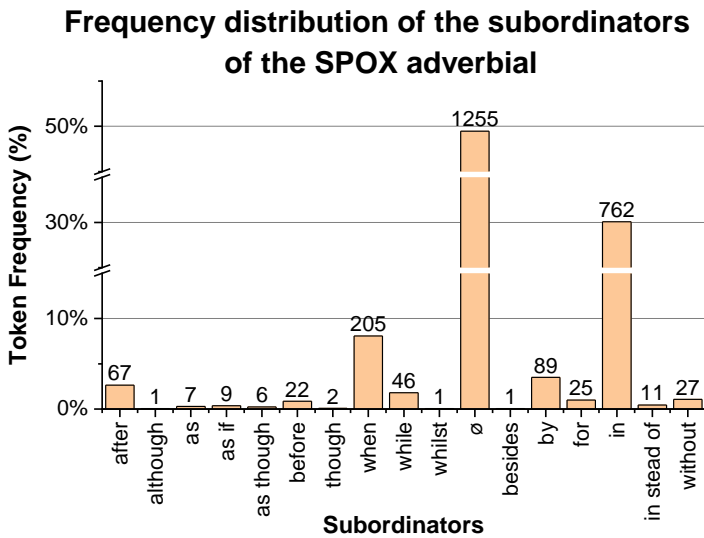


Figure 21 Frequency distribution of the subordinators of the SPOX adverbial in the COHA dataset (the Y-axis has a break between 15% and 25% and another between 35% and 45% for better visual presentation)

While the presence of subordinators determines (or delimits the interpretation of) the semantic relations between the adverbial clause and the matrix clause, asyndetic subordination allows highly versatile

interpretations (Quirk et al. 1985: 910; König & Van der Auwera 1990: 342). That means, asyndetic SPOX adverbials can be used to express various inter-clausal semantic relations, such as temporal, causal, conditional, etc., which can usually be successfully identified by the hearer based on pragmatic reasoning and their world knowledge (Kortmann 1991: 105–204).

Furthermore, Figure 21 also shows that, even when adverbial SPOX is indeed used with syndetic subordination, the most frequent subordinator among all is “in”, which allows versatile interpretations of the semantic relation as well.

In short, the dominant asyndetic subordination and the most often unspecified semantic relation between adverbial SPOX and the matrix clause made it a functionally versatile element in the sentence, which over time motivated constructional change of the phrase.

3.4.4.3 Co-occurrence with speech-act verbs

Looking into the verbal predicates in the matrix clauses to which the SPOX adverbials are attached, I discover that most of them (74.25% of all the incidents) are speech-act verbs, i.e. verbs that describe various acts of speaking in either verbal or written, specific or generic forms (Austin 1962; Searle 1969; 1976), as exemplified in (115) to (117).

- (115) *Hume, speaking of this adventure, says they burned St Anthony and St Helen, [...]* (1821, MAG, NorthAmRev, COHA)

- (116) “He can be the regenerator of his country,” *wrote* Shelley, *speaking of* Byron, in 1818, at Venice. (1869, NF, MyRecollections, COHA)
- (117) *Speaking of the Circassian women, the tourist thus descants*: “To how many thousands of Circassian maidens has this been the bright surveying point of a brilliant destiny! [...]” (1850, MAG, USDemRev, COHA)

The way SPOX is structured, consisting of the verb phrase “speaking of” and a variable X-element, shows that its semantic content does not really concern HOW the speech act is carried out, i.e. what speech act adverbials such as *frankly speaking* and *specifically speaking* do. Instead, the semantic content of SPOX refers to the action of speaking and points out the topic of the speech event. As an adverbial, it thus provides the information that the event described by the verbal predicate in the matrix clause is accompanied by the event of speaking of a certain topic. For example, in (118) and (119), the SPOX adverbial informs the reader that the events of “sitting” and “raising herself” are accompanied by the event of “speaking”.

- (118) *There they all sat at the round table, speaking of many things*; (1933, FIC, Harpers, COHA)
- (119) *She had raised herself on one naked elbow, speaking of this man as if he were a lover not a stranger to her*, [...] (1936, FIC, DeathMan, COHA)

However, as most of the co-occurring verbal predicates are speech-act verbs themselves depicting speech events, it would be

semantically redundant if one focuses on the “speaking of” part of the phrase when interpreting the semantic function of SPOX in these cases. This frequent co-occurrence pattern therefore gave prominence to the “topic” part of the phrase. Over time, the use of SPOX thus became more strongly associated with the introduction of topics and less strictly dependent on the lexical meaning of each component.

3.5 Summary

In this chapter I presented the results of an in-depth corpus-based study of the SPOX construction. I discussed its formal and semantic properties as well as its discourse marking functions, with both functional and frequency distributional analyses and from both synchronic and diachronic perspectives.

The SPOX construction appears to be polyfunctional as a discourse marker. It is typically used as a topic-introducing device, especially when the new topic represents some degree of topical movement from the previous discourse. It facilitates the perception of topical relevance and discourse coherence by linking the new topic to existing information, which is either contained in the discursive pretext or shared in the interlocutor relationship. Less frequently, the SPOX construction can also be used as a speech act adverbial, a commenting device, and a variant of the “speak of the devil” proverb.

The different meanings of SPOX as a discourse marker always require a constructional understanding and cannot be fully predicted from its lexical components or other structurally similar *ing*-form constructions.

Diachronically, the development of SPOX reflects the processes of grammatical constructionalization and pragmaticalization, i.e. the change from assuming a central propositional role in a sentence to playing an ancillary procedural role in the utterance. I proposed that such development originated from the once most frequent adverbial usage of SPOX before the emergence of discourse marker meanings. The dominant usage patterns associated with adverbial SPOX, including the sentence-initial position, asyndetic linking to the matrix clause, and co-occurrence with speech-act verbs, have all contributed to the emergence of discourse marker SPOX.

Chapter 4

The *huashuo* construction

This chapter aims to provide a comprehensive empirical analysis of the *huashuo* construction, based on both historical and contemporary corpus data and addressing both functional and frequency distributional aspects. In the following sections, I will first introduce the corpus data (Section 4.1), before moving on to analyzing the lexical structure and semantic content of *huashuo* (Section 4.2). After that, I will attend to the functional aspects, presenting the functions and usage patterns of *huashuo* as a discourse marker in modern Chinese (Section 4.3). And finally, I will turn to a diachronic perspective, tracing the development path of *huashuo* towards a discourse marker and discussing the motivating factors behind it (Section 4.4).

4.1 Data and methods

4.1.1 Corpora and datasets

The data sources for the study of the *huashuo* construction start with two large-scale corpora of the Chinese language: the CNCORPUS¹⁹ and the CCL corpus²⁰, which cover partly overlapping yet partly

¹⁹ The CNCORPUS is available online at <http://corpus.zhonghuayuwen.org/>. (Accessed: 19 March 2022) (Previously at <http://www.cncorpus.org>)

²⁰ The CCL corpus is available online at http://ccl.pku.edu.cn:8080/ccl_corpus. (Accessed: 18 March 2022)

complementing source material and time spans. The **CNCORPUS** is compiled by the Chinese National Operating Committee of Language Affairs. It consists of two sub-corpora of ancient and modern Chinese (further referred to as the **CN_Ancient** corpus and the **CN_Modern** corpus). The CN_Ancient corpus contains ca. 70 million words of written texts of classic literature ranging from the Zhou Dynasty to the Qing Dynasty (i.e. 1046 BC to 1921 AD), and the CN_Modern corpus contains ca. 100 million words of POS-tagged written texts from 1919 to 2002. The **CCL** corpus is developed by the Center for Chinese Linguistics of Peking University. It comprises two sub-corpora of ancient and modern Chinese as well (further referred to as the **CCL_Ancient** corpus and the **CCL_Modern** corpus). The CCL_Ancient corpus contains ca. 160 million words and covers a time span from the Zhou Dynasty to the Republic of China (i.e. 1046 BC to 1949 AD), and the CCL_Modern corpus contains over 500 million words of both written and spoken data from the end of 1910s to the beginning of the 2010s.

These corpora, with their substantial size and especially their considerable time span, support the diachronic analyses of the *huashuo* construction. They complement each other not only in terms of their content but also in terms of the different types of analysis allowed by virtue of their different corpus designs.

On one hand, the CNCORPUS does not provide information about the total number of words per each time period. This missing information makes it impossible to identify the normalized diachronic frequency distribution of the different uses of *huashuo*. That is to say, with the CNCORPUS data one can only compare which usage of

huashuo is more frequent in a certain period of time in relation to its other possible uses, but one cannot ascertain whether this usage is frequent or not in general. Therefore, the CCL corpus is used to back up the analyses in that respect.

On the other hand, as the CCL corpus is not POS-tagged, it falls short in supporting the semantic analysis of the *huashuo* construction, especially in time periods before it developed discourse marker meanings. The POS-tagged CNCORPUS proves to be especially useful in this respect, as it provides information about the absolute and relative token frequencies of POS-tagged linguistic items in the corpus. Specifically as this thesis is concerned, it provides the information about how frequently the components of *huashuo*, i.e. “hua” and “shuo”, are used as verbs and as nouns individually. This information facilitates the understanding of the semantic interpretation and the gradual development of *huashuo* (see Section 4.2.1 and Section 4.4.4 for the relevant discussion).

One limitation of these corpora is the weak representation of spoken language. The texts that are categorized as “spoken language” include literary works that are the compilation of oral narrations, more conversation-intensive types of literature such as play scripts, and transcripts of TV programs of various kinds. But these texts take up only a very small percentage of the whole data (ca. 1.76% in the CNCORPUS and ca. 2.47% in the CCL corpus). Consequently, the discourse marker usage of *huashuo* in modern Chinese is not well represented in either corpus (i.e. 2 tokens in CN_Modern and 11 tokens in CCL_Modern).

In order to compensate for this drawback, I compiled a small-sized corpus of computer-mediated conversation (further referred to as the **Chat** corpus). It comprises texts of chats via instant messaging applications such as *MSN*, *WeChat* and *WhatsApp*, containing ca. 3 million words in total and covering a 16-year time span from 2004 to 2019. This corpus serves as the foundation for the analyses of the current functions and usage patterns of *huashuo* in conversational contexts.

The instant-messaging type of computer-mediated communication (CMC) has an immediate and highly interactive nature, which makes it speech-like. It has some unique characteristics that do not appear in face-to-face oral conversations. Most notably, it allows the presence of multiple simultaneous threads (e.g. Garcia & Jacobs 1999; Herring 1999; Markman 2005; Paolillo & Zelenkauskaitė 2013: 119–121). Since the interlocutors cannot see or hear what each other is doing in a CMC setting, they lack the paralinguistic cues to coordinate their activities. That means, an interlocutor might introduce a new thread before a previous one is finished, or everyone could react at the same time, which generates multiple turns and potentially new threads, resulting in dispersed topic units and disrupted adjacency of turns. As a result, discourse coherence is constructed and perceived differently in the CMC setting. The speakers focus more on topicality and rely less on sequential connection to compute the coherent understanding of the utterances (e.g. Dresner & Barak 2006; Herring 1999; 2013; Markman 2013).

As far as the use of *huashuo* is concerned, even though the CMC data bring forth extra complexity in analyzing turn-taking

operations, they are sufficient for studying the topic-introducing and interpersonal functions: the use of *huashuo* is always conditioned by the topical relations between the utterances and the interlocutor relationships, regardless how dispersed the corresponding turns are.

Altogether, the working material for the study of *huashuo* is a combination of data from three different corpora. The initial datasets were obtained through KWIC searches of the text string “话说 [huashuo]” in all of the corpora. The searches returned 2051 tokens from CN_Ancient, 86 tokens from CN_Modern, 6725 tokens from CCL_Ancient, 20409 tokens from CCL_Modern, and 556 tokens from Chat. Each token was a cluster of texts of various lengths containing the keyword “话说 [huashuo]” in different positions. The searches in the CNCORPUS and the CCL corpus were powered by the corpus websites. Metalinguistic information of *year* and *source* (e.g. book title) were provided alongside the text tokens. The search in the Chat corpus was conducted using AntConc (Version 3.5.8). Metalinguistic information of *year* and *source* (i.e. folder, filename, application) were added to each token manually.

In the next step, manual screening was done to exclude duplicate tokens and invalid tokens (e.g. accidental juxtaposition of “hua” and “shuo” in a sentence), which turned out to be a large percentage of the initial datasets, especially for the CN_Modern and CCL_Modern datasets. Eventually, my final datasets contain 1580 tokens in CN_Ancient, 5 tokens in CN_Modern, 4456 tokens in CCL_Ancient, 205 tokens in CCL_Modern, and 407 tokens in Chat. The distribution of the data in the datasets by time period/year is shown in Table 6 and Table 7.

Table 6 Datasets CNCORPUS and CCL composition (number of tokens per time period)

Time Period	CNCORPUS		CCL	
	Ancient	Modern	Ancient	Modern
Six Dynasties [六朝] (222-589)	1	0	1	0
Tang [唐] (618-907)	0	0	2	0
Five Dynasties [五代] (907-979)	1	0	1	0
Song [宋] (960-1279)	6	0	9	0
Yuan [元] (1271-1368)	241	0	19	0
Ming [明] (1368-1644)	987	0	1171	0
Qing [清] (1636-1912)	344	0	2323	0
Modern [现代] (1912-2019)	0	5	930	205
Sum	1580	5	4456	205

Table 7 Dataset Chat composition (number of tokens per year)

Year	2004	2005	2006	2007	2008	2009	2010	2011
Tokens	0	0	0	4	16	21	87	75
Year	2012	2013	2014	2015	2016	2017	2018	2019
Tokens	51	29	26	18	33	15	23	9
Sum	407							

The examples presented in this thesis are mainly taken from these datasets and marginally cited from other studies. For privacy-protection reasons, no real names of the interlocutors or identity-revealing information (such as chat account or nicknames) will be shown in the examples without manipulation. The names of the source files, which are machine-generated sequences that contain the usernames of the chat participants, are also anonymized at the corresponding spots.

4.1.2 Classification and annotation

The *huashuo* construction has been used quite differently in different periods of time during the history of the Chinese language. The datasets thus contain not only the most recent discourse marker uses, but also other (archaic) uses of *huashuo*. It is therefore an essential part of the study to classify these uses in order to analyze the functions of *huashuo* as a discourse marker and to trace its development path.

I identify the following different uses of *huashuo* throughout the datasets, as listed in Table 8. Each type of usage is illustrated with at least one example. The raw token frequencies (*f*) of each type of usage in different datasets are provided accordingly.

Table 8 Different uses of *huashuo*

Type	Example		
Verb phrase	(120) 待 我 不 <u>话说</u> 时 则 闻。 Dǎi wǒ bú <u>huàshuō</u> shí zé wén. wait I NEG speak time CNJ listen “I’ll listen when I stop <i>speaking</i> .” (Zǔtángjí [祖堂集], 900s, CN_Ancient)		
	<i>f</i> (CNCORPUS)	<i>f</i> (CCL)	<i>f</i> (Chat)
	12	24	0

Noun phrase	(121) 西门庆 道: “干娘 ²¹ , Xīmén q òng d ào: “gānniáng, Ximen Qing say madam 和 你 说 正经 话说。 hé nǐ shuō zhèngjīng huàshuō.” with you say serious talk “Ximen Qing says, ‘Madam, (I) will have a serious <u>talk</u> with you.’” (Shuǐhǔzhuàn [水浒传], 1300s, CN_Ancient)		
	f(CNCORPUS)	f(CCL)	f(Chat)
	12	105	0
Subject- Verb clause	(122) 话说 人生 万事, Huàshuō rénshēng wànshì saying say life everything 前数 已 定。 qiánshù yǐ d òng. previous fate already determined “ <u>As the saying goes</u> , everything in life has been predestined.” (Chūkè Pāi‘àn Jīngqí [初刻拍案惊奇], 1627, CN_Ancient)		

²¹ “Gānniáng [干娘]” is an honorific used in ancient Chinese to address old women.

	<p>(123) <u>话说</u> 晋朝 有 一人， <u>Huàshuō</u> Jìn cháo yǒu yì rén, story say Jin Dynasty have one person</p> <p>姓 石， 名 崇， 字 季伦。 xìng shí míng chóng, zì jì lún. surname Shi first name Chong courtesy name Jilun</p> <p><i>“<u>The story tells</u>, there was a person in the Jin Dynasty, whose surname was Shi, first name was Chong, and courtesy name was Jilun.”</i></p> <p>(Huàběn [话本], 1200s, CCL_Ancient)</p> <table><tr><th><i>f</i>(CNCORPUS)</th><th><i>f</i>(CCL)</th><th><i>f</i>(Chat)</th></tr><tr><td>1559</td><td>4520</td><td>5</td></tr></table>	<i>f</i> (CNCORPUS)	<i>f</i> (CCL)	<i>f</i> (Chat)	1559	4520	5
<i>f</i> (CNCORPUS)	<i>f</i> (CCL)	<i>f</i> (Chat)					
1559	4520	5					
Discourse marker	<p>(124) 这回 我 在 市里， zhè huí wǒ zài shì lǐ, this time I LOC city in</p> <p>有 问题 打 电话。 yǒu wèn tí dǎ diàn huà have problem make phone call</p> <p><u>话说</u> 你 得 买 个 <u>Huàshuō</u> nǐ děi mǎi ge speaking of which you must buy CLF</p> <p>德国 手机 卡。 dé guó shǒu jī kǎ. German cellphone card</p>						

	<p>“I’ll be in town this time. If (you) have any problem, call me. <u>Speaking of which</u>, you need to buy a German SIM card.”</p> <p>(Chat History between K and W 2013-08-11.txt, 2013, Chat)</p> <p>(125) C: <u>话说</u> 你 今年 啥时候 回来 呀 <u>Huàshuō</u> nǐ jīnnián shíshíhòu huí lái ya so you this year what time come back Q</p> <p>W: 十一 Shíyī October first</p> <p>“C: <u>So</u> when are you coming back this year? W: October 1.”</p> <p>(Chat History between C and W 2017-08-28.txt, 2017, Chat)</p> <p>(126) 我 忽然 看到 你 发 我 Wǒ hūrán kàndào nǐ fā wǒ I suddenly see-COMPL you send me</p> <p>周日 两点 见 zhōurì liǎngdiǎn jiàn Sunday two o’clock see</p> <p><u>话说</u> 是 打错了 吧…… <u>Huàshuō</u> shì dǎcuò le ba… I mean²² COP type wrong-PST MOD</p>
--	--

²² See Schiffrin (1987: 296–304) for discussions of the intention-oriented explanatory discourse marker use of *I mean*.

	<p>“I just saw that you sent me: see you on Sunday at 2 o'clock. <u>I mean</u>, (it) was a mistake right?”</p> <p>(Chat History between S and W 2013-08-16.txt, 2013, Chat)</p>		
	$f(\text{CNCORPUS})$	$f(\text{CCL})$	$f(\text{Chat})$
	2	12	402

In all of the datasets, the *huashuo* tokens are annotated with “functional role (verb phrase, noun phrase, discourse marker, etc.)”, along with the metalinguistic information “time period” and “source” that were already recorded in the data collection phase. A number of additional parameters that are relevant to specific synchronic functional analyses and diachronic analyses were added to different datasets separately. In the Chat dataset, the additional parameters include “discourse marking function (topic introduction, interpersonal)”, “discourse event (topic change, expansion, initiation)”, and “position (sentence-initial/-medial/-final)”. In the CCL dataset, the additional parameters center on the narrative starter usage of *huashuo*, which I consider has motivated the constructionalization of discourse marker *huashuo* (detailed discussion follows in Section 4.4.4). Specifically, information about the type of the “initial position (chapter-/paragraph-/sentence-initial)” of this usage was added to the dataset. During the annotation process, unclear cases due to limited context size were searched again in the corpora in order to be clarified through expanded contexts.

4.2 Formal and semantic properties

4.2.1 Semantic content

The *huashuo* construction has a compound structure with two constituents: *huà* [话] and *shuō* [说]. Both constituents can function either as a verb or as a noun on their own, generally meaning ‘speak’ or ‘speech/utterance’, respectively.

The lexeme *huà* as a verb (further written as **hua_V**) means ‘talk about a certain topic’ and it requires the existence of an object in its usage, as shown in (127). The much more dominant part of speech of *huà* is however a noun (further written as **hua_N**)²³. It means ‘the verbally produced words that carry certain meaning’, as exemplified in (128). This general meaning is open to different interpretations in different contexts and can thereby develop more specific meanings, such as ‘topic’ in *huà í* [话题], ‘message’ in *shāohuà* [捎话], ‘conversation’ in *tánhuà* [谈话], ‘story’ in *huàběn* [话本], and even ‘chapter’ in *dìyīhuà* [第一话].

- (127) 话 旧
 huà jiù
 talk about old

“***talk about*** the old times”

²³ According to the frequency list provided by the CN_Modern corpus, *hua_N* (f=0.5539‰) is about 230 times more frequently used than *hua_V* (f=0.0024‰).

- (128) 没 说 多少 话
 mǎi shuō duōshǎo huà
 NEG say many words

“*did not say much*”

The lexeme *shuō* functions most dominantly as a verb (further written as **shuo_v**), meaning ‘verbally express meaning through words’, and it is the generic lexical item for a general speech act, as exemplified in (129). Used much less frequently as a noun²⁴, *shuō* (further written as **shuo_N**) means ‘the saying by someone or about something’ and could be interpreted as ‘theory’ or ‘doctrine’ in different contexts, as exemplified in (130).

- (129) 从 总体 上 说
 cóng zǒngtǐ shàng shuō
 from overall LOC speak

“*overall speaking*”

- (130) 仁学 之 说
 rén xué zhī shuō
 study of benevolence ADJ saying

“*the doctrine of benevolence*”

²⁴ According to the frequency list provided by the CN_Modern corpus, *shuo_v* (f=3.6675%) is almost 300 times more frequently used than *shuo_N* (f=0.0130%).

With a compound structure, *huashuo* thus lends itself to different ways of structural and semantic interpretation depending on the different combinations of the parts of speech and meanings of the components (see Table 9).

Table 9 Internal structural variations of *huashuo*

		shuo	
hua		Verb	Noun
	Verb	hua _V shuo _V	hua _V shuo _N
	Noun	hua _N shuo _V	hua _N shuo _N

However, not every combination shown in Table 9 obtains. First of all, the compound hua_Vshuo_N is not observed in my data. Although the VO structure of the phrase is grammatically permissible, the meaning of the construct is restricted by the non-dominant semantic content of both components, forcing an interpretation of ‘talk about the topic of “speaking”’ for instance, which does not make sense without any specification or context.

Secondly, if the compound hua_Nshuo_N is understood as a modifier-noun structure²⁵, its semantic interpretation is subject to very specific lexical meanings, e.g. ‘the theory of words’, which does not exist as an established term and is not observed in my data.

However, if hua_Nshuo_N is construed as a **copulative compound noun**, in which hua_N and shuo_N are interpreted as synonyms of

²⁵ See Dong (2017) for detailed discussion about the structural types and internal semantic relations of two-character words in Chinese.

‘speech’, it can mean ‘talk’ or ‘speech’, which is indeed observed in my data, as shown in examples (131) and (132).

- (131) 西门庆 道：“干娘， 和 你 说 正经 话说。”
Xīmén Qīng dào: “gānniáng, hé nǐ shuō zhèngjīng huàshuō.”
Ximen Qing say madam with you say serious **talk**
“Ximen Qing says, ‘Madam, (I) will have a serious talk with you.’”
(Shuǐhǔzhuàn [水浒传], 1300s, CN_Ancient)

- (132) 一家儿 听得 这般 话说 [...]
Yì jiā ér tīng dé zhèbān huàshuō [...]
one family person hear-COMPL this kind **speech**
“The whole family heard such a speech [...]”
(Xīyóujì [西游记], 1500s, CN_Ancient)

Thirdly, hua_Vshuo_V can function as a **copulative compound verb**, which expresses the shared semantic content of hua_V and $shuo_V$, meaning ‘talk’ or ‘speak’, as shown in (133) and (134).

- (133) 只好 话说 往来
zhǐhǎo huàshuō wǎnglái
can only **talk** back and forth
“could only make small talks”
(Chūkè Pāi’àn Jīngqí [初刻拍案惊奇], 1627, CN_Ancient)

- (134) 待 我 不 话说 时 则 闻。
Dǎi wǒ bú huàshuō shí zé wén.
wait I NEG **speak** time CNJ listen

“I’ll listen when I stop speaking.”

