Presuppositions at the Semantics-Pragmatics Interface:

Experimental Studies on Their Classification, Acquisition and Cross-Linguistic Comparison

Dissertation
zur Erlangung des philosophischen Doktorgrades
an der Philosophischen Fakultät der Georg-August-Universität Göttingen

vorgelegt von
Yuqiu Chen
aus Nanjing, China

Göttingen 2024

doi: 10.53846/goediss-10391

1. Gutachter: Prof. Dr. Markus Steinbach
2. Gutachterin: Dr. Mailin Antomo
3. Gutachter: Prof. Dr. Hannes Rakoczy

Acknowledgments

In this quite small section, I would like to express my very great appreciation to all people who have contributed to this work and supported me during my PhD study.

First and foremost, I am extremely grateful to my first supervisor, Prof. Dr. Markus Steinbach, for his constructive supervision and valuable advice, which have rescued me from the chaos and struggle with research several times, and for his continuous support and confidence in me and my research during the last 5 years. His competence and enthusiasm have also shown me that the journey in academia can be charming and enjoyable. I would also express my deepest appreciation to my second supervisor, Dr. Mailin Antomo. During my PhD study, she has not only accompanied me step by step from the practical suggestions to my very first exposé over collaboration in several studies to the patient discussion about every single chapter of my dissertation, but is also an inspiring example of female researchers who maintain a balance between life and career. My sincere thanks are extended to Prof. Dr. Hannes Rakoczy for agreeing to be my third supervisor and offering important advice from a psychological perspective at every meeting. Without my supervisors' support, I could not have undertaken this journey. I would also like to sincerely thank Prof. Dr. Marco Coniglio for agreeing to be the Chair of my defense committee.

Secondly, I want to especially and deeply thank the China Scholarship Council (CSC) for their financial support from 09.2018 to 08.2021, and the Graduiertenschule für Geisteswissenschaften Göttingen (GSGG) for their completion grants from 05.2022 to 08.2022. Having received these two scholarships means a lot to me.

Thirdly, I would like to express my special thanks to Maik Thalmann for his competent and valuable co-work in the first experiment. As this experiment has been published, I would also like to deeply thank the two anonymous reviewers and the editors of *Journal of Pragmatics* for their suggestions and evaluation. Many thanks also go to Pia Büsse, Linda Schleef and Imke Wedde for their help in creating the stimuli and testing the participants, and to the kindergarten Kunterbunt in Adelebsen for their support. For the second empirical study, I would like to sincerely thank Prof. Dr. Kong Deming (Nanjing University) for her generous help with the acquisition of participants, and the participants of the conference *XIV. Kongress der Internationalen Vereinigung für Germanistik*, section *Kontrastive Pragmatik* (Palermo, July 2021) for the instructive discussion. For the last experiment, I would like to give special thanks to Dr. Chang Xuan (Nanjing University),

Chen Hairong, Tang Wenli (University of Geneva) and Dr. Xia Yihui (Guangzhou College of Technology and Business) for their constructive discussions and suggestions for the construction of the Chinese items, and many thanks to Dr. Thomas Weskott for his valuable advice on statistics. Furthermore, I would like to thank all 213 participants from 4 to 65 years of age for contributing to the three experiments in this dissertation. I am also grateful to Susanne Müller for her help in testing the participants and for her hard work of proof-reading the whole dissertation.

Fourthly, I would like to thank the team of GSGG, the team of the PhD-program "Behavior and Cognition" and the Research Training Group "Form-meaning mismatches" (RTG 2636) for offering so many valuable workshops, enlightening colloquiums and all further interesting events, and also our Examination Office for the organizational support.

Last but not least, I would like to express my heartfelt appreciation to my dear mother, Chen Shuhua, who has not only helped me to maintain my life standard with her financial support, but has also looked after my emotional and spiritual health through ups and downs with her strong heart and open mind: 谢谢秋妈！I would also like to thank Evelyne and Matthias Gnatzy, Li Yu, Song Yue and all further family and friends for their sweet emotional support during the last 5 years.

Lastly, I would like to thank myself for keeping both mentally and physically stable during the whole study.