(Zǔtángjí [祖堂集], 900s, CN_Ancient)

Lastly, hua_Nshuo_V could be understood as a **subject-verb clause**, meaning roughly ‘it speaks of’ or ‘it talks about’. This semantic interpretation has its roots in the story-telling context, in which it is used as a narrative starter and therewith develops more concrete meanings such as ‘proverb says’ or ‘story tells’, as shown in (135) and (136).

- (135) 话说 人生 万事, 前数 已 定。
Huàshuō rénshēng wànshì qíánshù yǐ dìng.
saying say life everything previous fate already determined

“As the saying goes, everything in life has been predestined.”

(Chūkè Pāi’àn Jīngqí [初刻拍案惊奇], 1627, CN_Ancient)

- (136) 话说 石生 在 广州 做 巡抚。
Huàshuō shíshēng zài guǎngzhōu zuò xúnfǔ.
story say Shisheng LOC Guangzhou do governor

“The story tells that Shisheng is working as a governor in Guangzhou.”

(Huànzhōngyóu [幻中游], 1767, CN_Ancient)

The above shown semantic interpretations of the *huashuo* compound, different as they are, are all predictable from the lexical meanings of its components. However, as a discourse marker, *huashuo* can no longer be properly understood through a literal reading of its lexical components. As exemplified in (137), the copulative nominal and verbal interpretations would not fit in the sentence structure; the SV-clause interpretation would not suit the speech context either. Instead, *huashuo* is used as a topic-introducing discourse marker that prepares the hearer for a certain degree of topical movement, which might endanger discourse coherence. In example (137), without the use of *huashuo*, the transition from “J should fight more in a video game” to “J’s fight with her mum” would have appeared more abrupt.

(137) P: 你 也 放慢 速度 哈
 Nǐ yě fàngmǎn sùdù ha
 you also slow down speed MOD

多 打 大菠萝
 duō dǎ dàbōluó
 more fight Diablo

J: 话说 昨天 跟 我妈 打仗 来着
Huàshuō zuótiān gēn wǒ mā dǎzhàng lái zhe
 speaking of yesterday with my mum fight PST

“P: *You should slow down too. Fight more Diablo.*

J: *Speaking of*, *I fought with my mum yesterday.*”

(615163155(new).txt, 2010, Chat)

To sum up, the semantic content of *huashuo* as a discourse marker cannot be fully represented by the lexical meanings of its components. *Huashuo* needs to be recognized as a symbolic unit with a conventionalized form-function pairing. Its discourse marking functions and usage patterns will be discussed in detail in Section 4.3.

4.2.2 Formal fixedness

The *huashuo* construction has a more fixed form than the SPOX construction. It does not have a variable X-element in the structure. It assumes an invariable schema in usage: “*huashuo* + main clause”, in which only the position of *huashuo* in relation to the main clause may vary.

Sometimes due to comma punctuation which indicates a small pause in real speech, *huashuo* may appear to be part of a larger syntactic structure involving an X-element, for example in (138) and (139).

- (138) 话说 我的 红烧汁，
Huàshuō wǒ de hóngshāo zhī,
speaking of I-POSS red braise sauce
- 炖 过 一次 鸡腿 了，
dùn guò yí cì jītuǐ le,
braise PRF one time chicken leg PRF
- 可以 留着 再 炖 别的 东西 吗？
kěyǐ liú zhe zǎi dùn bié de dōngxi ma?
can keep-PROG again braise other-ADJ thing Q

“*Speaking of my red braise sauce, has been used once already to cook chicken legs, can it be kept to braise other things again?”

“Speaking of, my red braise sauce has been used once already to cook chicken legs, can it be kept to braise other things again?”

(2138162592 - Archive (2)(new).txt, 2011, Chat)

(139) 话说 你 睡 不 着 觉,

Huàshuō nǐ shuì bù zháo jiào,

speaking of you sleep NEG COMPL sleep

绝对 还是 有 咖啡 的 原因

juéduì hái shì yǒu kāfēi de yuányīn

absolutely still COP have coffee POSS reason

“*Speaking of you can't fall asleep, must be due to coffee.”

“Speaking of, (the fact that) you can't fall asleep must be due to coffee.”

(3332142874 - Archive (2)(new).txt, 2010, Chat)

In these examples, “speaking of my red braise sauce [话说我的红烧汁]” and “speaking of you can't fall asleep [话说你睡不着觉]” are separated from the rest of the utterances by comma punctuation. It thus seems as if the *huashuo* construction entailed a “huashuo X” structure, rendering the interpretations of the overall sentence structure to be “(huashuo + NP) + sentence” and “(huashuo + sentence) + sentence” in these cases. But it is clear that these alleged X-elements are actually an integral part of the post-*huashuo* segment of the utterances, both syntactically and semantically. Taking the “huashuo X”

chunk out of the utterances would leave the rest of the sentences syntactically fragmental and propositionally incomplete.

4.2.3 Loose syntactic connection

As a discourse marker, *huashuo* is not an integral part of the proposition in the utterance it introduces. In real speech situations, this quality is often manifested in a small pause between *huashuo* and the rest of the utterance. In the CMC setting, the representation of a pause is realized through different strategies of syntactic detachment, i.e. the use of punctuation (comma/ellipsis) or a space between *huashuo* and the rest of the utterance, or a change of line (thread separation), as exemplified in (140) to (143).

- (140) 话说, 你 在 家 了?
Huàshuō, nǐ zài jiā le?
speaking of you LOC home PRF

“Speaking of, are you home?”

(2912576732.xml, 2007, Chat)

- (141) 话说…… SJ 终于 要 出 2 辑 啦
Hu àshuō... SJ zhōngyú yào chū èr jí la
so SJ finally will release 2nd album PRF-MOD

“So... SJ is finally going to release their 2nd Album!”

(3981290848(backup1).txt, 2007, Chat)

- (142) 话说 [space] 这 人 什么 口音? ?
Hu àshuō [space] zhè rén shénme kǒuyīn?
 so this person what accent
 “So [space] *which accent does this person have?*”
 (3332142874(new).txt, 2010, Chat)

- (143) 1:06:46 话说
Hu àshuō
 you know
 1:07:00 你 脸 是 见 圆
 Nǐ liǎn shì jiàn yuán
 you face indeed seem round
 “You know, *your face does look rounder.*”
 (3332142874 - Archive(new).txt, 2009, Chat)

These syntactically detached cases take up 18.5% of all the discourse marker uses of *huashuo* in the Chat dataset. As CMC data only partially reflect prosodic features of real speech, the observations made in this section thus only confirm the syntactic and propositional non-integration properties of *huashuo* as a discourse marker, but do not fully reflect the formal or prosodic preferences of *huashuo* in general. Further studies of spoken data need to be done in order to better understand that aspect.

4.3 Functional aspects

The *huashuo* construction is polyfunctional as a discourse marker. The Chat data show that it can be used as a topic-introducing device, a self-repair tool, and a politeness marker, as exemplified respectively in (144) to (146).

- (144) 这回 我 在 市里, 有 问题 打 电话。
zhèhuí wǒ zài shìlǐ, yǒu wèntí dǎ diànhuà
this time I LOC city in have problem make phone call

话说 你 得 买 个 德国 手机 卡。
Huàshuō nǐ děi mǎi ge déguó shǒujī kǎ.
speaking of which you must buy CLF German cellphone card

“I’m in town this time. If (you) have any problem, call me. Speaking of which, you need to buy a German SIM card.”

(Chat History between K and W 2013-08-11.txt, 2013, Chat)

- (145) 以后 咱家的 男生 都是 胖胖的 了,
Yǐhòu zánjiā de nánshēng dōu shì pàngpàng de le,
later our family-POSS boy all COP chubby-ADJ PRF

话说 M 是 例外, 还 挺 极端
huàshuō M shì lìwài, hái tǐng jíduān
that said M COP exception even quite extreme

“All the guys in our family will be chubby ones from now on, that said, M is an exception, quite extremely actually.”

(2138162592 - Archive (2)(new).txt, 2011, Chat)

(146) 我 忽然 看到 你 发 我
 Wǒ hūrán kàn dào nǐ fā wǒ
 I suddenly see-COMPL you send me

周日 两点 见
 zhōurì liǎngdiǎn jiàn
 Sunday two o'clock see

话说 是 打错 了 吧……

Huàshuō shì dǎ cuò le ba...

I mean COP type wrong PST MOD

“I just saw that you sent me: see you on Sunday at 2 o'clock. **I mean,**
 (it) was a mistake right?”

(Chat History between S and W 2013-08-16.txt, 2013, Chat)

The frequency distribution of these functions is shown in Figure 22. It is clear that topic introduction is the most dominant discourse marking function of *huashuo*.

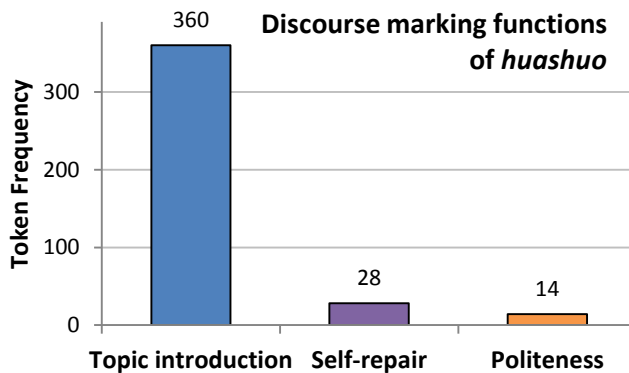


Figure 22 Frequency distribution of the discourse marking functions of *huashuo* in the Chat dataset (sum = 402)

In the following sections, I will discuss each of the discourse marking functions in detail (Section 4.3.1 – 4.3.2) and examine the usage patterns, including peripheral positions (Section 4.3.3) and dialogual-monologual uses (Section 4.3.4). Each of the functional analyses will be supported by corpus-based frequency distributional analyses.

4.3.1 Topic-introducing functions

In existing studies it is generally accepted that *huashuo* has topic-introducing functions. But different scholars make different observations within this understanding. For example, Shi [施] (2013) only talks about the topic change function of *huashuo* in comparison to its usage as an SV-structured reporting clause. Zhou [周] (2012) identifies an additional topic initiation function besides the topic change function, incorporating global positional differences into the functional analyses. Zhang [张] (2016) supports the distinction between topic initiation and topic change functions, but her definition of topic initiation equals topicalization and it is based on the understanding of the “*huashuo* X” structure of the construction (see counterargument towards this understanding in Section 4.2.2). Using the term “information change marker”, Lu [鲁] (2016) merges the concepts of topic initiation and topic change and argues that *huashuo* either signals the entrance of new information or adds explanations to old information. Besides introducing new information into the discourse, Chen [陈] and Huang [黄] (2017) further observe that

huashuo can also indicate that the new topic is propositionally related or even subordinated to the previous one.

In this thesis I incorporate the findings from the existing studies and observations from my own corpus data and distinguish the following topic-introducing functions of *huashuo*: **topic expansion**, **topic change**, and **topic initiation** (see Section 2.4.2 for the definition and classification of the concepts).

4.3.1.1 Topic expansion and topic change

Like the SPOX construction, *huashuo* can be used to mark both minor and major topical movements in discourse, i.e. *topic expansion* and *topic change* events, as exemplified in (147) and (148) respectively. In example (147), Y uses *huashuo* to bring up another point in their trip planning. The overall topic of the “Italian trip” stays unchanged with this operation. By contrast, in example (148), the use of *huashuo* marks a topical movement from “publishing academic papers” to “mitten crabs”, which is a rather drastic change of topic.

(147) (G and Y are planning a trip together to Italy.)

G: 我 知道 意大利 你 都 没去 那几个 城市,
 Wǒ zhīdào yìdàlì nǐ dōu méiqù nà jǐ ge chéngshì
 I know Italy you all NEG go those CLF city
 对吧
 duìba
 right-MOD

Y: 对滴, 话说 你 想 第一天
 Du ìdi, huàshuō nǐ xiǎng dì yī tiān
 right-MOD **speaking of** you want first day
 睡 在 米兰, 还是 去 维罗纳?
 shuì zài mǐlán, hái shì qù wéi luó nà?
 sleep LOC Milan or go Verona

“G: I know you haven’t been to those cities in Italy, right?”

Y: Right, **speaking of**, do you want to spend the night in Milan or go to Verona on the first day?”

(1581666176(new).txt, 2012, Chat)

(148) [D and W are talking about publishing academic papers.]

W: 英语 我 来 发 第一篇 吧
 Yīngyǔ wǒ lái fā dì yī piān ba
 English I come publish first CLF MOD
 D: 哦 第一个 吃 螃蟹 的 人
 Ōu dì yī gè chī pángxiè de rén
 oh first CLF eat crab ADJ person
 W: 话说 现在 是 大闸蟹 季节 呀
Huàshuō xiànzài shì dàzhàxiè jìjié ya
speaking of which now COP mitten crab season MOD

“W: Let me be the first to publish a paper (on this topic) in English then.

D: Oh, ‘the first person to eat the crab’²⁶.

W: **Speaking of which**, now it’s mitten crab season.”

(Chat History between D and W 2019-09-30.txt, 2019, Chat)

²⁶ It is a Chinese proverb that describes a brave and enterprising person.

In both of the cases, regardless of the different degrees of topical divergence, there is a clear textual and/or conceptual link between the new utterance and the previous discourse. In example (147), Y's new utterance is linked to "Italian cities" in G's previous speech; and in example (148), the new "mitten crab" topic is linked to D's prior use of the Chinese proverb "the first person to eat the crab". The use of *huashuo* not only informs the hearer about the (sub)topical movement, but also hints at the existence of a link between the new utterance and the discursive pretext and indicates that the new utterance is motivated by a certain piece of information contained in the previous discourse. By pointing at such a link, *huashuo* thus helps the hearer to better make sense of the transition and thereby facilitates the perception of discourse coherence.

Interestingly enough, my data also show that *huashuo* can as well mark topic change events which exhibit no textual or conceptual connection to the discursive pretext. As shown in (149), P uses *huashuo* to shift from talking about Z working too hard to telling him the news of a common friend J getting married. And in (150), Y uses *huashuo* to switch to ask G about her toothache in a conversation that was originally about flight booking. The new topics introduced by *huashuo* in these cases are not related to the previous conversations in any way.

(149) (Z is complaining to P about how busy he has been.)

P: 天哪, 你 咋 就 没 个 帮手 呢
 Tiān na, nǐ zǎ jiù méi gè bāngshǒu ne
 sky MOD you how just NEG have CLF helper Q

4 The huashuo construction

Z: 我 也 没 办法, 节省 成本 么
 Wǒ yě méi bànfa, jiéshěng chéngběn me
 I also NEG have solution save cost MOD

P: 唉唉, 你 也 真 是 太 强 大 了
 Ài ài, nǐ yě zhēn shì tài qiángdà le
 alas you INTS really COP too strong PRF

P: 话说 你 听说 J 要 结 婚 了 么?
Huàshuō nǐ tīngshuō J yào jiéhūn le me?
by the way you hear J will marry PRF Q

Z: 不 知 道 呀。 我 现 在 啥 也 不 知 道
 Bù zhīdào ya. Wǒ xiànzài shá yě bù zhīdào
 NEG know MOD I now what also NEG know
 呵呵, 啥时候 呀, [...]
 Hehe, sháshíhòu ya, [...]
 hehe what time Q

“P: *Gosh, why didn't you have anyone to help?*

Z: *Well I couldn't, to save costs you know.*

P: *Sigh, you are just incredible. **By the way**, have you heard that J is getting married?*

Z: *No! I don't know about anything these days! When? [...]*”

(528045637_new.txt, 2011, Chat)

(150) [G and Y are talking about flight tickets and baggage allowance.]

G: 我 觉 得 这 要 是 我 估 计 能 用 满……
 Wǒ juéde zhè yào shì wǒ gūjì néng yòng mǎn...
 I feel this if COP I estimate can use full

看 我的 包 就 知道 了…… 哈哈
 Kàn wǒ de bāo jiù zhīdào le... haha
 look I POSS purse CNJ know PRF haha

Y: 话说 你 牙 怎么 疼 了?
Huàshuō nǐ yá zěnmē tēng le?
 so you tooth how hurt PRF

“G: *I think if it were me I could probably use up (the entire quota)...*
Just look at my purse and (you) will see... haha

Y: So *how come you are having a toothache?*”

(2650361431(new).txt, 2012, Chat)

In these cases, the use of *huashuo* does not link the new utterance to the discursive pretext. It does not imply that the new utterance is motivated by anything contained in the previous discourse. Instead, it points at the existence of a pragmatic pretext that is established in the interlocutor relationship. It conveys the information that the new topic has a pending status on the speaker's interactional agenda and that it is based on information that already exists in the shared knowledge of the interlocutors. In example (149), the use of *huashuo* informs Z that “J getting married” is a topic that P has been meaning to talk about. And at the same time, it highlights the understanding that based on the nature and status of their relationship, bringing up updates about J should not be considered as completely unexpected or ungrounded. Similarly, in example (150), the use of *huashuo* not only informs G of the radical topical movement but also hints at the relationship status and the shared knowledge between G and Y, which justifies the relevance of the new topic.

In brief, when marking topic expansion and topic change events, *huashuo* links the new utterance to pre-existing information, which can either be linguistically encoded in the discursive pretext or pragmatically established in the interlocutor relationship. By establishing such a link, *huashuo* justifies the topical relevance of the new utterance and enhances the perception of discourse coherence.

4.3.1.2 Topic initiation

Besides topic expansion and topic change functions, *huashuo* is also observed in marking *topic initiation* events, i.e. introducing a brand new topic into the conversation without pre-existing topics to relate to or depart from, which often occurs at the very beginning of a new discourse event. For example in (151) and (152), *huashuo* is used to initiate a new topic and a whole new chat session at the same time.

- (151) 话说 我 昨天 梦见 你 了
Huàshuō wǒ zuótiān mèng jiàn nǐ le
so I yesterday dream-COMPL you PST

“So I dreamt about you yesterday.”

(Chat History between G and W 2018-02-03.txt, 2018, Chat)

- (152) 话说
Huàshuō
you know²⁷

²⁷ See Schiffrin (1987: 285–290) for the speech-opening variant of the discourse marker *you know*.

好 怀念 大家 一起 玩 三国杀 的 日子 啊
 hǎo huáiniàn dàjiā yìqǐ wán sānguóshā de rìzi a
 very miss all together play Sanguosha ADJ day MOD

“***You know***, (I) really miss the days when we were playing Sanguosha together.”

(Chat History of the WeChat-Group 2014-12-04.txt, 2014, Chat)

In topic initiation events, the use of *huashuo* does not intend to justify the topical relevance of the new utterance by linking it to any discursive pretext, as there is none. Instead, it relies again on the existence of a pragmatic pretext that contains the information about the pending status of the introduced topic and the interlocutors’ shared knowledge. In more elaborate words, when using *huashuo* to open a speech and/or a new chat session, the speaker indicates that the new utterance is motivated by the pending status of the topic on their interactional agenda, which practically directs the attention of the hearer to a non-verbalized, assumed discursive pretext, as if the newly introduced topic were situated in a continuous discourse structure, so that it may be perceived as less isolated and less sudden. At the same time, the use of *huashuo* also informs the hearer that the new utterance contains certain information which should be already known to them, so the new topic should not be taken as completely unanticipated.

In sum, the topic-introducing *huashuo* facilitates the perception of discourse coherence by pointing at a link between the new utterance and some pre-existing information. In topic expansion events, the link is established either textually or conceptually between the new utterance and the discursive pretext. In topic change events,

the connection can be found either in the discursive pretext or in the pragmatic pretext that exists in the interlocutor relationship. The topic initiation function points only at the pragmatic pretext.

4.3.1.3 Frequency distribution

According to the results of the frequency distributional analyses of the topic-introducing *huashuo* in the Chat dataset (Figure 23), it is most frequently used as a topic-changing device (53.61%), less often in marking topic expansions (38.33%), and least often in initiating a brand new topic (8.06%).

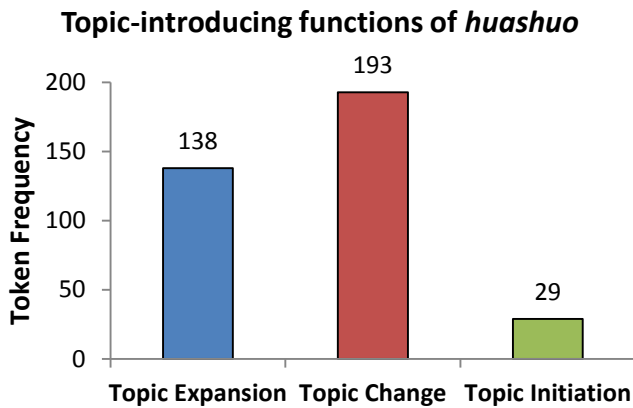


Figure 23 Frequency distribution of the topic-introducing functions of *huashuo* in the Chat dataset (sum = 360)

In terms of the different linking mechanisms of the topic-introducing functions, *huashuo* appears to be used more often in linking the new utterance to the discursive pretext, i.e. linking

(sub)topics on a textual or conceptual level, than linking the new utterance to a non-verbalized pragmatic pretext (see Figure 24).

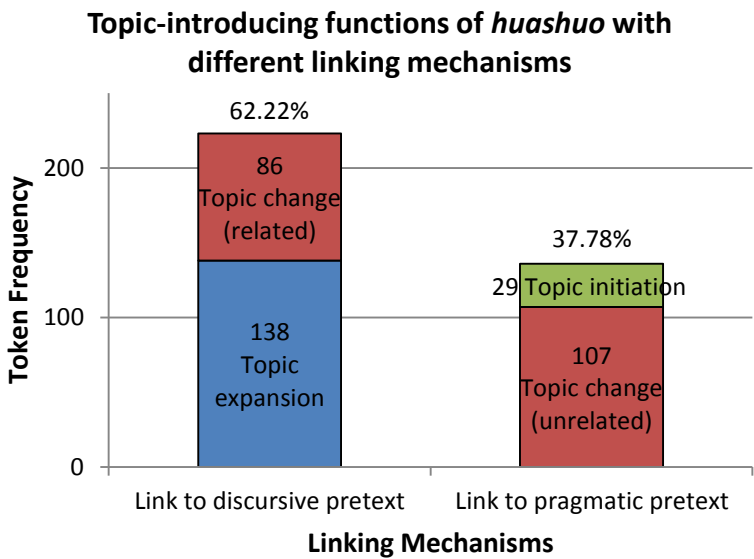


Figure 24 Frequency distribution of the topic-introducing functions of *huashuo* with different linking mechanisms in the Chat dataset (sum = 360)

In Figure 24, two types of topic change function are distinguished: “Topic change (related)” refers to cases in which the change of topic is motivated by a textual or conceptual connection to the discursive pretext, while “Topic change (unrelated)” refers to cases in which the transition is made between more distant and unrelated topics. The topic expansion function and the “related” type of the topic change function together represent the usage of *huashuo* in creating a link to the discursive pretext when introducing a new

(sub)topic into the conversation. This kind of usage appears to be more frequent than the other kind that entails linking the new utterance to a pragmatic pretext, as represented by the topic initiation function and the “unrelated” type of the topic change function.

4.3.2 Interpersonal functions

Occasionally, the use of *huashuo* does not concern topical relevance or discourse coherence, but rather the interlocutor relationship. There have been a few studies that observe the interpersonal functions of *huashuo*. For example, Shi [施] (2013) observes that *huashuo* conveys the speaker’s attitude in a mitigated way in the specific usage pattern “*huashuo* + sentence + *ne*”. Zhang [张] (2016) talks about “hesitation and negotiation” functions in using *huashuo* to raise a (rhetoric) question or request in the hope of a positive response and “reflection and specification” functions when the use of *huashuo* precedes or follows ellipsis that indicates a longer pause in speech. Comparing the use of *huashuo* in different sentence types (i.e. declarative, interrogative, and exclamatory sentences), Zhou [周] (2012) believes that the main interpersonal function of *huashuo* is “tone moderation” and that is why the *huashuo* construction is rarely used in exclamatory contexts. However, Lu [鲁] (2016) argues that the interpersonal functions of *huashuo* rather feature polarity: it carries mitigating functions when used for negotiation and suggestion in declarative and interrogative sentences, but amplifying functions when used in exclamatory sentences to convey the speaker’s strong subjective opinion. In Chen [陈] and Huang [黄]’s (2017) study, the mitigating

functions of *huashuo* are understood within the framework of the politeness theory (Brown & Levinson 1987) as a face-preserving strategy.

Drawing on findings in the existing studies and my own observations made in this thesis on the basis of the Chat dataset, I identify two interpersonal functions of *huashuo*: *self-repair* that involves elaboration or correction, and *politeness management*.