Abstract

While in traditional semantic accounts, presuppositions are widely accepted to be conceptualized as truth value preconditions, in most pragmatic accounts, they are considered to be a set of information that is normally or at least can be assumed as part of the common ground (see a.o. Frege 1892; Strawson 1950; Stalnaker 1972; Heim 1991; Stalnaker 1998 and summaries in e.g. Levinson 1983c; Beaver et al. 2021). Reviewing investigations of the last decades, one important observation is that although presuppositions are sorted into one category of projective meaning, they do not always behave uniformly but can have heterogeneous projection strength in certain contexts (see a.o. Karttunen 1971b; Simons 2001; Abusch 2002; Charlow 2009; Karttunen 2016). Addressing the heterogeneity of presuppositions, various theories have been proposed in order to discuss the necessity of their classification and criteria for it (see a.o. Zeevat 1992; Abusch 2010; Tonhauser et al. 2013; Abrusán 2016; Klinedinst 2016; Tonhauser et al. 2018). The classification issue remains highly controversial and is, therefore, at the heart of this dissertation. Of the variety of theories regarding this issue, I will explore mainly two: the first one is a distinction of soft and hard triggers, in a wider sense than firstly proposed by Abusch (2002, 2010), whereas the second one is rooted in the concept of at-issueness (see e.g. Simons et al. 2010; Tonhauser et al. 2018). The second main research concern of this dissertation is the acquisition of presuppositions in both first and second languages, i.e. the acquisition among children and among nonnative adult speakers. As there are only few cross-linguistic comparisons of presuppositions, and the comparisons are mainly made between European languages (see e.g. Amaral and Cummins 2015; Schwarz et al. 2020; Reins et al. 2021), the last main concern of my research addresses the universality of the properties and heterogeneity of presuppositions. These three main research concerns are further divided into five main research questions. Their aims are, inter alia, to verify the classification by testing the trigger split with regard to at-issueness, to apply the classification by measuring and sorting other triggers with typical soft and hard triggers functioning as anchors, to investigate the acquisition of the trigger distinction and at-issueness in first and second languages, and to assess the potential cross-linguistic stability of projection behaviors between German and Chinese.

These research questions were investigated by means of three experiments, in which the comparisons took place between German adults and preschool children, native and nonnative speakers of German, and native speakers of Chinese and German, respectively. In the first two experiments, video stimuli were used and participants were asked to judge the acceptability of soft and hard triggers whose presuppositions are forced to be at-issue

or remain not-at-issue. In the third experiment, a paper-and-pencil questionnaire was used and the projection strength of presuppositions in Chinese and German was measured by the certainty judgment task (Tonhauser et al. 2018).

The most relevant findings can be summarized as follows: firstly, presuppositions of hard triggers are significantly less suitable for being at-issue than those of soft triggers. This observation might indicate that the explanations for the soft-hard split and at-issueness, although based on very different assumptions, can be combined, namely via the concept of local accommodation. Secondly, the sub-class of soft triggers is more heterogeneous than most theories have predicted so far, and their softness might result from different mechanisms. This observation may illustrate that the soft-hard split may rather describe a contrast between triggers with less typical and typical presuppositions. Thirdly, although less sensitive than adults, preschool children are aware of at-issueness violations and the trigger split. This observation also confirms the importance of the developmental stage between 4 and 6 years of age. Fourthly, nonnative speakers are also aware of at-issueness violations and the trigger split but less sensitive than native speakers. Interestingly, their deviation in sensitivity from the native speakers was only significant for the hard triggers but not for the soft ones, challenging the uniform approach of soft triggers and scalar implicatures (Chemla 2008; Romoli 2014). Last but not least, the empirical results suggest that the projection behavior of presuppositions is stable between Chinese and German, providing a partial answer to the universality issue and a cross-linguistic baseline for intercultural pragmatic studies in the future. Methodologically, the empirical research in this dissertation highlights the relevance of contrastive pragmatics and provides some interesting information for further research, such as for the item drafting in experiments with children as participants.

Keywords: presuppositions, soft-hard distinction, at-issueness, language acquisition, L2 pragmatics, cross-linguistic comparison, experimental pragmatics, contrastive pragmatics

Zusammenfassung

Aus semantischer Sicht werden Präsuppositionen im Allgemeinen als Wahrheitswertvoraussetzungen verstanden, aus pragmatischer Sicht hingegen als Informationen, die normalerweise ein Teil des Common Ground sind oder zumindest als gegeben angesehen werden können (siehe z.B. Frege 1892; Strawson 1950; Stalnaker 1972; Heim 1991; Stalnaker 1998 sowie die Zusammenfassungen in Levinson 1983c; Beaver et al. 2021). Mit Blick auf die Untersuchungen der letzten Jahrzehnte beobachtet man vor allem, dass sich Präsuppositionen nicht immer einheitlich verhalten, obwohl sie als eine Kategorie projektiver Bedeutungen zusammengefasst werden, sondern in bestimmten Kontexten eine heterogene Projektionsstärke aufweisen können (siehe u.a. Karttunen 1971b; Simons 2001; Abusch 2002; Charlow 2009; Karttunen 2016). Angesichts der Heterogenität von Präsuppositionen wurden zahlreiche Theorien aufgestellt und die Notwendigkeit sowie die Kriterien für eine Klassifikation diskutiert (siehe u.a. Zeevat 1992; Abusch 2010; Tonhauser et al. 2013; Abrusán 2016; Klinedinst 2016; Tonhauser et al. 2018). Das Thema der Klassifikation bleibt hoch umstritten und steht daher im Mittelpunkt dieser Dissertation. Von den vielen Theorien zu diesem Thema werden vor allem zwei untersucht: die erste ist eine Unterscheidung zwischen weichen und harten Auslösern in einem weiteren Sinne als zuerst vorgeschlagen von Abusch (2002, 2010), während die zweite auf dem Konzept der At-issueness basiert (siehe z.B. Simons et al. 2010; Tonhauser et al. 2018). Der zweite Forschungsschwerpunkt dieser Dissertation ist der Erwerb von Präsuppositionen, und zwar sowohl in der Muttersprache als auch in der Fremdsprache, d.h. der Erwerb bei Kindern und bei erwachsenen Nicht-Muttersprachler/innen. Da es bisher nur wenige sprachübergreifende Vergleiche von Präsuppositionen gibt und die Vergleiche hauptsächlich innerhalb der europäischen Sprachen stattfinden (siehe z.B. Amaral and Cummins 2015; Schwarz et al. 2020; Reins et al. 2021), befasst sich der letzte Forschungsschwerpunkt mit der Universalität der Eigenschaften und Heterogenität von Präsuppositionen. Diese drei Forschungsschwerpunkte werden in fünf Forschungsfragen unterteilt. Sie zielen u.a. darauf ab, die Klassifikation zu überprüfen, indem die Unterscheidung zwischen den Auslösern im Hinblick auf At-issueness getestet wird; die Klassifikation anzuwenden, indem andere Auslöser mit typischen weichen und harten Auslösern als Ankerpunkten gemessen und sortiert werden; den Erwerb der At-issueness und des Unterscheids zwischen den Auslösern zu untersuchen, und zwar in der Muttersprache und in der Fremdsprache; und die sprachübergreifende Stabilität des Projektionsverhaltens zwischen Deutsch und Chinesisch zu untersuchen.

Diese Forschungsfragen wurden mit drei Experimenten untersucht, wobei die Vergleiche zwischen deutschen Erwachsenen und Vorschulkindern, Muttersprachler/innen und Nicht-Muttersprachler/innen des Deutschen, und Muttersprachler/innen des Chinesischen und Deutschen stattfanden. In den ersten beiden Experimenten wurden Videostimuli verwendet und die Teilnehmenden wurden gebeten, die Akzeptabilität von weichen und harten Auslösern zu beurteilen, deren Präsuppositionen at-issue oder nicht-at-issue sind. Im dritten Experiment wurde die Projektionsstärke von Präsuppositionen auf Chinesisch und Deutsch mithilfe der Sicherheit-Beurteilung (Tonhauser et al. 2018) und eines Paper-and-Pencil-Fragebogens gemessen.

Die wichtigsten Ergebnisse lassen sich wie folgt zusammenfassen: Erstens sind die Präsuppositionen von harten Triggern signifikant weniger geeignet als die von weichen Triggern, um at-issue zu sein. Dieser Zusammenhang könnte darauf hindeuten, dass die Theorien zur Weich-Hart-Unterscheidung und zur At-issueness kombiniert werden können, obwohl sie auf unterschiedlichen Annahmen beruhen. Zweitens ist die Unterklasse der weichen Auslöser heterogener als die meisten Theorien bisher angenommen haben, da ihre Weichheit auf unterschiedliche Mechanismen zurückzuführen sein kann. Das heißt, die Weich-Hart-Unterscheidung könnte vielleicht eher einen Kontrast zwischen Triggern mit weniger typischen und typischen Präsuppositionen darstellen. Drittens können Vorschulkinder At-issueness-Verletzungen und den Unterschied zwischen den Auslösern schon wahrnehmen, wenngleich sie dabei weniger sensibel als Erwachsene sind. Diese Beobachtung bestätigt die Relevanz der Entwicklungsstufe zwischen 4 und 6 Jahren. Viertens können Nicht-Muttersprachler/innen At-issueness-Verletzungen und den Unterschied zwischen den Auslösern ebenfalls wahrnehmen, sind dabei aber weniger sensibel als Muttersprachler/innen. Interessanterweise war ihre Sensitivitätsabweichung von den Muttersprachler/innen nur bei den harten Auslösern signifikant, nicht bei den weichen, was die Idee der Einheitlichkeit von weichen Auslösern und skalaren Implikaturen (z.B. Chemla 2008; Romoli 2014) in Frage stellt. Fünftens deuten die empirischen Ergebnisse darauf hin, dass das Projektionsverhalten von Präsuppositionen zwischen Chinesisch und Deutsch stabil ist, und sie bieten eine Teilantwort auf die Frage nach der Universalität sowie eine sprachübergreifende Baseline für zukünftige interkulturelle pragmatische Studien. In methodischer Hinsicht verdeutlichen die empirischen Studien in dieser Dissertation die Relevanz der kontrastiven Pragmatik und liefern einige interessante Hinweise für die weitere Forschung, z.B. für den Itemaufbau in Experimenten mit Kindern.