4.3.2.1 Self-repair

As a repair device, *huashuo* marks the addition of elaborative or corrective information to the previous utterance. As shown in example (153), after making the claim about all the guys in their family being chubby, the speaker uses *huashuo* to bring in a new piece of information that updates the truth value of the previous utterance: the claim can only be considered true when disregarding the aspect mentioned in the new utterance. And in example (154), N uses *huashuo* to add elaborative information to his previous utterance and clarifies that it is not for “cool” reasons that he is still using an ancient cellphone. The use of *huashuo* reveals N’s concern that his comment comparing his ancient cellphone to a brick might not be clear enough in terms of the point he is trying to make.

- (153) 以后 咱家的 男生 都是 胖胖的 了,
 Yǐhòu zán jiā de nánshēng dōu shì pàngpàng de le
 later our family-POSS boy all COP chubby-ADJ PRF

话说 M 是 例外, 还 挺 极端
hu àshuō M shì lìwài, hái tǐng jíduān
 that said M COP exception even quite extreme

“All the guys in our family will be chubby ones from now on, that said, M is an exception, quite extremely actually.”

(2138162592 - Archive (2)(new).txt, 2011, Chat)

(154) N: [...] 同事 全 Iphone 我 一 诺基亚 3108
 [...] tóngshì quán Iphone wǒ yí nuǒjīyà 3108...
 [...] colleague all Iphone I one Nokia 3108

P: 多 有 个性
 Duō yǒu gèxìng
 how have character

N: 是..... 直接 拿 砖头 完了~~~ 更 有 个性~~~
 Shì... zhǐjiē ná zhuāntóu wán le, gèng yǒu gèxìng
 yes direct take brick done-PRF more have character

话说 我 就是 对 手机
Hu àshuō wǒ jiùshì duì shǒujī
 you know²⁸ I just COP to cellphone

这 东西 没 兴趣
 zhè dōngxi méi xìngqù
 this thing NEG have interest

“N: [...] colleagues all (have) iPhones, while I (hold) a Nokia 3108...”

P: That is very cool.

²⁸ See Müller (2005: 164–167) for account of the explanatory discourse marking functions of *you know*.

N: *Yea right... then a brick would also do, even cooler. You know, I'm just not into cellphones."*

(3332142874(new).txt, 2010, Chat)

The fundamental aspect that distinguishes the self-repair *huashuo* from the topic-introducing *huashuo* is the absence of topical movements between the utterances. The elaborative or corrective new utterance does not represent any transition from the previous topic to the next. Instead, it contains information that updates the truth value of the previous utterance, which helps the hearer to achieve a better understanding of the speaker's intended meaning.

4.3.2.2 Politeness management

The politeness marker usage of *huashuo* is identified when the introduced utterance does not bring about a change of topicality or update the truth-condition of the previous utterance. In such cases, the introduced utterance is often a comment, a question, or a request, which could be interpreted as intrusive by the hearer. For example in (155), *huashuo* is used by S to preface her comment pointing out a mistake made by W. It helps to make the accusation sound casual and the subject matter less serious. In (156), *huashuo* is used to politely ask for the name of the other person. It dresses the utterance with an apologetic tone as if that question had been overdue. In the group chat (157), L uses *huashuo* in putting forward a request to be invited to an activity. It softens the tones of the request, which otherwise might be perceived as too direct or blunt.

(155) 我 忽然 看到 你 发 我

Wǒ hūrán kàn dào nǐ fā wǒ

I suddenly see-COMPL you send me

周日 两点 见

zhōuri liǎngdiǎn jiàn

Sunday two o'clock see

话说 是 打错 了 吧……

Huàshuō shì dǎ cuò le ba...

I mean COP type wrong PST MOD

“I just saw that you sent me: see you on Sunday at 2 o'clock. I mean, (it) was a mistake right?”

(Chat History between S and W 2013-08-16.txt, 2013, Chat)

(156) 学姐²⁹ 好!!

Xuéjiě hǎo!

school sister good

话说 还 不知道 你的 名字

Huàshuō hái bù zhīdào nǐ de míngzì

you know³⁰ still NEG know you-POSS name

“M: Hi Sis! You know, (I) still don't know your name.”

(Chat History between M and W 2016-10-19.txt, 2016, Chat)

²⁹ “学姐 [xuéjiě]” is a Chinese honorific that is used by younger/junior female students to address older/senior female students in any educational system.

³⁰ See Östman (1981: 19–21) for the hedging function of *you know* for politeness management.

(157) H: 劳动 一天 了, 感觉 很 快乐。

Láo dòng yì tiān le, gǎnjué hěn kuàilè

labor work one day PRF feel very happy

L: 话说 叫上 我们 一起 啊

Huàshuō jiào shàng wǒmen yìqǐ a

I mean call-COMPL we together MOD

“H: (I) did labor work the whole day, feeling very happy.

L: **I mean**, you should count us in!”

(Chat History of the Wechat Group 2015-03-29.txt, 2015, Chat)

To sum up, the interpersonal functions of *huashuo* concern the interlocutor relationship rather than topical relevance and discourse coherence. As a repair tool, it guides the hearer towards a more accurate understanding of the speaker’s intended meaning; and as a politeness marker, it softens the potential impact of the new utterance on the hearer.

4.3.3 Peripheral positions

As defined in Section 2.4.4, the term *periphery* in this thesis refers to the margin of an argument structure in a sentence. The *left periphery* (LP) and the *right periphery* (RP) refer to the beginning and the end of an argument structure and the *medial periphery* (MP) refers to the parenthetical position inside an argument structure. The usage of *huashuo* can be found in all three peripheral positions, as exemplified in (158) to (160).

- (158) 话说 你的 预产期 是 啥时候?
Huàshuō nǐ de yùchǎnqī shì shá shí hòu?
 by the way you-POSS due day COP what time

“*By the way when is your due day?*”

(2339984085.xml, 2011, Chat)

- (159) 今晚上 话说 你 有空 吗?
 Jīn wǎnshàng huàshuō nǐ yǒu kòng ma?
 today evening speaking of you have time Q

“*Are you free, speaking of, tonight?*”

(3071266691(new).txt, 2011, Chat)

- (160) 你 想不想 要 蜂蜜 话说?
 Nǐ xiǎng bù xiǎng yào fēngmì huàshuō?
 you will NEG will want honey speaking of which

“*Would you like to have some honey, speaking of which?*”

(Chat History between Q and W 2014-01-18.txt, 2014, Chat)

From the frequency distributional prospective, the LP position appears to be the most dominant position for the use of *huashuo* in both of the topic-introducing and interpersonal functional domains (see Figure 25). As a topic-introducing device, *huashuo* can occur in all of the peripheral positions, but the uses in the MP and RP positions are much rarer in comparison to those in the LP position. In the MP position, the interpersonal functions of *huashuo* seem to be more

dominant than the topic-introducing functions, while in the RP position they do not seem to occur at all³¹.

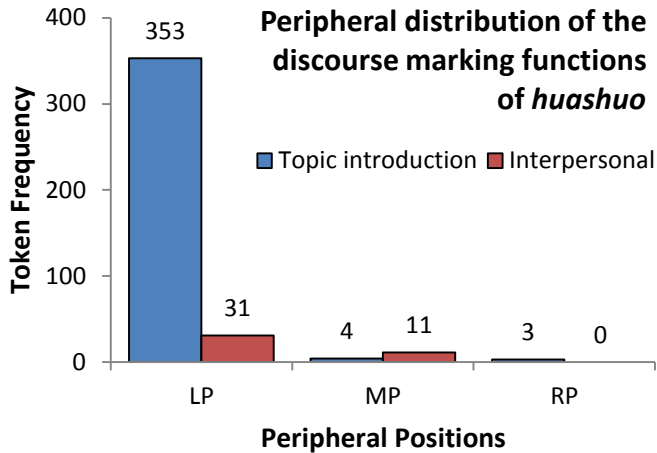


Figure 25 Frequency distribution of the discourse marking functions of *huashuo* in different peripheral positions in the Chat dataset (sum = 402)

In terms of functional asymmetry, *huashuo* thus represents a counterexample of the hypothesis. The interpersonal functions of *huashuo* that exhibit stronger intersubjectivity, i.e. showing more attention to the hearer, do not seem to favor the MP and RP positions more than the topic-introducing functions do. Both types of interpersonal function prevail in the LP position.

³¹ These observations, however, should only be taken as tentative for the time being, as the overall frequencies of these peripheral uses are too low for any conclusive remarks to be made at this point.

The functional asymmetry hypothesis thus appears to hold better for certain functional domains than for others. It does not seem to apply to the repair and politeness functions of *huashuo*.

4.3.4 Dialogual and monologual uses

The Chat data show that the use of *huashuo* can either be found in a *dialogual* context, i.e. in a turn-taking operation in which a different speaker initiates a new utterance/turn, or in a *monologual* context in which the same speaker continues with their own course of talk (see Section 2.4.5 for the definition of the terms). These two usage patterns are exemplified in (161) and (162) respectively.

(161) (K and P are catching up after not hearing from each other for quite some time.)

K: [...] 我们 计划 下月 下旬
 [...] wǒmén j hu à xi àyu è xi àx ún
 [...] we plan next month second half
 去 苏黎世 看 费德勒!
 qù Sūlīshì k àn Fěidēl ě
 go Zurich see Federer

P: 哈哈 话说 你 现在 在 哪里 啊? [...]
 Haha huàshuō nǐ xi ànz ài z ài nǎlǐ a? [...]
 haha **speaking of** you now LOC where Q

“K: [...] We plan to go to Zurich in the second half of next month to watch Federer (play)!”

P: Haha, speaking of, where are you right now? [...]”

(054093473722(new).txt, 2010, Chat)

(162) N: 周日 你 干 啥 去?

Zhōurì nǐ gàn shá qù?

Sunday you do what go

P: 可能, 跟 我妈 去 北京, 但是 还 不一定 了 [...]

Kěnéng, gēn wǒ mā qù běijīng, dànshì hái bù yí dìng le [...]

maybe with my mum go Beijing but still NEG sure PRF

哦, 另外, 明天 我 不 找 M 了。Z B 来

Òu, lìngwài, míngtiān wǒ bù zhǎo M le. Z B lái

oh moreover tomorrow I NEG find M PRF Z B come

话说, 你 要不要 也 来 呢?

Huàshuō, nǐ yào bú yào yě lái ne?

speaking of you want NEG want also come Q

“N: What are you doing on Sunday?

P: Maybe, go to Beijing with my mum, but it's not sure yet. [...] Oh and by the way, tomorrow I'm not going to meet up with M anymore. Z and B are coming. **Speaking of**, would you like to come too?”

(3332142874 - Archive (2)(new).txt, 2010, Chat)

According to the Chat dataset, the topic-introducing *huashuo* is used more often in monologual contexts than in dialogual contexts. In Figure 26 we can see that, both in topic expansion and topic change events, there are more monologual uses of *huashuo* than dialogual uses. The “not applicable (N/A)” cases refer to the topic initiation events, in which the turn-taking operation is not relevant.

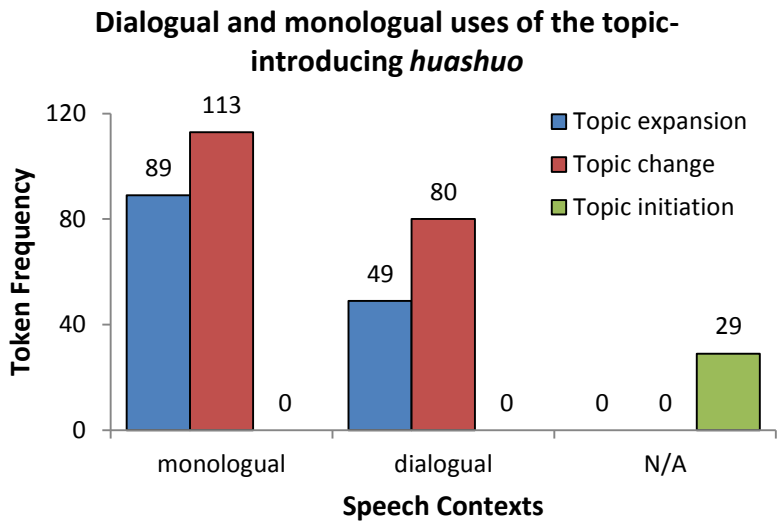


Figure 26 Frequency distribution of the topic-introducing functions of *huashuo* in dialogical and monological contexts in the Chat dataset (sum = 360)

As to the interpersonal functions (Figure 27), *huashuo* appears to be used equally frequently in both dialogual and monological contexts as a politeness marker; as a self-repair device, by contrast, since *huashuo* is meant to add new information that modifies the truth-condition of the speaker’s own previous utterance, it is observed more often in monological contexts. The few dialogual uses of the self-repair *huashuo* represent cases involving disrupted adjacency of turns between the interlocutors, which is a common phenomenon in the CMC setting (See Section 4.1.1 for more discussion on the characteristics of the CMC data).

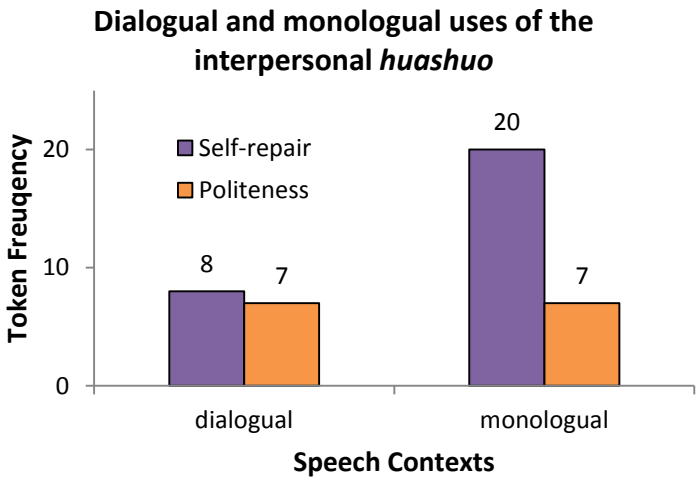


Figure 27 Frequency distribution of the interpersonal functions of *huashuo* in dialogual and monological contexts in the Chat dataset (sum = 42)

The unique characteristics of CMC also bring about the possibility of having multiple simultaneous threads and turns in a conversation. As the interlocutors lack the paralinguistic cues to coordinate their activities, an interlocutor might move on with their speech before they receive the other interlocutor’s response, and they might introduce a new topic into the ongoing conversation before the previous one is finished. As shown in example (163), D uses *huashuo* to bring up another topic about J’s medical exam before the previous “boat tour” topic is finished. It therefore results in the existence of two simultaneous threads of talk from that point on.

(163) W 18:16: 怎么? 想 带 我 游 珠江?
 Zěnmē? Xiǎng dǎi wǒ yóu zhūjiāng?
 how want take me tour Pearl River

D 18:17: 对, 夜游 珠江,
 Duì yèyóu zhūjiāng,
 right night tour Pearl River
 游完 再 去 吃 夜宵
 yóu wán zài qù chī yèxiǎo
 tour over then go eat late night food

D 18:18: 话说 J 报告 有出来 吗?
Huàshuō J bàogào yǒu chūlái ma?
by the way J report PRF-come out Q

W 18:19: 哇哦, 听起来 太 棒 啦
 Wā'òu, tīng qǐlái tài bàng la
 wow sound too great MOD

W 18:20: 出来了! 没有 病!
 Chūlái le! Méiyǒu bìng!
 come out-PRF no sickness

“W 18:16: *Why? Wanna take me on a boat tour on the Pearl River?*

D 18:17: *Yes, a boat tour at night on the Pearl River. After that we will then go and have some late night food.*

D 18:18: ***By the way**, is J's report out yet?*

W 18:19: *Wow, sounds amazing!*

W 18:20: *Yes! (She is) not sick!*”

(Chat history between D and W 2018-10-05.txt, 2018, Chat)

Due to this characteristic, the frequent monologal uses of *huashuo* should be understood as genre-specific at this stage. In a real-time face-to-face conversation, in which the interlocutors can better coordinate their conversations, the monologal uses of *huashuo* might be fewer in occurrence than as represented in this study. To better understand this aspect calls for further studies with a different type of data.

4.4 Diachronic changes

4.4.1 Grammatical constructionalization

As presented in Section 4.2.1, owing to different semantic interpretations of the phrase, *huashuo* can assume quite different roles in a sentence, including **verb phrase (VP)**, **noun phrase (NP)**, **subject-verb clause (SV)**, and **discourse marker (DM)**. The diachronic frequency distribution of these different uses is shown in Figure 28, on the basis of the CCL and Chat datasets combined. The X-axis represents the time periods from the ancient times to the modern days. The segmentation follows the original time periods provided in the CCL corpus, which is based on the timeline of the Chinese dynasties until the modern times. Each time period holds a different total number of words. The token frequency represented by the Y-axis is thus normalized per one million words. The color-coded data points represent the different uses of *huashuo*.

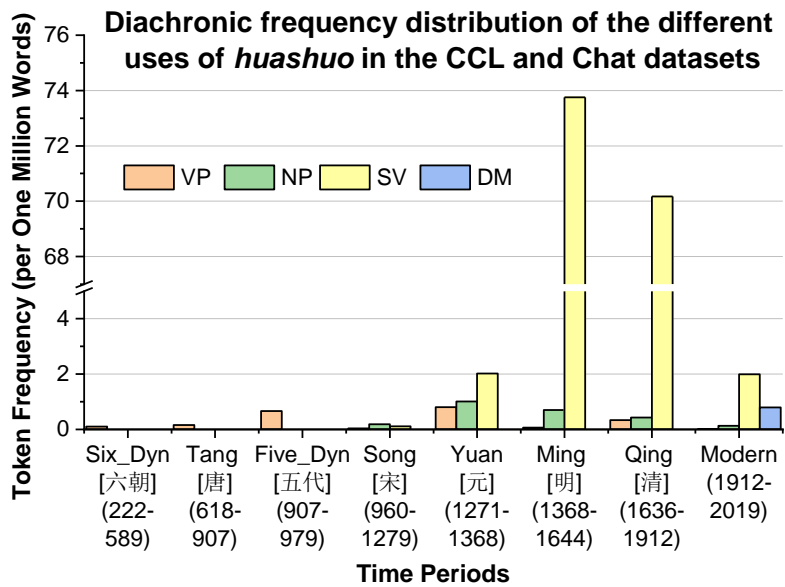


Figure 28 Diachronic frequency distribution of the different uses of *huashuo* in the CCL and Chat datasets (sum = 5069) (the Y-axis has a break between 5 and 67 for better visual presentation)

We can see from the diachronic frequency distribution that *huashuo* was first used as a verb phrase and a noun phrase in history, but those uses were never frequent. It only started to gain frequency in the Ming Dynasty around the 1300s when used as a subject-verb clause. As observed in a number of existing studies as well (i.e. Zhou [周] 2012; Shi [施] 2013; Lu [鲁] 2016; Chen [陈] & Huang [黄] 2017), the dominant SV-clausal usage of *huashuo* seems to be the starting point of its grammatical constructionalization process towards a discourse marker (see Section 4.4.4 for more discussion about the

influence of the SV-clausal usage on the emergent discourse marker usage of *huashuo*).

The SV-clausal use of *huashuo* originated in the late Song Dynasty around the 1200s. At that time, oral narrations as a performative art form (i.e. Shuohua [說話]) and the popular literary genre (i.e. Huaben [話本]) had some quite distinct formats and styles. Among others, it was a common practice to use *huashuo* to start the narrative, as exemplified in (164) and (165). In this usage, *huashuo* conveys the meaning of ‘story tells / talks about’, which can be understood as a structurally reduced yet semantically sufficient variant of the reporting clause “cǐ/zhè běn hu à shuō [此/这本话说] – ‘This piece of story tells / talks about that’” (Sun [孙] 1953: 40).

(164) 话说 本地 有 一 王妈妈，

Huàshuō bēndì yǒu yí wáng māma,

story tell local have one Wang mother

与 二边 说合， [...]

yǔ èrbiān shuō hé, [...]

with two sides mediate

“*The story tells that there was a woman here named Wang, (she) mediated between two families (about matchmaking), [...]*”

(Huàběn [话本], 1200s, CCL_Ancient)

(165) 话说 国朝 嘉靖 年间，

Huàshuō guócháo jiājīng niánjiān,

story tell state dynasty Jiajing year period

圣人	在位，	风调	雨顺，	国泰	民安。
shèngrén	zài wèi,	fēng tiáo	yǔ shùn,	guó táì	mín ān.
divine person	LOC seat	wind suited	rain obedient	state fit	people safe

*“**The story tells that** during the years of Jiājīng, with a divine sovereign in office, the whole country was blessed with good weather (and hence bountiful harvests), prosperity and security.”*

(Huàběn [话本], 1300s, CCL_Ancient)

The SV-clausal usage of *huashuo* prevailed for centuries. Over time, it started to exhibit signs of development. Originally, the SV-clausal *huashuo* was restricted in the *chapter-initial* position, i.e. the beginning of a chapter/book and the starting point of a story, as is the case in (164) and (165). However, the corpus data show that, during the course of the Ming Dynasty (i.e. 1300s – 1600s), *huashuo* started to appear in non-chapter-initial positions as well, i.e. *paragraph-initial* and *sentence-initial* positions³² (see Figure 29).

The *paragraph-initial* position refers to the beginning of a new paragraph inside a chapter, and the *sentence-initial* position refers to the beginning of a sentence/clause inside a paragraph. Neither of the non-chapter-initial positions corresponds to the beginning of a new story. Consequently, *huashuo* in these positions is not used to start a new story, but to carry out scene-setting or scene-changing operations within the existing storyline.

³² See also Zhou [周] (2012: 501–502) for similar observations with a smaller set of data.

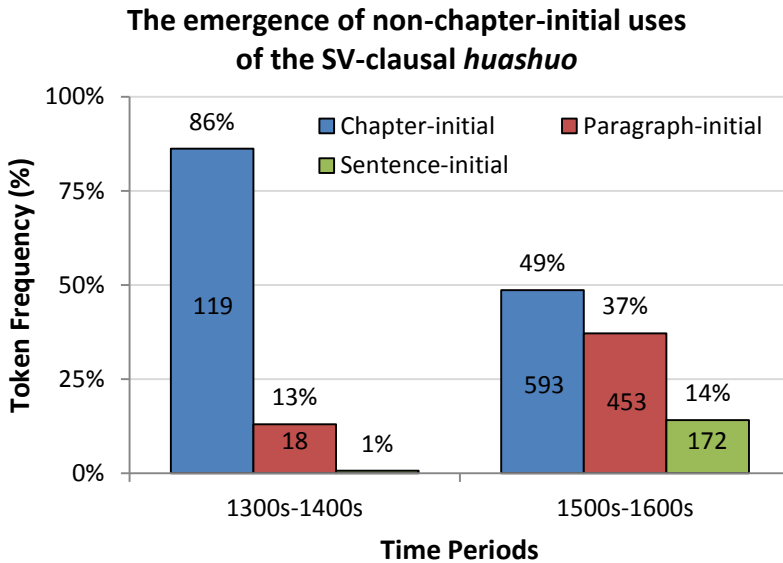


Figure 29 Emergence of non-chapter-initial uses of the SV-clausal *huashuo*

For example in (166), *huashuo* is used in the paragraph-initial position to switch scenes between different characters while the storyline continues; and in example (167), *huashuo* is used in the sentence-initial position to set the scene for a development in the storyline. In both examples, a literal interpretation of the SV-clausal *huashuo* as ‘story tells’ is no longer felicitous.

- (166) 玉姐 辞谢 先 回。
 Yù jiě cí xiè xiān huí
 Lady Yu refuse politely first return

亡八 又 请 众人 吃过 酒饭 方 散。 [...]
 Wángbā yòu qǐng zhòng rén chī guò jiǔfàn fāng sǎn. [...]
 Wangba further treat people eat-PRF feast only then part ways

话说 公子 在 路， 夜 住 晓 行，
Huàshuō gōngzǐ zài lù, yè zhù xiǎo xíng,
now Childe LOC road night stay morning go

不 数日， 来到 金陵 自家 门首 下马。
 Bú shù rì lái dào jīnlíng zì jiā mén shǒu xià mǎ
 NEG many days arrive-COMPL Jinling own home door front down horse

“Lady Yu politely refused [the invitation] and left first. Wangba then treated the rest to a feast before he could call it a day. [...]