Schlüsselwörter: Präsuppositionen, Weich-Hart-Unterscheidung, At-issueness, Spracherwerb, L2-Pragmatik, sprachübergreifender Vergleich, experimentelle Pragmatik, kontrastive Pragmatik

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CLMM ..... Cumulative Link Mixed Model

CP ..... Cooperative Principle

DaF ..... German as a foreign language

GCI ..... generalized conversational implicature

ISI ..... indirect scalar implicature

L1 ..... first language

L2 ..... second language

LMM ..... linear mixed model

NRRC ..... non-restrictive relative clause

PCI ..... particularized conversational implicature

PSP ..... presupposition

QUD ..... Question Under Discussion

SI ..... scalar implicature

TVJT ..... Truth-Value Judgment Task

1 Introduction

In 2016, there finally was a movie in which linguists played an essential role by saving the world (at least to some extent): *Arrival*, a science fiction film adapted from Ted Chiang's (1998) short story "Story of Your Life", is about communication between humans and aliens. The essential conception behind this story is assumed to be mainly rooted in the Sapir-Whorf Hypothesis (Sapir 1921; Whorf 1940), a theory that has been highly controversial since it was proposed. Briefly speaking, this hypothesis assumes that language can influence speakers' thought, perception and cognition (for more details and discussions, see e.g. Kay and Kempton 1984; Hussein 2012; Cibelli et al. 2016). In the movie, it is presented as Louise Banks, the linguist who learns the aliens' language (Heptapod), gains a different perception of time and can see into the future. Besides the disputations on the Sapir-Whorf Hypothesis and the thrilling plot, in the movie, there is an impressive and funny conversation that seems very interesting for my research topic. The conversation takes place between Louise and her physicist colleague Ian Donnelly, who is just slowly falling in love with her, after Louise has had another flashback and realized that this was not her memory but her future:

(1) Louise (mumbling): I just realize why my husband left me.[...]
Ian: You were married?!

The short sentence expressed by Louise is in fact very informative. Considering it word by word, firstly, the expression *my husband* indicates that Louise had a husband, in other word, she was married. Second, the word *left* expresses the change of state, in other words, the husband can leave Louise only if he was with her before. Third, if a person can realize something, saying p, then normally, p has to be the case. In this sentence, it means that Louise's husband did leave her (for some reasons). And fourth, Louise realizes the reason why he left her. The whole set of information contained in this one short sentence in (1) can be summarized as follows:

(2) a. Louise had a husband, i.e. she was married.
b. The husband was with her before.
c. The husband left her (for some reasons).
d. Louise realizes the reason of his leaving.

Among all these propositions, note that only (2d) is actually asserted, while the other three are in fact required by certain linguistic expressions in the sentence, such as the

possessive pronoun *my*, the change of state verb *leave* and the factive verb *realize*. As audience, we can understand her sentence properly because we see the 'flashback' into the future together with Louise and thus, from our view, (2a)-(2c) are all fulfilled. That is, this information can be regarded as true or already given in the context, and then be taken as shared knowledge between Louise and the audience. Thus, for the audience, uttering such a sentence is considered felicitous and understandable, and the sentence itself can be interpreted as true.