***The story tells that / Now (we come to the scene where) the Childe was on his way, (he) rested during the night and traveled during the day, in just a few days’ time, (he) arrived at the front door of his home in Jinliang and got off his horse.”**

(Huàběn [话本], 1300s, CCL_Ancient)

(167) 那 高照 又 有些 妙处， [...]
 Nà gāozhào yòu yǒu xiē miào chù, [...]
 that tall lamp also have some wonder

风 越 大， 灯 越 明。
 fēng yuè dà dēng yuè míng.
 wind the more big lamp the more bright

话说 这个 灯 倒 不怕 风，
Huàshuō zhè gè dēng dào bú pà fēng,
now this CLF lamp by contrast NEG fear wind

只是 天上的 云 倒 有些 怕 风。
 zhǐ shì tiānshàng de yún dào yǒu xiē pà fēng.
 only sky above-ADJ cloud by contrast somewhat fear wind

“Those tall lamps are surprisingly good, [...] the stronger the wind, the brighter the light. *The story tells that / Now (you see), the lamps do not fear the wind, but the clouds in the sky seem to.”

(Sānbǎo Tǎijiān Xīyáng Jì [三宝太监西洋记], 1400s, CCL_Ancient)

The expansion of *huashuo* to non-canonical positions reflects its early development as a grammatical construction. It shows that the form of *huashuo* started to be paired up with functions which are no longer fully predictable from the semantic meaning of its constituents. In comparison to the original narrative-starting function in the chapter-initial position, which is based on the literal interpretation of *huashuo* as an SV-clause, the new functions of *huashuo* appear to be procedural in nature, setting boundaries between different scenes or narrative sequences, independent of the meaning of its lexical components.

But the emergence of the procedural usage did not replace the original SV-clausal usage of *huashuo*. It is in fact a common phenomenon for newly developed functions to co-exist with the original/older functions of a linguistic element or for some traces of the original lexical meaning to persist in the new construction during its grammatical constructionalization process (see the notions of *divergence* and *persistence* in grammaticalization theory in Section 2.3.1).

In most of the cases, the chapter-initial uses of *huashuo* can be interpreted in both ways, either as an SV reporting clause that is a propositional part of the sentence or as a conventionalized narrative-starting phrase that is rather a procedural element in the utterance. There are only a handful of cases³³, mostly in the Qing Dynasty (i.e. 1600s – 1800s), in which the procedural interpretation of *huashuo* in the chapter-initial position is non-negotiable. For example in (168), the semantic interpretation of *huashuo* as ‘the story talks about’ is not felicitous, as it would be syntactically and semantically redundant with “*cǐ shū nǎi* [此书乃] – ‘this story is’”.

- (168) **话说** 此 书 乃 青石山 一段 故事。
Huàshuō cǐ shū nǎi qīngshí shān yī duàn gùshì
 so this story COP Mt. Black Stone one CLF story

“**The story talks about* / *So* this story is one that happens in the Black Stone Mountain.”

(*Hú lǐ yuán Quánzhuàn* [狐狸缘全传], 1888, CN_Ancient)

Entering the Modern (1912 – 2019) time period, we observe the emergence of discourse marking functions of *huashuo*. In Figure 30, the X-axis zooms in on the Modern period with a 10-year interval and the Y-axis represents the relative token frequency of the different functions of *huashuo*, i.e. the percentage of each type of usage in relation to the total number of occurrences in this time period. We can

³³ Both occurrences in the CCL_Ancient and CN_Ancient datasets are counted due to the scarcity of such cases.

see that the persisting SV-clausal usage concentrates mostly in the beginning of the 20th century, which has been losing frequency drastically with time. The discourse marker usage, by contrast, has been gaining ground and frequency since the beginning of the 21st century.

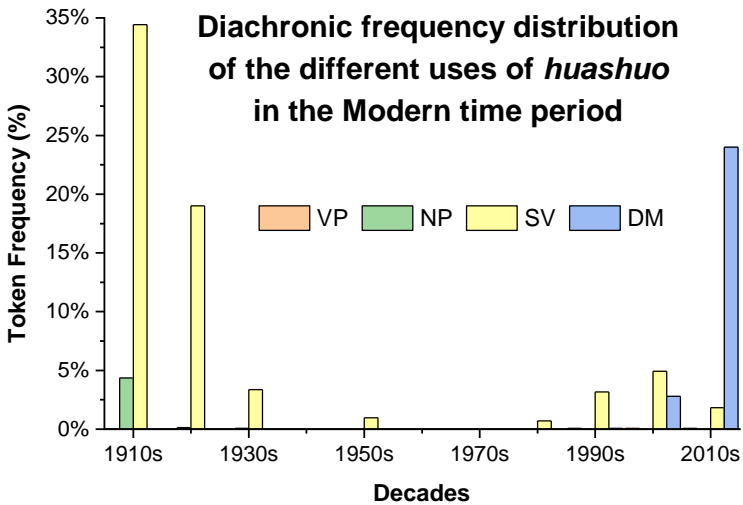


Figure 30 Diachronic relative frequency distribution of the different uses of *huashuo* in the Modern time period (1912 – 2019) in the CCL and Chat datasets

The emergence of discourse marking functions reflects further development of the *huashuo* construction. In succession to the development of procedural SV-clausal *huashuo*, the discourse marker *huashuo* exhibits new emergent features.

First of all, the use of *huashuo* is no longer restricted in the story-telling context. The topic-introducing *huashuo* can bring in subjects that are not necessarily (part of) a narrative but anything that is required by the communicative needs at the time of speech. And unlike the narrative-starting *huashuo*, which is used by the speaker solely to lay the groundwork for their own follow-up narration, the topic-introducing *huashuo* is often used to invoke the speech from the hearer.

Second, the topic-introducing *huashuo* has developed linking mechanisms. While the narrative-starting *huashuo* only concerns the following stretch of talk, the topic-introducing *huashuo* points at a link between the new utterance and the previous discourse or a pragmatic pretext, informing the hearer of the motivation behind the speech act.

Third, discourse marker *huashuo* has developed interpersonal functions. In the story-telling context, even though the narrative is addressed to an audience, the nature of the discourse is not interactive. The speech act is not performed to converse with the hearer on an interpersonal level. The narrative starter *huashuo* therefore does not concern interpersonal relations. However, as *huashuo* expanded to conversational contexts, in which discourse structuring involves interactions between the speaker and the hearer, it developed repair and politeness functions on the interpersonal level as well.

To sum up, the grammatical constructionalization process of *huashuo* started as a SV-structured reporting clause, which, by virtue of its semantic meaning, was frequently used as a customary phrase to start a story-telling event. With this usage becoming increasingly

established through time, the interpretation of *huashuo* became less dependent on the lexical meaning of the phrase but more conventionalized as a form-function pairing. More recently, *huashuo* developed new functions in discourse structuring, i.e. topic introduction and interpersonal functions, which further supports the validity and necessity of recognizing it as a grammatical construction and its versatile uses as different correspondences between form and functions.

4.4.2 Pragmaticalization

The development of the *huashuo* construction showcases a process of pragmaticalization, i.e. the rise of pragmatic meaning and interpersonal meaning out of lexical meaning. Specifically speaking, the grammatical constructionalization process of *huashuo* involves the emergence of discourse meanings about topical relevance and interpersonal relationship out of the literal interpretation of the phrase as an SV reporting clause.

Within the time frame of the Chat dataset, the discourse marking functions of *huashuo* have not exhibited further significant changes. In Figure 31, the X-axis represents the time span of the Chat dataset and the Y-axis shows the relative token frequency of each discourse marking function of *huashuo*, i.e. the percentage of each function among all occurrences of discourse marker *huashuo* within a certain time period. It shows that topic introduction is the most dominant function of *huashuo* as a discourse marker throughout the

dataset, while the interpersonal functions are always much less frequent in comparison.

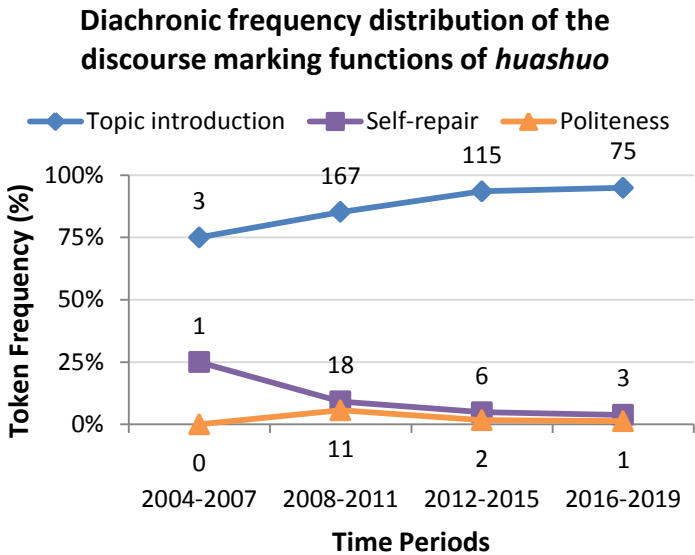


Figure 31 Diachronic frequency distribution of the discourse marking functions of *huashuo* in the Chat dataset (sum = 402)

The emergence of interpersonal functions has been considered a significant development of the *huashuo* construction as a discourse marker, as it further distinguishes discourse marker *huashuo* from the SV-clausal *huashuo* that was often used as a narrative starter (Zhou [周] 2012; Shi [施] 2013; Chen [陈] & Huang [黄] 2017). The low frequencies of these functions so far could be attributed to the fact that *huashuo* is still at its relatively early stage of development as a discourse marker. The future trend of development, especially regarding intersubjectification, calls for further studies.

Within the topic-introducing functional domain, *huashuo* indeed exhibits signs of further development. In Figure 32 we can see a steady increase in using *huashuo* to initiate a brand new topic. And in Figure 33 we can see that *huashuo* is becoming increasingly used in linking new topics to pragmatically established information instead of to the discursive pretext.

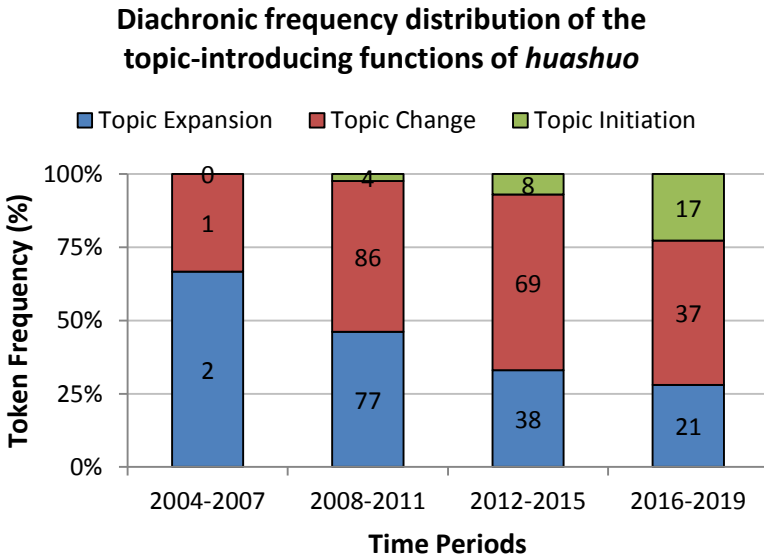


Figure 32 Diachronic frequency distribution of the topic-introducing functions of *huashuo* in the Chat dataset (sum = 360)

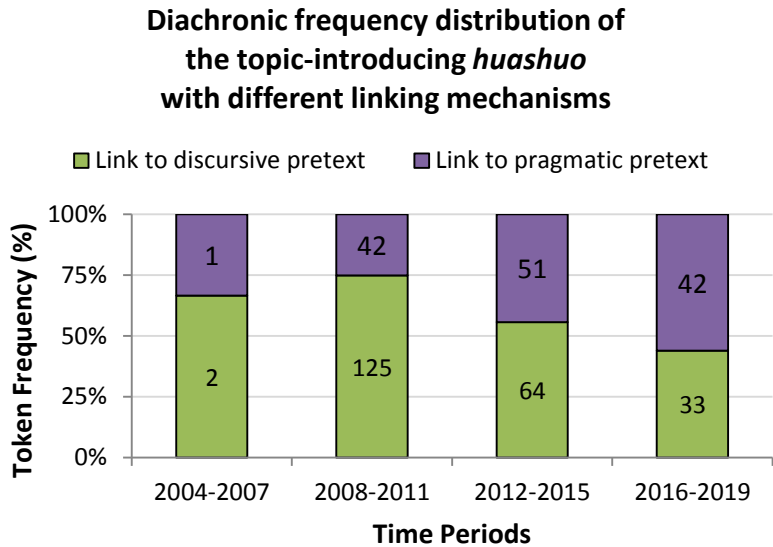


Figure 33 Diachronic frequency distribution of the topic-introducing *huashuo* with different linking mechanisms in the Chat dataset (sum = 360)

These changes show that the topic-introducing *huashuo* is becoming increasingly pretext-independent over time and thus represent further development along the pragmatization process (cf. Section 3.4.2).

4.4.3 Change in peripheral positions

As discussed in length in Section 4.3.3, discourse marker *huashuo* is observed in different peripheral positions. Diachronically, as shown in Figure 34, the LP position is the most dominant position throughout the Chat dataset; the MP and especially the RP uses seem to have

appeared later in time, i.e. around the end of the 2000s and the beginning of the 2010s respectively, and they have been rather low in frequency so far. Nonetheless, the emergent non-LP uses are a sign of further development. It shows that *huashuo* is gaining syntactic mobility as a non-propositional element in the utterance.

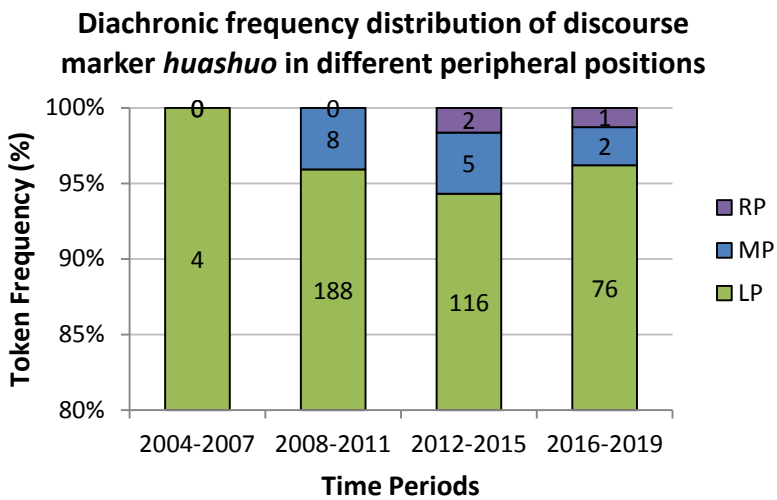


Figure 34 Diachronic frequency distribution of discourse marker *huashuo* in different peripheral positions in the Chat dataset (sum = 402)

In terms of the left-to-right movement hypothesis that is associated with the development of functional asymmetry between subjectivity and intersubjectivity, the case of *huashuo* does not represent any positive evidence. Firstly, there is no strict correlation between the peripheral positions and the (inter)subjectivity of the discourse marking functions of *huashuo* (see Section 4.3.3 for the

detailed discussion). Secondly, there is no significant diachronic change in the peripheral behavior of these functions. For the topic-introducing functions, it is clear that the LP position is the most dominant position (Figure 35); for the interpersonal functions, even though their usage in the LP position seems to be declining over the 26 years (Figure 36), their total number is yet too small for the frequency variation to be viewed as statistically significant.

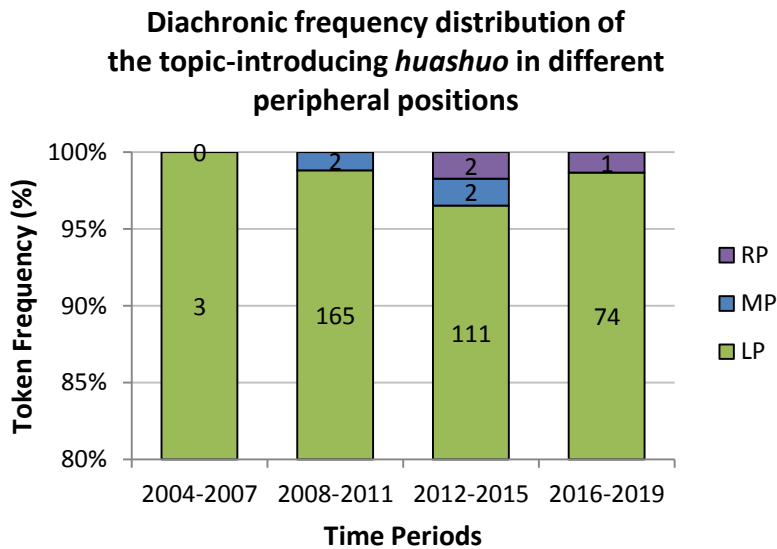


Figure 35 Diachronic frequency distribution of topic-introducing *huashuo* in different peripheral positions in the Chat dataset (sum = 360)

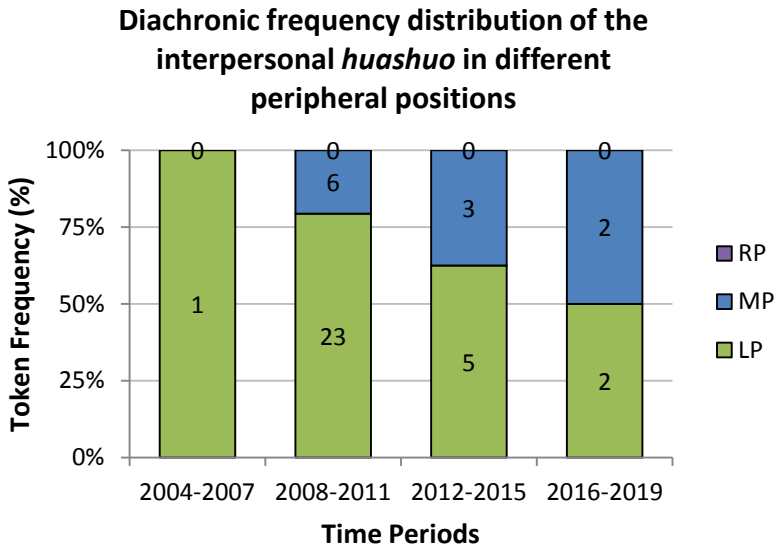


Figure 36 Diachronic frequency distribution of interpersonal *huashuo* in different peripheral positions in the Chat dataset (sum = 42)

For the time being, discourse marker *huashuo*'s expansion to the MP and RP positions from its original LP position, when viewed as a left-to-right movement at all, is not a manifestation of developing functional asymmetry and it does not coincide with intersubjectification. However, due to the limited corpus size and time span in the current study, close attention should be paid to the future development of *huashuo* and especially to its intersubjectification process in order to gain more insight into this aspect.

4.4.4 Motivation for change

As shown in Section 4.4.1, the SV-clausal usage of *huashuo* had been dominating for centuries and it is widely believed to be the starting point of *huashuo*'s development as a discourse marker. In order to understand how this usage motivated the emergence of topic-introducing and interpersonal functions, it is necessary to understand the dominance of the SV-clausal *huashuo* in the first place.

The SV-clausal *huashuo* as well as the VP and NP *huashuo* compounds are all products of the disyllabization process, which is a distinctive process during the development of the Chinese language (Hu [胡] 1923; Shi [石] 2011: 27–31; Wang [王] 2013: 333–337; Dong [董] 2017: 9–15). It refers to the lexicalization process of disyllabic words, which features the compounding of monosyllabic words and the degrading of once free-standing lexica to lexemes. The strategies of forming disyllabic words vary and are constrained by different semantic, syntactic and prosodic conditions. The VP and NP *huashuo* compounds were coined through the synonym-compounding strategy – the agglutination of two monosyllabic words that carry similar semantic meanings, i.e. $hua_V + shuo_V$ and $hua_N + shuo_N$, respectively, while the SV-clausal *huashuo* developed from connecting two elements from different syntactic categories, namely $hua_N + shuo_V$ (see Section 4.2.1 for the detailed semantic analysis of *huashuo*).

The emergence of the SV-clausal interpretation of *huashuo* coincided with, and was probably owing to, the nominalization of *hua* and the increasing dominance of the verbal meaning of *shuo* (c.f. Tian [田] 2007: 25–28). That is to say, the SV-clausal *huashuo* is composed

by the dominant meanings of both lexical constituents and is therefore semantically more salient in comparison to the VP and NP alternatives. And since its semantic content is fairly specific, unlike the broad concepts expressed by the VP and NP *huashuo*, it had few semantic and functional equivalents and hence less competition in obtaining a distinct spot in the lexicon (cf. Wang [王] 1998; Tian [田] 2007).

The semantic compatibility and hence the functional aptness in starting a narrative made the SV-clausal *huashuo* an optimal recruit in the story-telling context. Especially by virtue of its compact form in comparison to the other full-clause alternatives, such as (169) and (170), it soon became the conventionalized choice to start a story-telling event. This frequent usage thus associated *huashuo* strongly with the introduction of new information and motivated the emergence of the topic-introducing functions beyond the story-telling context later.

(169) 此 本 话 说 [...]

cǐ běn huà shuō [...]

this CLF story say

“*This piece of story tells [...]*”

(Xīxiāngjì Zhūgōngdiào [西厢记诸宫调], 1200s) (Sun [孙] 1953: 39)

(170) 今日 话 说 的 [...]

Jīn rì huà shuō de [...]

today story say ADJ

“*What today’s story/chapter talks about [...]*”

(Dàsòng Xuānhé Yíshì [大宋宣和遗事], 1300s, CCL_Ancient)

Moreover, the contextual characteristics associated with the SV-clausal usage could have further motivated the development of the linking mechanisms and the interpersonal functions of *huashuo*.

Firstly, the SV-clausal *huashuo* was often used to start a narrative that is part of a consecutive storyline, i.e. chapters and episodes that construct a whole story. That means, the SV-clausal *huashuo* often created sequential connections between different parts of a story. In relating new information to an existing information framework, this characteristic thus explains how discourse marker *huashuo* developed the mechanism of linking the new utterance to the discursive pretext and to existing information in general.

Secondly, the story-telling context distanced the speaker from their own utterances. As a narrative-starter, *huashuo* was constantly used to bring in information which was part of a story, which did not represent the speaker's personal beliefs or even the truth. This contextual characteristic thus later gave rise to the use of *huashuo* as a hedge that mitigates the speaker's commitment to the truth value and the potential impact of their utterance, i.e. the development of interpersonal functions that tend to the intended meaning and politeness.

4.5 Summary

In this chapter I discussed the formal and semantic aspects of the *huashuo* construction as well as its discourse marking functions and usage patterns both synchronically and diachronically. The findings

are supported by corpus-based observations from both functional and frequency distributional perspectives.

Even though the compound *huashuo* allows a number of different semantic interpretations, discourse marker *huashuo* carries meanings that are conventionalized and can no longer be fully constructed by its lexical components. It can be used as a topic-introducing device that connects the new utterance to pre-existing information. By doing so, it facilitates the perception of topical relevance and discourse coherence. On the interpersonal plane of the discourse, *huashuo* can be used as a self-repair tool and a politeness marker. Both functions center on hearer-orientation. They represent the speaker's extra efforts in designing their utterance for the hearer: the self-repair function aims to help the hearer achieve a better understanding of the speaker's intended meaning, and the politeness function aims to soften the potential impact of an utterance on the hearer.

The discourse marking functions of the *huashuo* construction are shown to have developed from the dominant SV-clausal usage of *huashuo* in the story-telling context in the past. The particular development path of *huashuo* showcases grammatical constructionalization and pragmaticalization. The use of *huashuo* has changed from a central propositional component in a sentence to an ancillary procedural element in the utterance. The dominant usage patterns and the contextual characteristics associated with the SV-clausal *huashuo*, namely the narrative-starter usage and the story-telling context, have contributed to the emergence and

development of different aspects of the functionalities of discourse marker *huashuo*.

Chapter 5

Conclusion

In this Chapter, I first summarize the findings of the contrastive study conducted in this thesis regarding the SPOX construction and the *huashuo* construction (Section 5.1). These findings subsequently shed light on the complex issue of language change involving the emergence and development of discourse markers in more general terms (Section 0). Some suggestions for further research are given by the end of this thesis in Section 5.3.

5.1 Summary of findings

With an interest in exploring cross-linguistic differences and generalizations in the linguistic category of discourse markers, this thesis analyzes the SPOX construction in English and the *huashuo* construction in Chinese, which share the semantic root of “speak” depicting a general speech act and both typically function as topic-introducing discourse markers.

As discourse markers are known to be context-dependent and polyfunctional on a language-specific basis, even between cognates, it is rare to find cross-linguistic correspondence that matches both on the semantic level and the functional level (e.g. English *actually* vs. French *actuellement*, Dutch *toch* vs. German *doch*) (see Aijmer & Simon-Vandenberg 2006; Simon-Vandenberg & Willems 2011; Auer & Maschler 2016). The issue is even more complex between

linguistically more distant languages and when adding diachronic aspects to the comparison.

Even though the SPOX construction and the *huashuo* construction are comparable in many ways, for which they are discussed as the English and Chinese counterparts in this thesis, they are not 100% equivalents of each other. My findings reveal both commonalities and differences in terms of their formal-semantic properties (Section 5.1.1), functionalities and usage patterns (Section 5.1.2), and pathways of change through time (Section 5.1.3).

5.1.1 Formal-semantic properties in contrast

The SPOX construction and the *huashuo* construction both have compact and formulaic forms. The SPOX construction exhibits more formal variability. Despite having a fixed head “speaking of”, it has an X-slot in the structure that can be filled with a variety of linguistic elements or left completely unfilled. The *huashuo* construction, by contrast, has an invariable compound structure with the components *hua* and *shuo*.

Both SPOX and *huashuo* have the formal representation of a clausal element in the sentence. The components of SPOX constitute an *ing*-form participle clause, which is typical for adnominal or adverbial modifier uses, providing additional propositional content to modify the matrix clause. The compound *huashuo* consists of two lexemes: *hua* meaning ‘utterance’ and *shuo* meaning ‘speak’. The compound allows highly versatile morpho-semantic interpretations, not only because *hua* and *shuo* can take on very different concrete

meanings in different contexts, but also because their lexical meanings have undergone significant changes during the long history of the Chinese language. The compound had been used as a verb phrase ‘talk about’ and a noun phrase ‘speech’ in ancient Chinese, while in modern Chinese it is rather interpreted as a subject-verb clause meaning ‘the story tells’.

However, the semantic content of neither construction can be fully described with a facile clausal interpretation based on the lexical meanings of their components. As discourse markers, SPOX and *huashuo* don’t participate in the propositional content of the sentence, but convey pragmatic and metalinguistic information about topical relevance and interlocutor relationship in order to achieve a better discourse organization. Therefore they both require a constructional understanding of the expressions in order to comprehend their meanings as discourse markers.

5.1.2 Functional aspects in contrast

Typical for discourse markers, SPOX and *huashuo* are both polyfunctional. Besides the common topic-introducing functions, SPOX can also be used as a speech act adverbial and a commenting device, while *huashuo* has additional interpersonal functions in self-repair and politeness management. Nonetheless, topic introduction is the most dominant function for both SPOX and *huashuo* (see Figure 37).

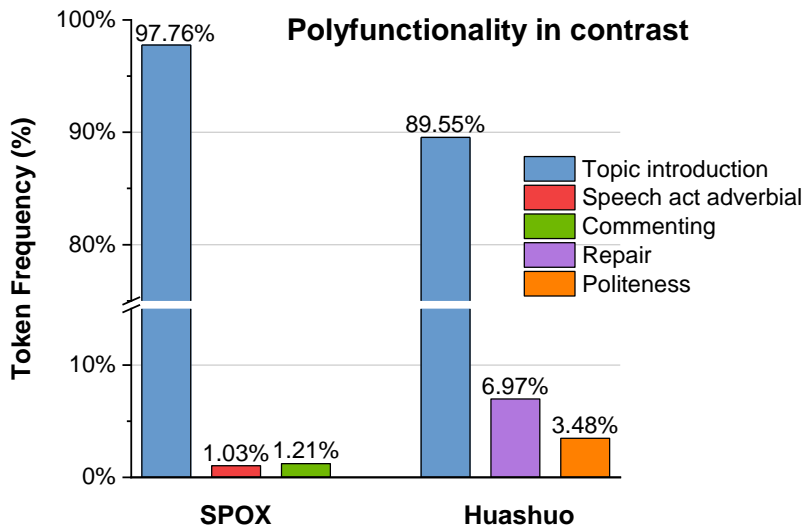


Figure 37 Relative frequency distribution of the discourse marking functions of SPOX and *huashuo* in the COCA-SPOK dataset and the Chat dataset respectively (the Y-axis has a break between 15% und 75% for better visual presentation)

As topic-introducing discourse markers, both SPOX and *huashuo* can be used either between utterances as a segue device for moving from one (sub)topic to the other (i.e. *topic expansion* and *topic change* events) or at the very beginning of an utterance or even a discourse event, introducing a brand new topic into the discourse (i.e. *topic initiation* events). But the mechanisms behind these functions exhibit some differences between SPOX and *huashuo*.

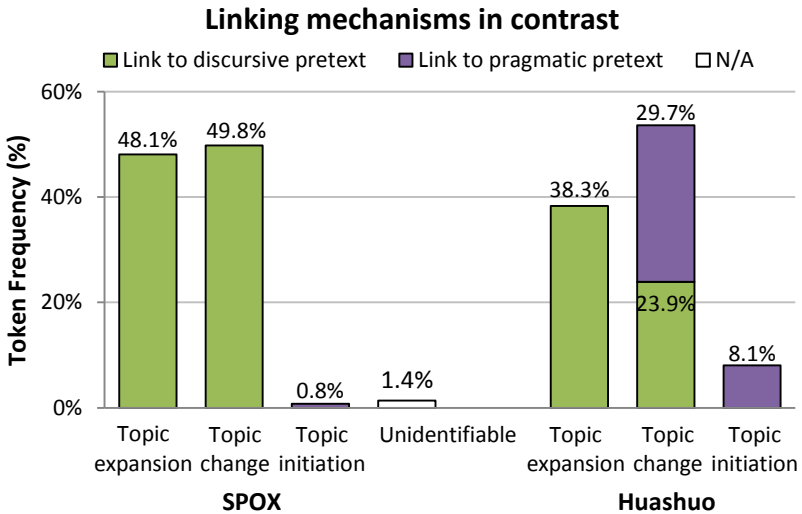


Figure 38 Relative frequency distribution of the linking mechanisms of topic-introducing SPOX and *huashuo* in the COCA-SPOK dataset and the Chat dataset respectively

Usually, the use of SPOX hints at a textual and/or conceptual link between the utterances, with which topical relevance is constructed. The link can be either explicit or implicit, depending on whether and how the X-element is given in the structure (see Section 3.3.1.1 for detailed discussion in this respect). In any case, the use of SPOX compels the hearer to actively search for a connection between the new utterance and the previous discourse so as to make sense of the transition and to assign topical relevance to the new utterance, which in turn facilitates the perception of discourse coherence.

By contrast, the linking mechanism of *huashuo* does not rely on any X-element and it does not necessarily point to the discursive pretext. The use of *huashuo* does link the new utterance to

pre-existing information, but it is often non-verbalized and pragmatically given information existing in the interlocutor relationship, instead of linguistically encoded information in the preceding discourse (see Sections 4.3.1.1 and 4.3.1.2 for detailed discussion in this respect). This linking mechanism allows *huashuo* to bring in both textually and conceptually unrelated topics into the ongoing discourse and to initiate a brand new topic without any discursive pretext.

As the use of SPOX implies a textual and/or conceptual connection between the different topics, it cannot be used in topic change events in which the new topic has no connection to the discursive pretext. For the same reason, it is rather counter-intuitive to observe SPOX in topic initiation events as a topicalization device, in which the X-element of SPOX highlights the topic of the following discourse but does not serve as any link to the discursive pretext as there is none. In this usage, SPOX atypically connects the new topic to a pragmatic pretext, as *huashuo* typically does, and therefore does not require the existence of a discursive pretext (see Section 3.3.1.2 for detailed discussion in this respect). But this usage needs to be understood as genre-specific for the time being. As limited by the type of corpus data used in this thesis, the topic initiation function of SPOX is only observed in media language use and is therefore conditioned by the special characteristics of the host-audience relationship. Further research needs to be done with a different type of spoken data to better understand this usage.

The frequency distributional analysis, as shown in Figure 38, confirms that the topic initiation function of SPOX is rather atypical in

comparison to that of *huashuo* (i.e. 0.8% vs. 8.1%). At the same time, it also shows that it is generally more typical for *huashuo* than SPOX to mark topical movements that are motivated by the pragmatic pretext instead of by the discursive pretext.

The comparison between the topic-introducing SPOX and *huashuo* regarding their usage in turn-taking operations shows that SPOX is used more often in dialogual contexts to pick up a turn, while *huashuo* is used more often in monologual contexts to continue with the speaker's own course of talk (Figure 39). The “not applicable (N/A)” tokens refer to a few topic initiation uses of SPOX and *huashuo*, which mark the start of a brand new discourse event and therefore do not involve turn-taking operations.

Further research on *huashuo* in this specific aspect still needs to be done with a different type of data in order to gain better insights. The CMC data used in this thesis may only reveal a biased pattern of *huashuo*'s usage in turn-taking operations, as the coordination between interlocutors in a CMC setting is different from that in face-to-face conversations.

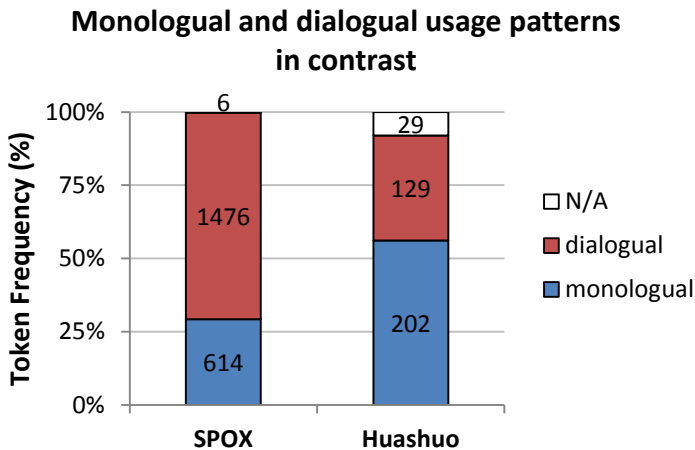


Figure 39 Frequency distribution of monologual and dialogual uses of topic-introducing SPOX and *huashuo* in the COCA-SPOK dataset (sum = 2096) and the Chat dataset (sum = 360) respectively

As to peripheral behavior, both SPOX and *huashuo* favor the LP position as discourse markers (Figure 40). The LP position is known to be typical for discourse markers that assist the interpretation and processing of the following utterance and mark its relation to the preceding part (Biber et al. 1999: 891; Virtanen 2004). Since SPOX and *huashuo* are most typically used as topic-introducing devices, it is natural for them to prevail in the LP position. Different kinds of discourse marker uses of SPOX and *huashuo* are observed in other peripheral positions as well, but they are much lower in frequency in comparison. The “not applicable (N/A)” tokens refer to the commenting uses of SPOX, in which it is not attached to any matrix clause.

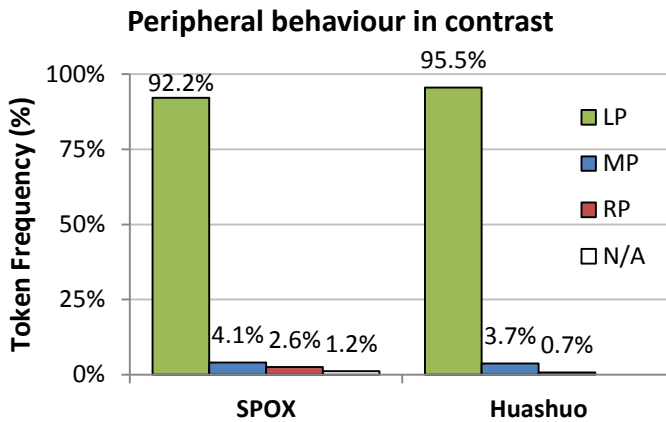


Figure 40 Relative frequency distribution of SPOX and *huashuo* as discourse markers in different peripheral positions in the COCA-SPOK dataset and the Chat dataset respectively

In terms of subjectivity and intersubjectivity, both SPOX and *huashuo* index subjective meaning (speaker-orientation) and intersubjective meaning (hearer-orientation), but *huashuo* seems to express stronger intersubjective meaning than SPOX in certain functions (see Table 10).

As topic-introducing devices, SPOX and *huashuo* mainly indicate how the utterance is related to the speaker's perspectives or viewpoints, i.e. they reveal the speaker's evaluation of the topical relevance of the new utterance. Intersubjective meaning may be postulated, when the expressions are used in non-initial positions, as such usage pattern shows additional efforts in the hearer design.

In other functional domains, SPOX expresses subjectivity as well: as a speech act adverbial, it indicates the speaker's viewpoint on

5 Conclusion

the quality and quantity of the information contained in their utterance; and as a commenting device, it communicates the speaker’s subjective opinion and attitude. The use of *huashuo*, by contrast, expresses mainly intersubjectivity in its additional interpersonal functions: as a repair tool, it shows the speaker’s attention to the hearer’s state in the conversation and particularly to their potential need for assistance in getting the intended meaning; and as a politeness marker, it shows the speaker’s specific attention to the face of the hearer.

Table 10 Subjectivity and intersubjectivity of different discourse marking functions of SPOX and *huashuo*

<i>speaking of X</i>		
Discourse marking functions	subjective	intersubjective
Topic introduction	√	(√ in non-initial positions)
Speech act adverbial	√	–
Commenting	√	–
<i>huashuo</i>		
Discourse marking functions	subjective	intersubjective
Topic introduction	√	(√ in non-initial positions)
Self-repair	–	√
Politeness management	–	√

On the basis of the findings regarding both peripheral behavior and (inter)subjective meanings, it is clear that neither SPOX nor *huashuo* evidences functional asymmetry at the current stage of their development. Even though due to the dominance of the topic-introducing functions and their prevalence in the LP position, some less dominant function may appear to favor a different periphery, e.g. the speech act adverbial function of SPOX that seems to occur more often in non-initial positions, in general, the discourse marking functions of SPOX and *huashuo* and the subjectivity and intersubjectivity they index in usage are not categorically associated with any specific peripheral positions (see detailed discussion in Section 3.3.5 and Section 4.3.3 in this respect.).

5.1.3 Diachronic development in contrast

The SPOX construction and the *huashuo* construction share the semantic root of “speak” at the very beginning of their constructionalization and pragmaticalization processes. Their overall development paths thus follow by and large the same line of change: i.e. from a speech-act verb phrase with full lexical meaning to a discourse marker with primarily topic-introducing functions. The original lexical uses of SPOX and *huashuo*, namely “speaking of” and “*huashuo*” as verb phrases in a sentence, are exemplified in (171) and (172).

- (171) *Some of the guests were speaking of it when he entered, [...]*
(1836, FIC, Philothea, COHA)

5 Conclusion

(172) 一日，夫人犹能力为行坐，**话说**家事。

Y í rì fū rén yóu néng lì wéi xíng zuò, **huàshuō** jiā shì
one day lady still can try do walk sit **talk about** family matter

“One day, Lady was still able to walk and sit and **talk about** the family matters.”

(Tángdài Mùzhì míng Huìbiān Xùjí [唐代墓志汇编续集], 800s, CCL_Ancient)

Both SPOX and *huashuo*, with their original verbal meanings, were often used as clausal elements in a sentence prior to the emergence of their discourse marking functions. Looking into the individual processes involved, the development paths of SPOX and *huashuo* have their own distinctive courses and motivating factors.

The emergence of discourse marker SPOX is motivated by the clausal usage of the verb phrase in the present participle form, i.e. “speaking of”, as an adverbial modifier, as exemplified in (173).

(173) *Speaking of the acquittal of Captain Preston and his soldiers, tried at Boston in 1770, the author says, it was a thing truly remarkable, [...]*
(1821, MAG, NorthAmRev, COHA)

The most dominant usage patterns associated with the adverbial SPOX, including sentence-initial position, asyndetic subordination and co-occurrence with speech-act verbs, motivated scope expansion of the phrase from clause-internal to sentential and semantic change of the expression from propositional to procedural (see detailed discussion in Section 3.4.4).

The origin of discourse marker *huashuo* is rooted in the subject-verb clausal usage of the phrase as a customary way to start a narrative, as exemplified in (174).

- (174) 话说 元朝 时，
Huàshuō yuán cháo shí
 story tells Yuan Dynasty time
 都 下 有 个 李总管，
 dū xià yǒu ge lǐ zǒngguǎn，
 capital under have CLF Li general manager
 官 居 三品， 家业 巨富。
 guān jū sān pǐn, jiāyè jùfù
 office place third rank family property great wealth

“The story tells that, in the Yuan Dynasty, there was a general manager Li in the capital, who was a third-ranked official with a wealthy family.”

(Chūkè Pāi‘àn Jīngqí [初刻拍案惊奇], 1627, CN_Ancient)

The special characteristics of the story-telling context regarding how topicality is structured and construed and how interlocutor relationship is perceived and maintained gave rise to the specific topic-introducing functions with different linking mechanisms and the interpersonal functions (see detailed discussion in Section 4.4.4).

Both SPOX and *huashuo* exhibit pragmatic strengthening during their development towards discourse markers, i.e. their meanings in a sentence have both changed from primarily propositional to primarily pragmatic. For both SPOX and *huashuo*,

their semantic content has changed from referring to the action of speaking of a certain topic to providing non-truth-conditional information about the speech act, e.g. shift of topicality, connection to pretext, etc. In terms of pragmatic strengthening on the (inter)personal level, i.e. (inter)subjectification processes, *huashuo* shows further development than SPOX, as it has developed intersubjective meanings represented by its interpersonal functions.

5.2 Implications of findings

In reference to the observations made in Lehmann (1995; 2015), Hopper (1991), and Claridge and Arnovick (2010) in terms of grammaticalization, pragmaticalization, and mechanisms and processes leading towards discourse markers (see Sections 2.3.1 and 2.3.3), SPOX and *huashuo* exhibit the following behavior:

- (a) The development of SPOX and *huashuo* showcases a process of pragmaticalization. They have developed textual/discourse-oriented meanings and (inter)subjective meanings out of propositional meanings.
- (b) During the process of pragmaticalization, the semantic content of neither SPOX nor *huashuo* has been bleached completely. Some traces of the original lexical meanings of “speak of (a certain topic)” and “(the story) talk(s) about” still persist in the discourse marker usage, even though their participation in the sentence becomes more procedural than propositional.

- (c) Both SPOX and *huashuo* have become syntactically and prosodically detached from the rest of the utterance (reflecting propositional non-integration), often manifested by the use of comma punctuation in a sentence or a pause in speech.
- (d) Both SPOX and *huashuo* show scope expansion in their development towards discourse markers. They have developed from propositional elements with a local/clause-internal scope to discourse markers with a scope over the entire proposition.
- (e) Both SPOX and *huashuo* have become polyfunctional as discourse markers. From being once monofunctional conceptual elements in a sentence, they have developed multiple metalinguistic and discourse organizing functions.
- (f) The emergence of some discourse marking functions of SPOX and *huashuo* has brought about layering with existing expressions: e.g. the commenting SPOX layers with the set phrase “Talk about X!”; the topic initiating SPOX layers with topicalization devices such as *as to* and *as for*; the politeness marker *huashuo* layers with the politeness marker *nàgè* [那个] – ‘that’ (cf. Wang 2011).
- (g) The development of discourse marker meanings of SPOX and *huashuo* features divergence from their original meanings as adverbial clauses and subject-verb reporting clauses, which continue to exist alongside the new usage.
- (h) Both SPOX and *huashuo* have undergone decategorialization and paradigmaticization as discourse markers: i.e. they lost their

original verbal/clausal status and became integrated into the new paradigm/grammatical category of discourse markers.

- (i) The newest variant of SPOX – *speaking of* without any X-element – can be viewed as a sign of formal reduction (phonological attrition) or the initial stage of one. By contrast, *huashuo* so far has not shown any sign of formal reduction as a discourse marker.

As SPOX and *huashuo* represent linguistically heterogeneous and typologically distant languages, the contrastive study, especially from the diachronic perspective, helps to make the complex issues involved in the emergence and development of discourse markers become more transparent. Based on the observations made in this thesis, pragmatic strengthening, syntactic/prosodic detachment, scope expansion and development of functional polysemy appear to be universal processes during the constructionalization process of discourse markers, while formal reduction and semantic bleaching seem to be less categorical and contingent upon the type of the discourse markers, and differ from one case to another. The phenomena of layering, persistence, divergence, decategorialization and paradigmaticization seem to be universal common properties of emergent discourse markers as well.

5.3 Suggestions for further research

5.3.1 Dataset expansion

The corpus-based approach adopted in this thesis facilitates the empirical studies immensely. At the same time, it also reveals specific demands on the type of corpus data suited to the research questions at the crossroads of pragmatics, historical linguistics, and contrastive linguistics. On account of the limitations presented by the corpus data available in this thesis, I suggest that the research questions can be further explored if the datasets are expanded in the following ways:

First of all, the functional analyses of SPOX and *huashuo* can be enriched by investigating datasets with different types of spoken data. As already pointed out in Section 3.3.1.2, the spoken data of SPOX can be expanded to other genres than media language in order to better understand the topic initiation usage: Is it indeed a genre-specific phenomenon? If not, how is SPOX used as a topicalization device in other speech genres? Are there any special characteristics or constraints to observe? The functional analyses of *huashuo* can be improved if real-time face-to-face spoken data are added to the dataset. It would help to better understand *huashuo*'s usage patterns in turn-taking operations, as it eliminates the bias caused by the special characteristics of CMC data (see the relevant discussion in Sections 4.1.1 and 4.3.4).

Second, datasets with broader time spans will reveal more insights into the development of SPOX and *huashuo* as discourse

markers. For SPOX, further studies can look into corpus data before the 1810s and examine the development of the adverbial usage of SPOX. The observations may provide us with a fuller picture of the constructionalization process of SPOX and perhaps reveal when and how the adverbial SPOX start to gain prominence among all the semantic roles it can assume in a sentence and which constructional changes are involved in the process. For both SPOX and *huashuo*, I also suggest to keep track of their future development, especially in terms of (inter)subjectification, functional asymmetry and left-to-right movement, and formal reduction.

5.3.2 Comedic effect

Due to the conventionalized topic-introducing functions of SPOX and its linking mechanisms, comedic effect can occur when SPOX is “exploited” to imply a connection between utterances when there is clearly none or when the connection it implies is rather unexpected, such as in examples (175) and (176) respectively.

(175) [Will is sitting on the couch reading. Naked. Grace enters and makes her way to Will’s fridge.]

GRACE: *Hi. Just wanted to get some pudding. You got pudding at the market. I didn’t.* [taking a pudding cup from the fridge] *This is fat-free, right?*

WILL: *Yeah.*

GRACE: *Great. Thanks. Bye.*

[Grace leaves to re-enter immediately.]

GRACE: *Oh, it’s chocolate. I thought I grabbed the swirl.*

WILL: *You know, it's funny. Speaking of chocolate versus swirl. I'm naked! I'm devoid of clothing. I'm in my nude.*

GRACE: *I know. You're so naked these days. Since when did you start being naked all the time?*

WILL: *Well, since I started living alone.*

(Will and Grace, Season 2 Episode 1, 1999)

- (176) *You probably noticed that there were a little more security than normal here today. Hillary Clinton is here. [Audience cheering] So I know you had to go through a lot: you stood in line, someone checked your ID – it's like going to the bathroom in North Carolina. [Audience laughing] Speaking of things that are hard getting into, it's almost swimsuit season. And that means getting into shape.*

(The Ellen Show, 2016)

The comedic effect in these examples can be explained by the *incongruous juxtaposition theory* (Freud 1905) in humor research: i.e. humor can be found in fundamentally incompatible concepts or unexpected resolutions³⁴. It is not the focus of this thesis, but it could be an interesting perspective in further studies of the SPOX construction.

³⁴ I thank Prof. Dr. Stephan Packard for his generous advice on the humor research and comics studies.

References

- Abraham, Werner. 1991. *Discourse Particles: descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German* (Pragmatics & Beyond New Series 12). Amsterdam/Philadelphia: John Benjamins.
- Aijmer, Karin. 1997. *I think* - an English modal particle. In Toril Swan & Olaf Jansen Westvik (eds.), *Modality in Germanic Languages* (Trends in Linguistics. Studies and Monographs 99), 1–47. Berlin/New York: Mouton de Gruyter.
- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus* (Studies in Corpus Linguistics 10). Amsterdam/Philadelphia: John Benjamins.
- Aijmer, Karin. 2013. *Understanding pragmatic markers: a variational pragmatic approach*. Edinburgh: Edinburgh University Press.
- Aijmer, Karin. 2016. Pragmatic Markers as Constructions: The Case of *Anyway*. In Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds.), *Outside the Clause: Form and function of extra-clausal constituents* (Studies in Language Companion Series 178), 29–57. Amsterdam/Philadelphia: John Benjamins.
- Aijmer, Karin & Simon-Vandenberg, Anne-Marie. 2006. *Pragmatic markers in contrast*. Amsterdam [u.a.]: Elsevier.
- Alm, Maria, Behr, Janina & Fischer, Kerstin. 2018. Modal particles and sentence type restrictions: A construction grammar perspective. *Glossa* 3(1). 1–32.
- Andersen, Gisle. 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance-theoretic Approach to the Language of Adolescents*

References

- (Pragmatics & Beyond New Series 84). Amsterdam/Philadelphia: John Benjamins.
- Andersen, Gisle. 2015. Discourse topic markers from a cross-linguistic and variational perspective. (Presented at the 14th International Pragmatics Conference [IPrA 14], Antwerp.)
- Antonopoulou, Eleni & Nikiforidou, Kiki. 2011. Construction grammar and conventional discourse: A construction-based approach to discoursal incongruity. *Journal of Pragmatics* 43. 2594–2609.
- Ariel, Mira. 1998. Discourse Markers and Form-function Correlations. In Andreas H. Jucker & Yael Ziv (eds.), *Discourse markers: descriptions and theory* (Pragmatics & Beyond New Series 57), 223–259. Amsterdam/Philadelphia: John Benjamins.
- Ariel, Mira. 2008. *Pragmatics and Grammar* (Cambridge Textbooks in Linguistics). Cambridge [u.a.]: Cambridge University Press.
- Ariel, Mira. 2009. Discourse, Grammar, Discourse. *Discourse Studies* 11(1). 5–36.
- Ariel, Mira. 2010. *Defining Pragmatics* (Research Surveys in Linguistics). Cambridge [u.a.]: Cambridge University Press.
- Auer, Peter & Maschler, Yael (eds.). 2016. *NU / NÅ: A Family of Discourse Markers Across the Languages of Europe and Beyond* (Linguae & Litterae 58). Berlin/Boston: Walter de Gruyter.
- Auer, Peter & Pfänder, Stefan (eds.). 2011. *Constructions: Emerging and Emergent* (Linguae & Litterae 6). Berlin/Boston: Walter de Gruyter.
- Austin, John L. 1962. *How to do things with words: the William James lectures delivered at Harvard University in 1955*. Oxford [u.a.]: Oxford University Press.

- Bai, Yinchun. 2014. A usage-based study of the *just me* construction. In Martin Hilpert & Susanne Flach (eds.), *Yearbook of the German Cognitive Linguistics Association*, vol. 2, 127–146. Berlin/Boston: Walter de Gruyter.
- Bai, Yinchun. 2015. The *just me* construction. Is it just me, or does it work like a discourse marker? In Mieke Vandenbroucke, Tine Defour, Bart Defrancq, Stavros Kelepouris & Jonah Rys (eds.), *Studies van de BKL*, vol. 9. Ghent: Belgische Kring voor Linguïstiek (BKL). (<https://sites.uclouvain.be/bkl-cbl/nl/publicaties/studies-van-de-bkl/volume-9-2015/>)
- Bai, Yinchun. 2021. Functional asymmetry and left-to-right movement: Speaking of peripheries. In Daniël van Olmen & Jolanta Šinkūnienė (eds.), *Pragmatic Markers and Peripheries* (Pragmatics & Beyond New Series 325), 303–326. Amsterdam/Philadelphia: John Benjamins.
- Barðdal, Jðanna, Smirnova, Elena, Sommerer, Lotte & Gildea, Spike (eds.). 2015. *Diachronic construction grammar* (Constructional Approaches to Language 18). Amsterdam/Philadelphia: John Benjamins.
- Barth-Weingarten, Dagmar & Couper-Kuhlen, Elizabeth. 2002. On the development of final *though*: A case of grammaticalization? In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization* (Typological Studies in Language 49), 345–361. Amsterdam/Philadelphia: John Benjamins.
- Bayer, Josef & Struckmeier, Volker (eds.). 2017. *Discourse Particles* (Linguistische Arbeiten 564). Berlin/Boston: Walter de Gruyter.
- Beeching, Kate. 2016. *Pragmatic markers in British English: meaning in social interaction*. Cambridge [u.a.]: Cambridge University Press.

References

- Beeching, Kate & Detges, Ulrich (eds.). 2014. *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12). Leiden/Boston: Brill.
- Beeching, Kate & Wang, Yu-fang. 2014. Motivation for meaning shift at the left and right periphery: *well, bon and hao*. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12), 47–71. Leiden/Boston: Brill.
- Beijering, Karin. 2015. The lexicalization-grammaticalization-pragmaticalization interface: The case of Mainland Scandinavian *jeg tror*. In Andrew D. M. Smith, Graeme Trousdale & Richard Wälchli (eds.), *New Directions in Grammaticalization Research* (Studies in Language Companion Series 166), 67–91. Amsterdam/Philadelphia: John Benjamins.
- Bender, Emily M. & Kathol, Andreas. 2001. Constructional Effects of “Just Because . Doesn’t Mean .” In Andrew Simpson (ed.), *Proceedings of the Twenty-seventh Annual Meeting of the Berkeley Linguistics Society, Berkeley, March 22-25, 2001*, 13–26. Berkeley, CA: Berkeley Linguistics Society.
- Bergs, Alexander & Diewald, Gabriele (eds.). 2008. *Constructions and language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 194). Berlin/New York: Mouton de Gruyter.
- Bergs, Alexander & Diewald, Gabriele (eds.). 2009. *Contexts and Constructions* (Constructional Approaches to Language 9). Amsterdam/Philadelphia: John Benjamins.

- Biber, Douglas, Conrad, Susan & Reppen, Randi. 1998. *Corpus linguistics: Investigating language structure and use* (Cambridge Approaches to Linguistics). Cambridge [u.a.]: Cambridge University Press.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman grammar of spoken and written English*. London/Harlow: Longman.
- Bisang, Walter. 1998. Grammaticalization and Language Contact, Constructions and Positions. In Anna Giacalone Ramat & Paul J. Hopper (eds.), *The Limits of Grammaticalization* (Typological Studies in Language 37), 13–58. Amsterdam/Philadelphia: John Benjamins.
- Bisang, Walter. 2008. Precategoriality and Argument Structure in Late Archaic Chinese. In Jaakko Leino (ed.), *Constructional Reorganization* (Constructional Approaches to Language 5), 55–88. Amsterdam/Philadelphia: John Benjamins.
- Blakemore, Diane. 1987. *Semantic constraints on relevance*. Oxford [u.a.]: Blackwell.
- Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers* (Cambridge Studies in Linguistics 99). Cambridge [u.a.]: Cambridge University Press.
- Blank, Andreas. 2001. *Pathways of Lexicalization. Language Typology and Language Universals* (Handbücher Zur Sprach- Und Kommunikationswissenschaft 20), vol. 2, 1596–1608. Berlin/New York: Walter de Gruyter.
- Boas, Hans Christian. 2003. *A constructional approach to resultatives* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

References

- Boas, Hans Christian (ed.). 2010. *Contrastive Studies in Construction Grammar* (Constructional Approaches to Language 10). Amsterdam/Philadelphia: John Benjamins.
- Boas, Hans Christian & González-García, Francisco (eds.). 2014. *Romance Perspectives on Construction Grammar* (Constructional Approaches to Language 15). Amsterdam/Philadelphia: John Benjamins.
- Boas, Hans Christian & Höder, Steffen (eds.). 2018. *Constructions in Contact: Constructional Perspectives on Contact Phenomena in Germanic Languages* (Constructional Approaches to Language 24). Amsterdam/Philadelphia: John Benjamins.
- Bolden, Galina B. 2009. Implementing incipient actions: The discourse marker “so” in English conversation. *Journal of Pragmatics* 41(5). 974–998.
- Bouveret, Myriam & Legallois, Dominique (eds.). 2012. *Constructions in French* (Constructional Approaches to Language 13). Amsterdam/Philadelphia: John Benjamins.
- Boyd, Jeremy K. & Goldberg, Adele E. 2011. Learning What NOT to Say: The Role of Statistical Preemption and Categorization in A-Adjective Production. *Language* 87(1). 55–83.
- Boye, Kasper & Harder, Peter. 2012. A Usage-Based Theory of Grammatical Status and Grammaticalization. *Language* 88(1). 1–44.
- Breban, Tine. 2009. Structural persistence: a case based on the grammaticalization of English adjectives of difference. *English Language and Linguistics* 13(1). 77–96.
- Brinton, Laurel J. 1996. *Pragmatic markers in English: grammaticalization and discourse functions* (Topics in English Linguistics 19). Berlin/New York: Mouton de Gruyter.

- Brinton, Laurel J. 2002. Grammaticalization versus lexicalization reconsidered: On the late use of temporal adverbs. In Teresa Fanego, María José López-Couso & Javier Pérez-Guerra (eds.), *English Historical Syntax and Morphology: Selected papers from 11 ICEHL, Santiago de Compostela, 7-11 September 2000* (Current Issues in Linguistic Theory 223), 67–97. Amsterdam/Philadelphia: John Benjamins.
- Brinton, Laurel J. 2007. The development of *I mean*: Implications for the study of historical pragmatics. In Susan M. Fitzmaurice & Irma Taavitsainen (eds.), *Methods in historical pragmatics* (Topics in English Linguistics 52), 37–79. Berlin/New York: Mouton de Gruyter.
- Brinton, Laurel J. 2008. *The comment clause in English syntactic origins and pragmatic development* (Studies in English Language). Cambridge [u.a.]: Cambridge University Press.
- Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge [u.a.]: Cambridge University Press.
- Brinton, Laurel J. & Traugott, Elizabeth Closs. 2005. *Lexicalization and Language Change* (Research Surveys in Linguistics). Cambridge [u.a.]: Cambridge University Press.
- Brinton, Laurel J. & Traugott, Elizabeth Closs. 2007. Lexicalization and grammaticalization all over again. In Joseph C. Salmons & Shannon Dubenion-Smith (eds.), *Historical Linguistics 2005: Selected papers from the 17th International Conference on Historical Linguistics, Madison, Wisconsin, 31 July - 5 August 2005* (Current Issues in Linguistic Theory 284), 3–19. Amsterdam/Philadelphia: John Benjamins.

References

- Briz Gómez, Antonio & Hidalgo Navarro, Antonio & Padilla García, Xosé & Pons Bordería, Salvador & Ruiz Gurillo, Leonor & Sanmartín Sáez, Julia & Albelda Marco, Marta. 2003. Un sistema de unidades para el estudio del lenguaje coloquial. *Oralia* 6. 7–63.
- Brown, Penelope & Levinson, Stephen C. 1987. *Politeness: some universals in language usage* (Studies in Interactional Sociolinguistics 4). Cambridge [u.a.]: Cambridge University Press.
- Bybee, Joan. 2003. Mechanisms of Change in Grammaticization: The Role of Frequency. In Brian D. Joseph & Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 602–623. Malden, Mass. [u.a.]: Blackwell.
- Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4). 711–733.
- Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge [u.a.]: Cambridge University Press.
- Campbell, Lyle. 2001. What's wrong with grammaticalization? *Language Sciences* 23. 113–161.
- Cao [曹], Xiuling [秀玲]. 2016. 汉语话语标记多视角研究 [A Multi-dimensional Approach to Chinese Discourse Markers] (学思语言学丛书 [Learning and Reasoning Linguistics Series]). Beijing: 中国社会科学出版社 [China Social Science Press].
- Carter, Ronald & McCarthy, Michael. 2006. *Cambridge Grammar of English. A Comprehensive Guide. Spoken and Written English Grammar and Usage*. Cambridge [u.a.]: Cambridge University Press.
- Celle, Agnès & Huart, Ruth (eds.). 2007. *Connectives as discourse landmarks* (Pragmatics & Beyond New Series 161). Amsterdam/Philadelphia: John Benjamins.

- Chen, Jiajun. 2017. Dialogicity in dialogue: deriving Chinese discourse marker *bieshuo* from the negative imperative. *Journal of Pragmatics* 110. 34–49.
- Chen, Jiajun. 2018. (Inter)subjectification at the left and right periphery: Deriving Chinese pragmatic marker *bushi* from the negative copula. *Language Sciences* 66. 83–102.
- Chen [陈], Li [荔] & Huang [黄], JiaLi [佳丽]. 2017. 语法化演进中的次语法化现象——以“话说”为例 [On Secondary Grammaticalization of Introductory Particle – A case study of “Huashuo”]. 齐齐哈尔大学学报 [Journal of Qiqihar University] 239. 125–128.
- Chomsky, Noam. 1957. *Syntactic structures*. Berlin/New York: Mouton de Gruyter.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass. [u.a.]: The MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass. [u.a.]: The MIT Press.
- Claridge, Claudia & Arnovick, Leslie. 2010. Pragmaticalisation and Discursisation. In Andreas H. Jucker & Irma Taavitsainen (eds.), *Historical Pragmatics* (Handbooks of Pragmatics 8), 165–192. Berlin/New York: Mouton de Gruyter.
- Coussé, Evie & Andersson, Peter & Olofsson, Joel (eds.). 2018. *Grammaticalization meets construction grammar* (Constructional Approaches to Language 21). Amsterdam/Philadelphia: John Benjamins.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford [u.a.]: Oxford University Press.

References

- Croft, William. 2003. Lexical Rules vs. Constructions: A False Dichotomy. In Hubert Cuyckens, Thomas Berg, René Driven & Klaus-Uwe Panther (eds.), *Motivation in Language* (Current Issues in Linguistic Theory 243). Amsterdam/Philadelphia: John Benjamins.
- Croft, William. 2005. Logical and typological arguments for Radical Construction Grammar. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (Constructional Approaches to Language 3), 273–314. Amsterdam/Philadelphia: John Benjamins.
- Croft, William & Barðdal, Jónanna & Hollmann, Willem & Soitirova, Violeta & Taoka, Chiaki. 2010. Revising Talmy's typological classification of complex event constructions. In Hans Christian Boas (ed.), *Contrastive Studies in Construction Grammar* (Constructional Approaches to Language 10), 210–235. Amsterdam/Philadelphia: John Benjamins.
- Croft, William & Cruse, David A. 2004. *Cognitive linguistics*. Cambridge [u.a.]: Cambridge University Press.
- Dąbrowska, Ewa. 2009. Words as Constructions. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New Directions in Cognitive Linguistics* (Human Cognitive Processing 24), 201–223. Amsterdam/Philadelphia: John Benjamins.
- Dascal, Marcelo. 1977. Conversational relevance. *Journal of Pragmatics* 1(4). 309–328.
- Dascal, Marcelo. 2003. *Interpretation and understanding*. Amsterdam/Philadelphia: John Benjamins.
- Dascal, Marcelo & Katriel, Tamar. 1979. Digressions: A Study in Conversational Coherence. *PTL* 4. 203–232.

- Davidse, Kristin. 2000. A Constructional Approach to Clefts. *Linguistics* 38(6). 1101–1131.
- Degand, Liesbeth. 2014. “So very fast then” Discourse Markers at Left and Right Periphery in Spoken French. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12), 151–178. Leiden/Boston: Brill.
- Degand, Liesbeth & Cornillie, Bert & Pietrandrea, Paola (eds.). 2013. *Discourse Markers and Modal Particles: Categorization and Description* (Pragmatics & Beyond New Series 234). Amsterdam/Philadelphia: John Benjamins.
- Degand, Liesbeth & Evers-Vermeul, Jacqueline. 2015. Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. *Journal of Historical Pragmatics* 16(1). 59–85.
- Degand, Liesbeth & Fagard, Benjamin. 2011. *Alors* between Discourse and Grammar: The role of syntactic position. *Functions of Language* 18(1). 29–56.
- Dé, Csilla Ilona. 2010. On the Status of Discourse Markers. *Acta Linguistica Hungarica* 57(1). 3–28.
- Detges, Ulrich & Waltereit, Richard. 2011. Turn-taking as a trigger for language change. In Sarah Dessì Schmid & Peter Koch (eds.), *Rahmen des Sprechens: Beiträge zu Valenztheorie, Varietätenlinguistik, Kreolistik, kognitiver und historischer Semantik*, 175–190. Tübingen: Narr.
- Diessel, Holger. 2015. Usage-Based Construction Grammar. In Ewa Dąbrowska & Dagmar Divjak (eds.), *Handbook of Cognitive Linguistics* (Handbooks of Linguistics and Communication Science 39), 203–220. Berlin/Boston: Mouton de Gruyter.

References

- Diessel, Holger. 2019. *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. Cambridge [u.a.]: Cambridge University Press.
- Diewald, Gabriele. 2006a. Discourse particles and modal particles as grammatical elements. In Kerstin Fischer (ed.), *Approaches to discourse particles* (Studies in Pragmatics 1), 403–425. Amsterdam [u.a.]: Elsevier.
- Diewald, Gabriele. 2006b. Context types in grammaticalization as constructions. *Constructions* Special Volume 1(9). 1–29.
- Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2). 365–390.
- Diewald, Gabriele. 2013. “Same same but different” – Modal particles, discourse markers and the art (and purpose) of categorization. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Discourse Markers and Modal Particles: Categorization and Description* (Pragmatics & Beyond New Series 234), 19–45. Amsterdam/Philadelphia: John Benjamins.
- Diewald, Gabriele & Fischer, Kerstin. 1998. Zur diskursiven und modalen Funktion der Partikeln *aber*, *auch*, *doch* und *ja* in instruktionsdialogen. *Linguistica* 38(1). 75–99.
- Dong [董], Xiufang [秀芳]. 2003. “X 说”的词汇化 [On the Lexicalization of “X shuo”]. *语言科学* [*Language Sciences*] 2(2). 46–57.
- Dong [董], Xiufang [秀芳]. 2017. *词汇话：汉语双音词的衍生和发展* [*Lexicalization: The Origin and Evolution of Chinese Disyllabic Words*]. 修订版 [Revised Edition]. Beijing: 商务印书馆 [The Commercial Press].

- Dresner, Eli & Barak, Segev. 2006. Conversational multitasking in interactive written discourse as a communication competence. *Communication Reports* 19(1). 70–78.
- Du Bois, John W. 1987. The Discourse Basis of Ergativity. *Language* 63(4). 805–855.
- Du Bois, John W. 2003. Discourse and Grammar. In Michael Tomasello (ed.), *The new psychology of language: cognitive and functional approaches to language structure*, vol. 2, 47–87. Mahwah, New Jersey, London: Lawrence Erlbaum Associates Publishers.
- Erman, Britt & Kotsinas, Ulla-Britt. 1993. Pragmaticalization: The Case of *ba'* and *you know*. *Studier i Modern Språkvetenskap* 10. 76–93.
- Fagard, Benjamin & Blumenthal, Peter. 2020. Topic shifts in contrast: Ways to change the subject in French and German. *Journal of Pragmatics* 156. 206–222.
- Ferrara, Kathleen W. 1997. Form and function of the discourse marker *anyway*: implications for discourse analysis. *Linguistics* 35. 343–378.
- Fillmore, Charles J. 1988. The Mechanisms of “Construction Grammar.” In Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society, Berkeley, February 13-15, 1988*. 35–55. Berkeley, CA: Berkely Linguistics Society.
- Fillmore, Charles J. & Kay, Paul & O'Connor, Mary Catherine. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64(3). 501–538.
- Firbas, Jan. 1964. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1. 267–280.

References

- Firbas, Jan. 1975. On the thematic and the non-thematic section of the sentence. In Håkan Ringbom (ed.), *Style and text: studies presented to Nils Erik Enkvist*, 317–334. Stockholm: Språkförlaget Skriptor.
- Fischer, Kerstin. 2000. *From cognitive semantics to lexical pragmatics: the functional polysemy of discourse particles*. Berlin/Boston: Mouton de Gruyter.
- Fischer, Kerstin (ed.). 2006a. *Approaches to discourse particles* (Studies in Pragmatics 1). Amsterdam [u.a.]: Elsevier.
- Fischer, Kerstin. 2006b. Frames, constructions, and invariant meanings: the functional polysemy of discourse particles. In Kerstin Fischer (ed.), *Approaches to discourse particles* (Studies in Pragmatics 1), 427–447. Amsterdam [u.a.]: Elsevier.
- Fischer, Kerstin. 2007. Grounding and Common Ground: Modal Particles and their Translation equivalents. In Anita Fetzer & Kerstin Fischer (eds.), *Lexical Markers of Common Ground* (Studies in Pragmatics 3), 47–66. Amsterdam [u.a.]: Elsevier.
- Fischer, Kerstin. 2010. Beyond the sentence: Constructions, frames and spoken interaction. *Constructions and Frames* 2(2). 185–207.
- Fischer, Kerstin. 2015. Conversation, Construction Grammar, and Cognition. *Language and Cognition* 7. 563–588.
- Fischer, Kerstin & Alm, Maria. 2013. A radical construction grammar perspective on the modal particle-discourse particle distinction. In Liesbeth Degand & Bert Cornillie & Paola Pietrandrea (eds.), *Discourse Markers and Modal Particles: Categorization and Description* (Pragmatics & Beyond New Series 234), 47–87. Amsterdam/Philadelphia: John Benjamins.

- Fischer, Olga. 2007a. *Morphosyntactic change: functional and formal perspectives* (Oxford Surveys in Syntax and Morphology). Oxford [u.a.]: Oxford University Press.
- Fischer, Olga. 2007b. The development of English parentheticals: A case of grammaticalization? In Ute Smit, Stefan Dollinger, Julia Hüttner, Gunther Kaltenböck & Ursula Lutzky (eds.), *Tracing English through time: explorations in language variation: in honour of Herbert Schendl on the occasion of his 65th birthday* (Austrian Studies in English 95), 99–114. Wien: Braumüller.
- Fischer, Olga. 2008. On analogy as the motivation for grammaticalization. *Studies in Language* 32(2). 336–382.
- Fischer, Olga. 2018. Analogy: Its role in language learning, categorization, and in models of language change such as grammaticalization and constructionalization. In Sylvie Hancil, Tine Breban & José Antonio Vicente Lozano (eds.), *New Trends in Grammaticalization and Language Change* (Studies in Language Companion Series 202), 75–104. Amsterdam/Philadelphia: John Benjamins.
- Fox, Barbara A. & Thompson, Sandra A. 1990. A Discourse Explanation of the Grammar of Relative Clauses in English Conversation. *Language* 66(2). 297–316.
- Frank-Job, Barbara. 2006. A dynamic-interactional approach to discourse markers. In Kerstin Fischer (ed.), *Approaches to discourse particles* (Studies in Pragmatics 1), 359–374. Amsterdam [u.a.]: Elsevier.
- Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguistica Hungarica* 38(1/4). 19–33.
- Fraser, Bruce. 1990. An Approach to Discourse Markers. *Journal of Pragmatics* 14(3). 383–395.

References

- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6(2). 167–190.
- Fraser, Bruce. 1999. What Are Discourse Markers? *Journal of Pragmatics* 31(7). 931–952.
- Fraser, Bruce. 2001. The case of the empty S1. *Journal of Pragmatics* 3. 1625–1630.
- Fraser, Bruce. 2009. Topic Orientation Markers. *Journal of Pragmatics* 41(5). 892–898.
- Freud, Sigmund. 1905. *Jokes and their relation to the unconscious*. Reprint 2013. Worcestershire: Read Books Ltd.
- Fried, Mirjam. 2008. Constructions and constructs: mapping a shift between predication and attribution. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 194), 47–79. Berlin/Boston: Mouton de Gruyter.
- Fried, Mirjam. 2009a. Construction Grammar as a Tool for Diachronic Analysis. *Constructions and Frames* 1(2). 262–291.
- Fried, Mirjam. 2009b. Representing Contextual Factors in Language Change: Between Frames and Constructions. In Alexander Bergs & Gabriele Diewald (eds.), *Contexts and Constructions* (Constructional Approaches to Language 9), 63–94. Amsterdam/Philadelphia: John Benjamins.
- Fried, Mirjam. 2010. Grammar and interaction: New directions in constructional research. *Constructions and Frames* 2(2). 125–133.
- Fried, Mirjam. 2013. Principles of Constructional Change. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 419–437. Oxford [u.a.]: Oxford University Press.

- Fried, Mirjam & Östman, Jan-Ola. 2004. *Construction Grammar in a Cross-language Perspective* (Constructional Approaches to Language 2). Amsterdam/Philadelphia: John Benjamins.
- Fried, Mirjam & Östman, Jan-Ola. 2005. Construction Grammar and spoken language: The case of pragmatic particles. *Journal of Pragmatics* 37(11). 1752–1778.
- Garcia, Angela Cora & Jacobs, Jennifer Baker. 1999. The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. *Research on Language and Social Interaction* 32(4). 337–367.
- Gast, Volker. 2010. Contrastive Topics and Distributed Foci as Instances of Sub-Informativity: A Comparison of English and German. In Carsten Breul & Edward Göbbel (eds.), *Comparative and Contrastive Studies of Information Structure* (Linguistik Aktuell/Linguistics Today 165), 15–50. Amsterdam/Philadelphia: John Benjamins.
- Geluykens, Ronald. 1992. *From discourse process to grammatical construction: on left-dislocation in English* (Studies in Discourse and Grammar 1). Amsterdam/Philadelphia: John Benjamins.
- Gilquin, Gaëtanelle. 2015. Contrastive Collostructional Analysis: Causative Constructions in English and French. *Zeitschrift für Anglistik und Amerikanistik* 63(3). 253–272.
- Gisborne, Nikolas & Patten, Amanda. 2011. Construction grammar and grammaticalization. In Heiko Narrog & Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*, 92–104. Oxford [u.a.]: Oxford University Press.
- Givón, Talmy. 1993. *English Grammar: A Function-Based Introduction*. Vol. 2. Amsterdam/Philadelphia: John Benjamins.

References

- Givón, Talmy. 2018. *On Understanding Grammar*. Revised Edition. Amsterdam/Philadelphia: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago/London: University of Chicago Press.
- Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.
- Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford [u.a.]: Oxford University Press.
- Goldberg, Adele E. & van der Auwera, Johan. 2012. This Is to Count as a Construction. *Folia Linguistica* 46(1). 109–132.
- González-García, Francisco. 2017. Exploring inter-constructional relations in the constructicon: A view from Contrastive (Cognitive) Construction Grammar. In Francisco José Ruiz de Mendoza Ibáñez, Alba Luzondo Oyón & Paula Pérez Sobrino (eds.), *Constructing Families of Constructions: Analytical Perspectives and Theoretical Challenges* (Human Cognitive Processing 58), 135–172. Amsterdam/Philadelphia: John Benjamins.
- Greenbaum, Sidney. 1969. *Studies in English adverbial usage*. London/Harlow: Longman.
- Grice, H. Paul. 1975. Logic and Conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics 3: Speech acts*, 41–58. New York: Academic Press.
- Grosz, Barbara J. & Sidner, Candace L. 1985. Discourse Structure and the Proper Treatment of Interruptions. In Aravind Joshi (ed.), *Proceedings of the 9th international joint conference on Artificial*

- intelligence (IJCAI'85), Los Angeles, August 18-23, 1985*, vol. 2, 832–839. San Francisco: Morgan Kaufmann Publishers Inc.
- Günthner, Susanne & Imo, Wolfgang. 2006. *Konstruktionen in der Interaktion*. Berlin/Boston: Mouton de Gruyter.
- Günthner, Susanne & Mutz, Katrin. 2004. Grammaticalization v. pragmaticalization? The development of pragmatic markers in German and Italian. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components* (Trends in Linguistics. Studies and Monographs [TiLSM] 158), 77–107. Berlin/Boston: Mouton de Gruyter.
- Halliday, Michael A. K. 1985. *An introduction to functional grammar*. London [u.a.]: Arnold.
- Hancil, Sylvie. 2018. (Inter)subjectification and paradigmaticization: The case study of the final particle *but*. In Sylvie Hancil, Tine Breban & José Antonio Vicente Lozano (eds.), *New Trends in Grammaticalization and Language Change* (Studies in Language Companion Series 202), 291–313. Amsterdam/Philadelphia: John Benjamins.
- Hansen, Maj-Britt Mosegaard. 1998. *The Function of Discourse Particles : A Study with Special Reference to Spoken Standard French* (Pragmatics & Beyond New Series 53). Amsterdam/Philadelphia: John Benjamins.
- Haselow, Alexander. 2011. Discourse marker and modal particle: The functions of utterance-final *then* in spoken English. *Journal of Pragmatics* 43(14). 3603–3623.
- Haselow, Alexander. 2015. Final particles in spoken German. In Sylvie Hancil, Alexander Haselow & Margje Post (eds.), *Final Particles*

References

- (Trends in Linguistics. Studies and Monographs [TiLSM] 284), 77–108. Berlin/Boston: Mouton de Gruyter.
- Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline - the nature of grammaticalization* (Typological Studies in Language 59), 17–44. Amsterdam/Philadelphia: John Benjamins.
- Hayase, Naoko. 2011. The cognitive motivation for the use of dangling participles in English. In Klaus-Uwe Panther & Günter Radden (eds.), *Motivation in Grammar and the lexicon*, 89–106. Amsterdam/Philadelphia: John Benjamins.
- Heine, Bernd. 2013. On Discourse Markers: Grammaticalization, Pragmaticalization, or Something Else?. *Linguistics* 51(6). 1205–1247.
- Heine, Bernd. 2018. Are there two different ways of approaching grammaticalization? In Sylvie Hancil, Tine Breban & José Antonio Vicente Lozano (eds.), *New Trends in Grammaticalization and Language Change* (Studies in Language Companion Series 202), 23–54. Amsterdam/Philadelphia: John Benjamins.
- Heine, Bernd & Kaltenböck, Gunther & Kuteva, Tania & Long, Haiping. 2013. An Outline of Discourse Grammar. In Shannon Bischoff & Carmen Jany (eds.), *Functional Approaches to Language* (Trends in Linguistics. Studies and Monographs [TiLSM] 248), 155–206. Berlin/Boston: Mouton de Gruyter.
- Herring, Susan C. 1999. Interactional coherence in CMC. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, Maui, January 5-8, 1999*, 1–13. Maui, HI: IEEE Computer Society.

- Herring, Susan C. 2013. Relevance in computer-mediated conversation. In Susan C. Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of Computer-Mediated Communication* (Handbooks of Pragmatics 9), 245–268. Berlin/Boston: Mouton de Gruyter.
- Hilpert, Martin. 2006. A Synchronic Perspective on the Grammaticalization of Swedish Future Constructions. *Nordic Journal of Linguistics* 29(2). 151–172.
- Hilpert, Martin. 2008. *Germanic Future Constructions: A Usage-based Approach to Language Change* (Constructional Approaches to Language 7). Amsterdam/Philadelphia: John Benjamins.
- Hilpert, Martin. 2009. The German *mit*-predicative construction. *Constructions and Frames* 1(2). 29–55.
- Hilpert, Martin. 2013. *Constructional change in English: developments in allomorphy, word formation, and syntax*. Cambridge [u.a.]: Cambridge University Press.
- Hilpert, Martin. 2014. *Construction grammar and its application to English* (Edinburgh Textbooks on the English Language). Edinburgh: Edinburgh University Press.
- Hilpert, Martin. 2018. Three open questions in Diachronic Construction Grammar. In Evie Coussé, Peter Andersson & Joel Olofsson (eds.), *Grammaticalization meets construction grammar* (Constructional Approaches to Language 21), 21–39. Amsterdam/Philadelphia: John Benjamins.
- Hilpert, Martin & Östman, Jan-Ola. 2016. *Constructions Across Grammars* (Benjamins Current Topics 82). Amsterdam/Philadelphia: John Benjamins.

References

- Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components* (Trends in Linguistics. Studies and Monographs [TiLSM] 158), 21–42. Berlin/New York: Mouton de Gruyter.
- Höder, Steffen. 2012. Multilingual Constructions: A Diasystematic Approach to Common Structures. In Kurt Braunmüller & Christoph Gabriel (eds.), *Multilingual Individuals and Multilingual Societies* (Hamburg Studies on Multilingualism 13), 241–257. Amsterdam/Philadelphia: John Benjamins.
- Hoffmann, Thomas & Trousdale, Graeme (eds.). 2013. *The Oxford Handbook of Construction Grammar*. Oxford [u.a.]: Oxford University Press.
- Hopper, Paul. 1987. Emergent Grammar. In Jon Aske, Natasha Beery, Laura Michaelis and Hana Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, Berkeley, February 14-16, 1987*, 139–157. Berkeley, CA: Berkeley Linguistics Society.
- Hopper, Paul J. 1991. On Some Principles of Grammaticization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization II: Focus on Types of Grammatical Markers* (Typological Studies in Language 19), 17–35. Amsterdam/Philadelphia: John Benjamins.
- Hopper, Paul J. 1998. Emergent Grammar. In Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, vol. 2, 155–175. Mahwah, New York, London: Lawrence Erlbaum Associates Publishers.

- Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251–299.
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization* (Cambridge Textbooks in Linguistics). 2nd Edition. Cambridge [u.a.]: Cambridge University Press.
- Hu [胡], Yilu [以鲁]. 1923. 说国语后天发展 [Speaking of the development of the Chinese Language]. *国语学草创* [*Establishing Chinese Linguistics*], 51–57. Beijing: 商务印书馆 [The Commercial Press].
- Huddleston, Rodney D. & Pullum, Geoffrey K. 2002. *The Cambridge grammar of the English language*. Cambridge [u.a.]: Cambridge University Press.
- Hundt, Marianne. 2014. The Demise of the *Being to V* Construction. *Transactions of the Philological Society* 112(2). 167–187.
- Imo, Wolfgang. 2006. A Construction Grammar Approach to the Phrase *I mean* in Spoken English. *gidi Arbeitspapierreihe* 4. 1–39.
- Imo, Wolfgang. 2014. Appositions in monologue, increments in dialogue? On appositions and apposition-like patterns in spoken German and their status as constructions. In Ronny Boogaart, Timothy Colleman & Gijbert Rutten (eds.), *Extending the scope of construction grammar* (Cognitive Linguistics Research 54), 321–351. Berlin/Boston: Mouton de Gruyter.
- Israel, Michael. 1996. The way constructions grow. In Adele E. Goldberg (ed.), *Conceptual structure, discourse, and language*, 217–230. Stanford, CA: CSLI Publications.
- Izutsu, Katsunobu & Izutsu, Mitsuko Narita. 2013. From discourse markers to modal/final particles. What the position reveals about the continuum. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea

References

- (eds.), *Discourse Markers and Modal Particles: Categorization and Description* (Pragmatics & Beyond New Series 234), 217–235. Amsterdam/Philadelphia: John Benjamins.
- Janda, Richard D. 2001. Beyond “pathways” and “unidirectionality”: on the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* 23(2–3). 265–340.
- Jucker, Andreas H. 1993. The discourse Marker *well*: A relevance-theoretical account. *Journal of Pragmatics* 19. 435–452.
- Jucker, Andreas H. 1997. The discourse marker *well* in the history of English. *English Language & Linguistics* 1(1). 91–110.
- Jucker, Andreas H. & Ziv, Yael (eds.). 1998. *Discourse markers: descriptions and theory* (Pragmatics & Beyond New Series 57). Amsterdam/Philadelphia: John Benjamins.
- Kaltenböck, Gunther & Heine, Bernd. 2014. Sentence grammar vs. thetical grammar: Two competing domains? In Brian MacWhinney, Andrej L. Malchukov & Edith A. Moravcsik (eds.), *Competing motivations in grammar and usage* (Oxford Linguistics), 348–363. Oxford [u.a.]: Oxford University Press.
- Kaltenböck, Gunther & Heine, Bernd & Kuteva, Tania. 2011. On thetical grammar. *Studies in Language* 35(4). 852–897.
- Kay, Paul. 1984. The *Kind of/Sort of* Construction. In Claudia Brugman and Monica Macaulay (eds.), *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society, Berkeley, February 17-20, 1984*, 157–171. Berkeley, CA: Berkeley Linguistics Society.
- Kay, Paul. 2002. English Subjectless Tagged Sentences. *Language* 78(3). 458–481.

- Kay, Paul. 2004. Pragmatic aspects of grammatical constructions. In Laurence R. Horn & Gregory L. Ward (eds.), *The Handbook of Pragmatics* (Blackwell Handbooks in Linguistics 16), 675–700. Malden, Mass. [u.a.]: Blackwell.
- Kay, Paul & Fillmore, Charles J. 1999. Grammatical Constructions and Linguistic Generalizations: The *What's X Doing Y?* Construction. *Language* 75(1). 1–33.
- Kay, Paul & Michaelis, Laura A. 2012. Constructional Meaning and Compositionality. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning* (Handbooks of Linguistics and Communication Science 33/3), 2271–2296. Berlin/Boston: Mouton de Gruyter.
- Keizer, Evelien. 2016. The *(the) fact is (that)* construction in English and Dutch. In Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds.), *Outside the clause: Form and function of extra-clausal constituents* (Studies in Language Companion Series 178), 59–95. Amsterdam/Philadelphia: John Benjamins.
- Kong, Lei & Qin, Hongwu. 2017. The development of manner of speaking markers in English and Chinese: pragmaticalization, grammaticalization and lexicalization. *Journal of Pragmatics* 107. 16–30.
- König, Ekkehard & van der Auwera, Johan. 1990. Adverbial participles, gerunds and absolute constructions in the languages of Europe. In Johannes Bechert, Giuliano Bernini & Claude Buridant (eds.), *Toward a typology of European languages* (Empirical Approaches to Language Typology [EALT] 8), 337–355. Berlin/New York: Mouton de Gruyter.

References

- Kortmann, Bernd. 1991. *Free adjuncts and absolutes in English: Problems of control and interpretation*. London [u.a.]: Routledge.
- Kortmann, Bernd & König, Ekkehard. 1992. Categorical reanalysis: the case of deverbal prepositions. *Linguistics* 30(4). 671–698.
- Krug, Manfred. 1998. British English is Developing a New Discourse Marker, *innit?* - A Study in Lexicalisation Based on Social, Regional and Stylistic Variation. *Arbeiten aus Anglistik und Amerikanistik* 23(2). 145–197.
- Labov, William. 1972. *Sociolinguistic patterns* (Conduct and Communication Series 4). Philadelphia: University of Pennsylvania Press.
- Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago, Ill. [u.a.]: University of Chicago Press.
- Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus, and the mental representations of discourse referents* (Cambridge Studies in Linguistics 71). Cambridge [u.a.]: Cambridge University Press.
- Langacker, Ronald W. 1985. Observations and speculations on subjectivity. In John Haiman (ed.), *Iconicity in Syntax: Proceedings of a Symposium on Iconicity in Syntax, Stanford, June 24-26, 1983* (Typological Studies in Language 6), 109–150. Amsterdam/Philadelphia: John Benjamins.
- Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1(1). 5–38.
- Langacker, Ronald W. 1997. Consciousness, Construal, and Subjectivity. In Maxim I. Stamenov (ed.), *Language Structure, Discourse and the*

- Access to Consciousness* (Advances in Consciousness Research 12), 49–75. Amsterdam/Philadelphia: John Benjamins
- Langacker, Ronald W. 2002. Deixis and Subjectivity. In Frank Brisard (ed.), *Grounding: The Epistemic Footing of Deixis and Reference* (Cognitive Linguistics Research 21), 1–28. Berlin/New York: Mouton de Gruyter.
- Langacker, Ronald W. 2005. Integration, grammaticization, and constructional meaning. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical Constructions: Back to the roots* (Constructional Approaches to Language 4), 157–189. Amsterdam/Philadelphia: John Benjamins.
- Langacker, Ronald W. 2006. Subjectification, grammaticization, and conceptual archetypes. In Angeliki Athanasiadou, Costas Canakis & Bert Cornillie (eds.), *Subjectification: Various Paths to Subjectivity* (Cognitive Linguistics Research 31), 17–40. Berlin/New York: Mouton de Gruyter.
- Lauwers, Peter & Vanderbauwhede, Gudrun & Verleyen, Stijn (eds.). 2012. *Pragmatic Markers and Pragmaticalization: Lessons from false friends*. Amsterdam/Philadelphia: John Benjamins.
- Leech, Geoffrey N. 1974. *Semantics*. Harmondsworth [u.a.]: Penguin.
- Lehmann, Christian. 1995. *Thoughts on Grammaticalization* (LINCOM Studies in Theoretical Linguistics 1). Revised and expanded version. München/Newcastle: LINCOM EUROPA.
- Lehmann, Christian. 2002. New reflections on grammaticalization and lexicalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization* (Typological Studies in Language 49), 1–18. Amsterdam/Philadelphia: John Benjamins.

References

- Lehmann, Christian. 2015. *Thoughts on Grammaticalization* (Classics in Linguistics 1). 3rd Edition. Berlin: Language Science Press.
- Leino, Pentti & Östman, Jan-Ola. 2008. Language change, variability, and functional load: Finnish genericity from a constructional point of view. In Jaakko Leino (ed.), *Constructional reorganization* (Constructional Approaches to Language 5), 37–54. Amsterdam/Philadelphia: John Benjamins.
- Lenk, Uta. 1998a. Discourse markers and global coherence in conversation. *Journal of Pragmatics* 30(2). 245–257.
- Lenk, Uta. 1998b. *Marking discourse coherence: functions of discourse markers in spoken English* (Language in Performance 15). Tübingen: Narr.
- Lewis, Diana M. 2007. From temporal to contrastive and causal: The emergence of connective *after all*. In Agnès Celle & Ruth Huart (eds.), *Connectives as discourse landmarks* (Pragmatics & Beyond New Series 161), 89–99. Amsterdam/Philadelphia: John Benjamins.
- Lewis, Diana M. 2011. A discourse-constructional approach to the emergence of discourse markers in English. *Linguistics* 9(2). 415–443.
- Lewis, Diana M. 2018. Grammaticalizing connectives in English and discourse information structure. In Sylvie Hancil, Tine Breban & José Antonio Vicente Lozano (eds.), *New Trends in Grammaticalization and Language Change* (Studies in Language Companion Series 202), 135–157. Amsterdam/Philadelphia: John Benjamins.
- Li [李], Bingzhen [秉震]. 2009. “说”类话题转换标记的语义演变 [The semantic evolution of the “shuo”-type topic change markers]. *中国语文* [Chinese Literature and Language] 32(5). 419–428, 479–489.

- Li [李], Bingzhen [秉震] & Zhang [张], Quansheng [全生]. 2012. “说到”的话语功能分析 [Discourse functional analyses of “shuodao”]. *Studies in Language and Linguistics* [语言研究] 32(4). 27–34.
- Li [李], Yanyan [艳艳]. 2010. 论话题标记“要说” [On the Topic Marker “yaoshuo”]. *唐山师范学院学报* [Journal of Tangshan Teacher's College] 32(1). 25–27.
- Li [李], Zongjiang [宗江]. 2007. 话题标引成分“要说”的由来和去向 [Origin and development of the topic marker “yaoshuo”]. In Jiaxuan [家煊] Shen [沈], Fuxiang [福祥] Wu [吴] & Zongjiang [宗江] Li [李] (eds.), *语法化与语法研究* [Grammaticalization and Grammar Studies], vol. 3, 86–105. Shanghai: 商务印书馆 [The Commercial Press].
- Li [李], Zongjiang [宗江]. 2017. 近代汉语的话题标记及其演变 [Topic Markers in Modern Chinese and Their Development]. *汉语学报* [Chinese Linguistics] 60. 36–44.
- Lightfoot, Douglas. 2011. Grammaticalization and lexicalization. In Heiko Narrog & Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*, 438–461. Oxford [u.a.]: Oxford University Press.
- Lin, Hsueh-O. 2004. Grammaticalization of the Reported Speech Frames *Wo shuo* and *Ni shuo* in Mandarin Conversation. *Concentric: Studies in Linguistics* 30(2). 157–182.
- Linell, Per. 2009. Grammatical Constructions in Dialogue. In Alexander Bergs & Gabriele Diewald (eds.), *Contexts and Constructions* (Constructional Approaches to Language 9), 97–110. Amsterdam/Philadelphia: John Benjamins.

References

- Liu [刘], Fang [芳]. 2009. 析言说义动词“说”的语法化 [Grammaticalization of the Speech Act Verb “shuo”]. *语文学刊* [Journal of Literature and Language] 3. 50–52.
- Liu [刘], Qin [嵇]. 2008. “我说”的语义演变及其主观化 [Semantic change and subjectification of “woshuo”]. *语文研究* [Linguistic Research] 108. 18–23.
- Lu [鲁], Jinfang [瑾芳]. 2016. 话语标记“话说”的关联认知解读 [Analysis of the discourse marker “huashuo” in the relevance theory]. *语文学刊* [Journal of Literature and Language] 9. 24–26.
- Markman, Kris M. 2005. To send or not to send: Turn construction in computer-mediated chat. *Texas Linguistic Forum (TLF): Proceedings of the Twelfth annual Symposium About Language and Society, Austin, April 16-18, 2004 (SALSA XII Proceedings)* (Texas Linguistic Forum 48), 115–124. Austin: University of Texas Press.
- Markman, Kris M. 2013. Conversational coherence in small group chat. In Susan C. Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of Computer-Mediated Communication (Handbooks of Pragmatics 9)*, 539–564. Berlin/Boston: Mouton de Gruyter.
- McEnery, Tony & Wilson, Andrew. 2001. *Corpus linguistics: an introduction* (Edinburgh Textbooks in Empirical Linguistics). 2nd Edition. Edinburgh: Edinburgh University Press.
- Michaelis, Laura A. 2013. Sign-based Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 133–152. Oxford [u.a.]: Oxford University Press.
- Michaelis, Laura A. & Lambrecht, Knud. 1996. Toward a Construction-Based Theory of Language Function: The Case of Nominal Extraposition. *Language* 72(2). 215–247.

- Mithun, Marianne. 2008. Borrowed rhetorical constructions as starting points for grammaticalization. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 194), 195–230. Berlin/New York: Mouton de Gruyter.
- Moreno Cabrera, Juan C. 1998. On the relationships between grammaticalization and lexicalization. In Anna Giacalone-Ramat & Paul J. Hopper (eds.), *The limits of grammaticalization* (Typological Studies in Language 37), 211–227. Amsterdam/Philadelphia: John Benjamins.
- Mosegaard Hansen, Maj-Britt. 2008. *Particles at the Semantics/pragmatics Interface: Synchronic and Diachronic Issues: A Study with Special Reference to the French Phasal Adverbs* (Current Research in the Semantics/Pragmatics Interface 19). Oxford [u.a.]: Elsevier.
- Müller, Simone. 2005. *Discourse markers in native and non-native English discourse* (Pragmatics & Beyond New Series 138). Amsterdam/Philadelphia: John Benjamins.
- Narrog, Heiko. 2017. Three types of subjectivity, three types of intersubjectivity, their dynamicization and a synthesis. In Daniël van Olmen, Hubert Cuyckens & Lobke Ghesquière (eds.), *Aspects of Grammaticalization: (Inter)Subjectification and Directionality* (Trends in Linguistics. Studies and Monographs 305), 19–46. Berlin/Boston: Mouton de Gruyter.
- Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, Mass. [u.a.]: The MIT Press.
- Noël, Dirk. 2007. Diachronic Construction Grammar and Grammaticalization Theory. *Functions of Language* 14(2). 177–202.

References

- Nođ, Dirk & Coleman, Timothy. 2009. The Nominative and Infinitive in English and Dutch: An Exercise in Contrastive Diachronic Construction Grammar. *Languages in Contrast* 9(1). 144–181.
- Nođ, Dirk & Coleman, Timothy. 2010. Believe-Type Raising-to-Object and Raising-to-Subject Verbs in English and Dutch: A Contrastive Investigation in Diachronic Construction Grammar. *International Journal of Corpus Linguistics* 15(2). 157–182.
- Nøke, Henning. 2006. The semantics of polyphony (and the pragmatics of realization). *Acta Linguistica Hafniensia* 38(1). 137–160.
- Norde, Muriel. 2001. Deflexion as a counterdirectional factor in grammatical change. *Language Sciences* 23(2–3). 231–264.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford [u.a.]: Oxford University Press.
- Norde, Muriel & Trousdale, Graeme. 2016. Exaptation from the perspective of construction morphology. In Muriel Norde & Freek van de Velde (eds.), *Exaptation and Language Change* (Current Issues in Linguistic Theory 336), 163–195. Amsterdam/Philadelphia: John Benjamins.
- Norrick, Neal R. 2001. Discourse markers in oral narrative. *Journal of Pragmatics* 33(6). 849–878.
- Nuyts, Jan. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33(3). 383–400.
- Nuyts, Jan. 2014. Notions of (inter)subjectivity. In Lieselotte Brems, Lobke Ghesquière & Freek Van de Velde (eds.), *Intersubjectivity and Intersubjectification in Grammar and Discourse: Theoretical and descriptive advances* (Benjamins Current Topics 65), 53–76. Amsterdam/Philadelphia: John Benjamins

- Onodera, Noriko O. 2014. Setting Up a Mental Space: A Function of Discourse Markers at the Left Periphery (LP) and Some Observations about LP and RP in Japanese. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12), 92–116. Leiden/Boston: Brill.
- Östman, Jan-Ola. 1981. *You Know: A Discourse-Functional Approach* (Pragmatics & Beyond II:7). Amsterdam/Philadelphia: John Benjamins.
- Östman, Jan-Ola. 2005. Construction Discourse: A prolegomenon. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive Grounding and Theoretical Extensions* (Constructional Approaches to Language 3), 121–144. Amsterdam/Philadelphia: John Benjamins.
- Östman, Jan-Ola. 2006. Constructions in Cross-Language Research: Verbs as Pragmatic Particles in Solv. In Karin Aijmer & Anne-Marie Simon-Vandenberghe (eds.), *Pragmatic markers in contrast*, 237–257. Amsterdam [u.a.]: Elsevier.
- Overstreet, Maryann. 1999. *Whales, Candlelight, and Stuff Like That: General Extenders in English Discourse* (Oxford Studies in Sociolinguistics). New York and Oxford: Oxford University Press.
- Paolillo, John C. & Zelenkauskaitė, Asta. 2013. Real-time chat. In Susan C. Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of Computer-Mediated Communication* (Handbooks of Pragmatics 9), 109–133. Berlin/Boston: Mouton de Gruyter.
- Pedersen, Johan. 2008. The Construction of Macro-events: A Typological Perspective. In Chris Butler & Javier M. Arista (eds.),

References

- Deconstructing Constructions*, 25–62. Amsterdam/Philadelphia: John Benjamins.
- Pedersen, Johan. 2009. Lexical and constructional organization of argument structure: a contrastive analysis. In Jordan Zlatev, Mats Andr  n, Marlene Johansson Falck & Carita Lundmark (eds.), *Studies in Language and Cognition*, 241–256. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Pedersen, Johan. 2013. The Way-construction and Cross-linguistic Variation in Syntax: Implications for Typological Theory. In Carita Paradis, Jean Hudson & Ulf Magnusson (eds.), *The construal of spatial meaning: windows into conceptual space* (Explorations in Language and Space 7), 236–262. Oxford [u.a.]: Oxford University Press.
- Peng, Rui. 2017. *Pivotal Constructions in Chinese: Diachronic, synchronic, and constructional perspectives* (Studies in Chinese Language and Discourse 8). Amsterdam/Philadelphia: John Benjamins.
- Perek, Florent. 2015. *Argument Structure in Usage-Based Construction Grammar: Experimental and corpus-based perspectives* (Constructional Approaches to Language 17). Amsterdam/Philadelphia: John Benjamins.
- Petr  , Peter & Cuyckens, Hubert. 2008. Bedusted, yet not beheaded: The role of *be*-'s constructional properties in its conservation. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 194), 133–169. Berlin/New York: Mouton de Gruyter.
- Pirc, Ana Monika. 2013. Construction Grammar and “Non-native Discourse.” *Theory and Practice in English Studies* 1(1). 55–73.
- Pons Border  , Salvador. 2018. Paths of Grammaticalization: Beyond the LP/RP Debate. In Salvador Pons Border   &   scar Loureda Lamas

- (eds.), *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change* (Studies in Pragmatics 18), 334–383. Leiden/Boston: Brill.
- Pons Border á, Salvador & Estellés Arguedas, Maria. 2009. Expressing digression linguistically: Do digressive markers exist? *Journal of Pragmatics* 41(5). 921–936.
- Pons Border á, Salvador & Loureda Lamas, Óscar (eds.). 2018. *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change* (Studies in Pragmatics 18). Leiden/Boston: Brill.
- Prevost, Sophie. 2011. *A propos* from Verbal Complement to Discourse Marker: A Case of Grammaticalization? *Linguistics* 49(2). 391–413.
- Prince, Ellen F. 1981. Toward a Taxonomy of Given-New Information. In Peter Cole (ed.), *Radical pragmatics*, 223–255. New York [u.a.]: Academic Press.
- Proost, Kristel. 2007. *Conceptual Structure in Lexical Items: The Lexicalisation of Communication Concepts in English, German and Dutch* (Pragmatics & Beyond New Series 168). Amsterdam/Philadelphia: John Benjamins.
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1985. *A comprehensive grammar of the English language*. London/Harlow: Longman.
- Ramat, Paolo. 1992. Thoughts on Degrammaticalization. *Linguistics* 30. 549–560.
- Redeker, Gisela. 1990. Ideational and Pragmatic Markers of Discourse Structure. *Journal of Pragmatics* 14(3). 367–381.

References

- Reig Alamillo, Assela. 2007. El digresor *por cierto* y la digresión. *Oralia* 10. 233–254.
- Rostila, Jouni. 2004. Lexicalization as a Way to Grammaticalization. In Fred Karlsson (ed.), *Proceedings of the 20th Scandinavian Conference of Linguistics, Helsinki, January 7-9, 2004*, 1–13. Helsinki: University of Helsinki Press.
- Rostila, Jouni. 2006. Storage as a way to grammaticalization. *Constructions* 1(1). 1–59.
- Roulet, Eddy. 1984. Speech acts, discourse structure, and pragmatic connectives. *Journal of Pragmatics* 8(1). 31–47.
- Sacks, Harvey & Schlegoff, Emanuel A. & Jefferson, Gail. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.
- Schank, Roger C. 1977. Rules and Topics in Conversation. *Cognitive Science* 1(4). 421–441.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge [u.a.]: Cambridge University Press.
- Schourup, Lawrence C. 1982. *Common discourse particles in English conversation* (Working Papers in Linguistics 28). Columbus: Ohio State University Press.
- Schourup, Lawrence C. 1999. Discourse Markers. *Lingua* 107(3–4). 227–265.
- Schutz, Alfred. 1970. *Reflections on the Problem of Relevance*. New Haven and London: Yale University Press.
- Schwenter, Scott. 2000. Viewpoints and polysemy: Linking adversative and causal meanings of discourse markers. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – Condition – Concession –*

- Contrast: Cognitive and Discourse Perspectives* (Topics in English Linguistics 33), 257–281. Berlin/New York: Mouton de Gruyter.
- Schwenter, Scott A. & Traugott, Elizabeth Closs. 1995. The Semantic and Pragmatic Development of Substantive Complex Prepositions in English. In Andreas H. Jucker (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English* (Pragmatics & Beyond New Series 35), 243–273. Amsterdam/Philadelphia: John Benjamins.
- Schwenter, Scott A. & Traugott, Elizabeth Closs. 2000. Invoking scalarity: The development of *in fact*. *Journal of Historical Pragmatics* 1(1). 7–25.
- Searle, John R. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, John R. 1976. A Classification of Illocutionary Acts. *Language in Society* 5(1). 1–23.
- Shi [施], Lingli [玲丽]. 2013. “话说”的语义演变及主观化 [Semantic change and subjectification of “huashuo”]. *语文学刊 [Journal of Literature and Language]* 2. 45–46, 57.
- Shi [石], Yuzhi [毓智]. 2011. *语法化理论——基于汉语发展的历史 [Grammaticalization Theory: with special reference to Chinese]* (外教社认知语言学丛书(应用系列) [SFLEP Cognitive Linguistics Series (Practical Series)]). Shanghai: 上海外语教育出版社 [Shanghai Foreign Language Education Press].
- Simon-Vandenberg, Anne-Marie & Willems, Dominique. 2011. Cross-linguistic Data as Evidence in the Grammaticalization Debate: The Case of Discourse Markers. *Linguistics* 49(2). 333–364.

References

- Smirnova, Elena. 2015. Constructionalization and constructional change: The role of context in the development of constructions. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic construction grammar* (Constructional Approaches to Language 18), 81–106. Amsterdam/Philadelphia: John Benjamins.
- Sohn, Sung-Ock & Kim, Stephanie Hyeri. 2014. The Interplay of Discourse and Prosody at the Left and Right Periphery in Korean: An Analysis of *Kuntŭ* “but.” In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change* (Studies in Pragmatics 12), 221–250. Leiden/Boston: Brill.
- Sommerer, Lotte. 2015. The influence of constructions in grammaticalization: Revisiting category emergence and the development of the definite article in English. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic construction grammar* (Constructional Approaches to Language 18), 107–137. Amsterdam/Philadelphia: John Benjamins.
- Sommerer, Lotte & Smirnova, Elena (eds.). 2020. *Nodes and Networks in Diachronic Construction Grammar* (Constructional Approaches to Language 27). Amsterdam/ Philadelphia: John Benjamins.
- Sperber, Dan & Wilson, Deirdre. 1986. *Relevance: communication and cognition*. Oxford: Blackwell.
- Steen, Francis & Turner, Mark. 2013. Multimodal Construction Grammar. In Michael Borkent, Barbara Dancygier & Jennifer Hinnell (eds.), *Language and the Creative Mind*, 1–20. Stanford, CA: CSLI Publications.

- Stefanowitsch, Anatol. 2020. *Corpus linguistics: A guide to the methodology* (Textbooks in Language Sciences 7). Berlin: Language Science Press.
- Stefanowitsch, Anatol & Gries, Stefan Th. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Sun [孙], Hongwei [洪威]. 2013. 反预期标记“别说” [Counter-expectation Marker “bieshuo”]. *社科纵横 (新理论版)* [Social Sciences Review (New Theories Edition)] 1. 347–349.
- Sun [孙], Kaidi [楷第]. 1953. 说话考 [The Study of Narration]. *论中国短篇白话小说* [Studies of Mandarin Chinese Short Stories] (中国古典文学研究丛刊 [Chinese Classical Literature Studies Series]), 37–41. Shanghai: 棠棣出版社 [Tangdi Publishing].
- Sun [孙], Liping [利萍]. 2017. 现代汉语言说类话语标记研究 [The study on speaking discourse markers in modern Chinese] (华侨大学哲学社会科学文库：文学系列 [Huaqiao University Library of Philosophical and Social Sciences: Literary Studies Series]). Beijing: 社会科学文献出版社 [Social Sciences Academic Press].
- Swan, Toril & Breivik, Leiv Egil. 2011. English Sentence Adverbials in a Discourse and Cognitive Perspective. *English Studies* 92(6). 679–692.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics. Vol. II: Typology and Process in Concept Structuring*. Cambridge, Mass. [u.a.]: The MIT Press.
- Talmy, Leonard. 2007. Lexical Typologies. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, vol. III: Grammatical Categories and the Lexicon*, 66–168. Cambridge [u.a.]: Cambridge University Press.

References

- Tanno, Koji. 2017. Expanding to the peripheries: A corpus-based study of the development of the Japanese discourse marker *toiuka*. *Ampersand* 4. 58–72.
- Thompson, Sandra A. & Mulac, Anthony. 1991. A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization II: Focus on Types of Grammatical Markers* (Typological Studies in Language 19), 313–329. Amsterdam/Philadelphia: John Benjamins.
- Tian [田], Yuan [源]. 2007. 汉语“说”类动词研究 [A study of the *shuo*-type verbs in Chinese]. Wuhan: 华中师范大学 [Central China Normal University]. (Master's Thesis).
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics* (Current issues in linguistic theory 24), 245–271. Amsterdam/Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs. 1989. On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. *Language* 65(1). 31–55.
- Traugott, Elizabeth Closs. 1995a. The role of the development of discourse markers in a theory of grammaticalization. (Presented at the 12th International Conference on Historical Linguistics (ICHL XII), Manchester.)
- Traugott, Elizabeth Closs. 1995b. Subjectification in Grammaticalisation. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 31–54. Cambridge [u.a.]: Cambridge University Press.

- Traugott, Elizabeth Closs. 1997. Subjectification and the development of epistemic meaning: the case of *promise* and *threaten*. In Toril Swan & Olaf Jansen Westvik (eds.), *Modality in Germanic Languages: Historical and Comparative Perspectives* (Trends in Linguistics. Studies and Monographs [TiLSM] 99), 185–210. Berlin/New York: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 1999. The rhetoric of counter-expectation in semantic change: a study in subjectification. In Andreas Blank & Peter Koch (eds.), *Historical semantics and cognition* (Cognitive Linguistics Research 13), 177–196. Berlin/New York: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for Language Change*, 124–139. Cambridge [u.a.]: Cambridge University Press.
- Traugott, Elizabeth Closs. 2008a. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of Degree Modifiers in English. In Regine Eckhardt, Gerhard Jäger & Tonjes Veenstra (eds.), *Variation, selection, development probing the evolutionary model of language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 197), 219–250. Berlin/New York: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 2008b. The grammaticalization of NP of NP patterns. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs [TiLSM] 194), 23–45. Berlin/New York: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, Intersubjectification and*

References

- Grammaticalization (Topics in English Linguistics 66), 29–71. Berlin/New York: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 2012. Intersubjectification and clause periphery. *English Text Construction* 5(1). 7–28.
- Traugott, Elizabeth Closs. 2015a. Toward a coherent account of grammatical constructionalization. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic construction grammar* (Constructional Approaches to Language 18), 51–79. Amsterdam/Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs. 2015b. Investigating “periphery” from a functionalist perspective. *Linguistics Vanguard* 1(1). 119–130.
- Traugott, Elizabeth Closs. 2020. The development of “digressive” discourse-topic shift markers in English. *Journal of Pragmatics* 156. 121–135.
- Traugott, Elizabeth Closs & Dasher, Richard B. 2002. *Regularity in semantic change* (Cambridge Studies in Linguistics 97). Cambridge [u.a.]: Cambridge University Press.
- Traugott, Elizabeth Closs & König, Ekkehard. 1991. The Semantics-Pragmatics of Grammaticalization Revisited. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization II: Focus on Types of Grammatical Markers* (Typological Studies in Language 19), 189–218. Amsterdam/Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs & Trousdale, Graeme. 2013. *Constructionalization and constructional changes* (Oxford Studies in Diachronic and Historical Linguistics 6). Oxford [u.a.]: Oxford University Press.
- Trousdale, Graeme. 2008a. Constructions in grammaticalization and lexicalization: Evidence from the history of a composite predicate construction in English. In Graeme Trousdale & Nikolas Gisborne

- (eds.), *Constructional Approaches to English Grammar* (Topics in English Linguistics 57), 33–67. Berlin/New York: Mouton de Gruyter.
- Trousdale, Graeme. 2008b. Words and constructions in grammaticalization: The end of the English impersonal construction. In Susan M. Fitzmaurice & Donka Minkova (eds.), *Studies in the History of the English Language IV: Empirical and Analytical Advances in the Study of English Language Change* (Topics in English Linguistics 61), 301–326. Berlin/New York: Mouton de Gruyter.
- Trousdale, Graeme. 2008c. A constructional approach to lexicalization processes in the history of English: Evidence from possessive constructions. *Word Structure* 1(2). 156–177.
- Trousdale, Graeme. 2010. Issues in Constructional Approaches to Grammaticalization in English. In Katerina Stathi, Elke Gehweiler & Ekkehard König (eds.), *Grammaticalization: Current Views and Issues* (Studies in Language Companion Series 119). Amsterdam/Philadelphia: John Benjamins.
- Trousdale, Graeme. 2012. Grammaticalization, constructions and the grammaticalization of constructions. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and Language Change: New Reflections* (Studies in Language Companion Series 130), 167–198. Amsterdam/Philadelphia: John Benjamins.
- Trousdale, Graeme. 2014. On the relationship between grammaticalization and constructionalization. *Folia Linguistica* 48(2). 557–577.
- Tsujimura, Natsuko. 2005. A constructional approach to mimetic verbs. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical Constructions:*

References

- Back to the roots* (Constructional Approaches to Language 4), 137–156. Amsterdam/Philadelphia: John Benjamins.
- Urgelles-Coll, Miriam. 2010. *The syntax and semantics of discourse markers* (Continuum Studies in Theoretical Linguistics). London [u.a.]: Continuum.
- Van de Velde, Freek. 2014. Degeneracy: The maintenance of constructional networks. In Ronny Boogaart, Timothy Coleman & Gijsbert Rutten (eds.), *Extending the scope of construction grammar* (Cognitive Linguistics Research 54), 141–179. Berlin/Boston: Mouton de Gruyter.
- Van der Auwera, Johan. 2002. More thoughts on degrammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization* (Typological Studies in Language 49), 19–29. Amsterdam/Philadelphia: John Benjamins.
- Van Dijk, Teun A. 1979. Pragmatic connectives. *Journal of Pragmatics* 3(5), 447–456.
- Van Eynde, Frank. 2007. The Big Mess Construction. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford, July 20–22, 2007*, 415–433. Stanford, CA: CSLI Publications.
- Virtanen, Tuija. 2004. Point of departure: cognitive aspects of sentence-initial adverbials. In Tuija Virtanen (ed.), *Approaches to cognition through text and discourse* (Trends in Linguistics. Studies and Monographs [TiLSM] 147), 79–98. Berlin/New York: Mouton de Gruyter.
- Waltereit, Richard. 2012. On the origins of grammaticalization and other types of language change in discourse strategies. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and Language Change: New Reflections*

- (Studies in Language Companion Series 130), 51–72. Amsterdam/Philadelphia: John Benjamins.
- Wang, Yan. 2011. A Discourse-Pragmatic Functional Study of the Discourse Markers: Japanese *Ano* and Chinese *Nage*. *Intercultural Communication Studies* 20(2). 41–61.
- Wang [王], Li [力]. 1998. 古代汉语 [Classical Chinese]. Reprint (2001). Vol. 1. Beijing: 中华书局 [Zhonghua Book Company].
- Wang [王], Li [力]. 2013. 汉语史稿 [Manuscripts on the History of the Chinese Language] (王力文集 [Collected Works of Li Wang] 1). Beijing: 中华书局 [Zhonghua Book Company].
- Warner, Richard George. 1985. *Discourse connectives in English*. New York/London: Garland Publishing.
- Watts, Richard J. 1988. A Relevance-Theoretic Approach to Commentary Pragmatic Markers: the Case of *Actually*, *Really* and *Basically*. *Acta Linguistica Hungarica* 38(1–4). 235–260.
- Wiemer, Björn & Bisang, Walter. 2004. What makes grammaticalization? An appraisal of its components and its fringes. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components* (Trends in Linguistics. Studies and Monographs [TiLSM] 158), 3–20. Berlin/New York: Mouton de Gruyter.
- Wilson, Deirdre & Sperber, Dan. 1993. Linguistic Form and Relevance. *Lingua* 90. 1–25.
- Wischer, Ilse. 2000. Grammaticalization versus lexicalization: “Methinks” there is some confusion. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in*

- English* (Studies in Language Companion Series 53), 355–370. Amsterdam/Philadelphia: John Benjamins.
- Xian [鲜], Lixia [丽霞]. 2012. 话语标记“你说”的语境及其功能 [Contexts and functions of discourse marker “nishuo”]. *四川师范大学学报: 社会科学版* [*Journal of Sichuan Normal University (Social Science)*] 5. 50–56.
- Xuan [玄], Yue [玥]. 2011. “说”的一种新用法——客观叙述标记词 [On a new usage of “shuo” – marker of objective narration]. *汉语学报* [*Chinese Linguistics*] 34. 28–35.
- Yao [姚], Zhanlong [占龙]. 2008. “说、想、看”的主观化及其诱因 [The Subjectification of “shuo, xiang, kan” and the Motivation]. *语言教学与研究* [*Language Teaching and Linguistic Studies*] 5. 47–53.
- Yin [尹], Hailiang [海良]. 2009a. 自然会话中“我说”的语用标记功能 [Pragmatic marking functions of “woshuo” in natural discourse]. *修辞学习* [*Rhetoric Learning*] 1. 45–50.
- Yin [尹], Hailiang [海良]. 2009b. 现代汉语“别说”的篇章衔接功能及其语法化 [Discourse linking functions of “bieshuo” in modern Chinese and its grammaticalization]. *西南农业大学学报 (社会科学版)* [*Journal of Southwest University (Social Science)*] 4. 111–116.
- Yin [尹], Hailiang [海良]. 2009c. 自然口语中的话语标记“别说” [Discourse marker “bieshuo” in natural speech]. *宁夏大学学报 (人文社会科学版)* [*Journal of Ningxia University (Humanities and Social Science)*] 31(6). 56–61.
- Yin [殷], Shulin [树林]. 2012. *现代汉语话语标记研究* [*Studies of discourse markers in modern Chinese*]. Beijing: 中国社会科学出版社 [China Social Science Press].

- Yu [喻], Wei [薇] & Yao [姚], Shuangyun [双云]. 2018. 言说动词“说”的语法化考察 [Studies of the grammaticalization of the speech act verb “shuo”]. *湖北师范大学学报(哲学社会科学版)* [*Journal of Hubei Normal University (Philosophy and Social Science)*] 38(2). 7–16.
- Zhang, Liulin. 2018. The Chinese Notional Passive Construction under the View of Cognitive Construction Grammar. *Acta Linguistica Asiatica* 8(2). 83–110.
- Zhang [张], Jinwen [苓文]. 2016. 论“话说”的话语标记功能——以网络论坛及贴吧语言为例 [On the discourse marking functions of “huashuo”: a case study of language used in online forums and post bars]. *现代语文* [*Modern Chinese Literature and Language*] 4. 122–125.
- Zhang [张], Qingqing [青青]. 2012. 言及类词语及其标记的话题 [*The reference words and the topics they marked - Take “shuodao” “shuoqi” “tiqi” for example*]. Ningbo: Ningbo University. (Master’s Thesis).
- Zhou [周], Chenlei [晨磊]. 2012. 从语篇到人际——“话说”的意义和功能演变 [From Textual to Interpersonal: the Development of the Meaning and Function of “Huashuo”]. *语言科学* [*Linguistic Sciences*] 11(5). 499–508.
- Zhou [周], Li [莉]. 2013. “别说”类语用标记来源探讨 [On the Origin of the Pragmatic Marker “bieshuo”]. *汉语学报* [*Chinese Linguistics*] 2. 48–56.
- Ziegeler, Debra. 2015. *Converging Grammars: Constructions in Singapore English*. Berlin/Boston: Mouton de Gruyter.
- Zima, Elisabeth & Bergs, Alexander. 2017. Multimodality and Construction Grammar. *Linguistics Vanguard* 3(s1). 1–9.

References

- Zipf, George Kingsley. 1949. *Human Behavior and the Principle of Least Effort: an Introduction to Human Ecology*. Oxford: Addison-Wesley Press.
- Zwicky, Arnold M. 2002. I Wonder What Kind of Construction That This Example Illustrates. In Luis D. David Beaver, Casillas Martínez, Brady Z. Clark & Stefan Kaufmann (eds.), *The construction of meaning*, 219–248. Stanford, CA: CSLI Publications.

This thesis takes an interest in the emergence and development of discourse markers. It develops within the theoretical framework of Construction Grammar and addresses the following questions: (1) Which consequent incremental changes are involved during the process that leads a linguistic element towards a discourse marker? (2) What motivational factors are behind specific constructional changes? (3) Are there cross-linguistic generalizations to be made, both in terms of the semantic and/or syntactic sources and the development paths? It examines a pair of linguistically heterogeneous expressions that typically function in the same pragmatic domain: i.e. topic-introducing discourse markers *speaking of X* (SPOX) in English and *huashuo* in Chinese. The study is corpus-based and includes both functional and frequency distributional analyses both panchronically and diachronically. The results, detailing both commonalities and differences in terms of formal-semantic properties, pragmatic functions and usage patterns, as well as processes of diachronic development, thus enrich cross-linguistic investigations of discourse markers at large and open up new directions for further studies.

About the author:

Yinchun Bai studied English and International Business at the Tianjin Foreign Studies University, China and received her M.A. in European Linguistics at the University of Freiburg, Germany. She is an experienced researcher with a demonstrated history of working in higher education and research institutes. This book is based on her joint Ph.D project conducted at the University of Freiburg, Germany and the University of Antwerp, Belgium. The project was partly funded by the German federal LGFG doctoral scholarship "Empirical Linguistics".

Die Publikationsreihe NIHIN – New Ideas in Human Interaction – entstand 2010 und ist ein Kooperationsprojekt zwischen der Hermann Paul School of Linguistics (HPSL) und der Universitätsbibliothek Freiburg (UB).

NIHIN bietet eine moderne, frei zugängliche Plattform für wissenschaftliche Essays erfahrener WissenschaftlerInnen sowie Prädikatsdissertationen, Textsammlungen zum Thema Sprache in der Interaktion und multimodale Sprachkorpora.

ISBN 978-3-928969-98-7



9 783928 969987 >



Universität
Basel