However, from Ian's perspective, the case is more complex. As only Louise is able to see the future but her colleague cannot, (2a)-(2c) are not given in the context or mutually known by the speaker and hearer in the conversation. Moreover, based on Ian's knowledge about Louise, she is neither married nor can be assumed to be married – from his perspective, she never had a husband, not to mention other events that have to do with this not-existing person. Now, under this condition, how can Louise's utterance be interpreted? Should it be regarded as true or false? Is it appropriately uttered in the first place? As we have seen in the film, due to the information clash, Louise's sentence is regarded by Ian neither as true nor as false, but rather as surprising, inappropriate or even a little weird, so that he is very confused and has to question (2a) explicitly – for more information about Louise's marriage status, and perhaps also to signal that something goes wrong in this conversation.

This required information that needs to be true, given in context, or at least can be supposed to be true or given before the expressions themselves are uttered, are called presuppositions. Concerning these presupposed contents, a key observation is that in contrast to asserted contents like (2d), presuppositions like (2a)-(2c) can hardly be affected by operators like negation, question, conditionalization and modalization, the so-called family of sentences (Chierchia and McConnell-Ginet 2000). For a better illustration, slightly modifying Louise's sentence in (1) into third person, as shown in (3), the difference between presuppositions and assertions can be illustrated as in (4):

(3) Louise realizes why her husband left her.
(3) $\rightsquigarrow$ (2a)-(2d)

(4) a. Negation: Louise does not realize why her husband left her.
b. Question: Does Louise realize why her husband left her?
c. Conditionalization: If Louise realizes why her husband left her, she may also understand the perception of non-linear time.
d. Modalization: Perhaps Louise realizes why her husband left her.

(4a)-(4d) $\rightsquigarrow$ (2a)-(2c)
(4a)-(4d) $\not\rightsquigarrow$ (2d)

As shown above, once the utterance is embedded under an operator from the family of sentences, the asserted proposition is modified, while the presupposed content is not affected – it survives or projects out of the scope of these otherwise entailment-canceling operators. The projection behavior is one of the relevant properties which distinguish presuppositions from other meanings like assertions, implicatures and entailments.[1]

As a phenomenon at the semantics-pragmatics interface, presuppositions have been investigated for a long time. Starting from the philosophical area (see e.g. research like Frege 1892; Russell 1905; Strawson 1950), investigations on presuppositions began with research on the referent demands of definite expressions like *the king of France* in (5), and have received linguistic attention mainly since the second half of the 20th century (see e.g. Horn 1969; D. Langendoen and Savin 1971; Stalnaker 1972; Katz 1973).

(5) ϕ: *The king of France* lived in Versailles.

$\rightsquigarrow \psi$: There was one and exactly one king of France.

From the traditional semantic perspective, the presupposition ψ is defined either as a truth condition for the sentence, similar with propositions in assertions (mainly traced back to Russell 1905), or a condition for the sentence being able to be interpreted as true or false, i.e. a truth value precondition, whose failure makes the sentence neither true nor false (mainly traced back to Frege 1892; Strawson 1950). In comparison, the later one has been more widely accepted in later research (see e.g. definitions given by Van Fraassen 1966; Katz 1973; Heim 1991, among others). However, at least since Stalnaker (1972), the pragmatic aspects have been considered increasingly relevant to the investigation of presuppositions. Instead of truth value, the pragmatic definitions of presuppositions primarily focus on their constraints on the felicity in conversation and on the common ground. From these points of view, presuppositions are considered as information that is normally mutually shared or can be assumed to be part of the common ground when a sentence with certain expressions is uttered (see e.g. definitions in Stalnaker 1973, 1977 and summary in e.g. Levinson 1983c; Beaver et al. 2021).

Besides possessive pronouns, change of state verbs, factive verbs and definite expressions that have been mentioned above, in our daily communication, there are a great number of linguistic forms that also have such pre-requirements on the context and can therefore trigger presuppositions. These linguistic forms are called presupposition triggers. Among various lists of presupposition triggers, the most frequently cited and widely discussed one is given by Levinson (1983c) and contains 13 sub-classes, although several of

[1]Note, however, that projection is neither a sufficient nor a necessary property of presuppositions. On the one hand, there are also non-presupposed projective contents, and on the other hand, some presuppositions can also lose their projectivity in certain contexts. I will come to this point later in Chapter 2 and 3.

them have caused controversies and were then excluded later. This list is adopted below with examples and projection tests by means of negation.[2]

Table 1: Summary of Levinson's list of presupposition triggers (Levinson 1983c, pp. 181–184), with selected examples and in modified form.

Sub-classes	Triggers	Example with presuppositions
1. Definite descriptions (Strawson 1950, 1952)	the	John saw/didn't see *the man with two heads*. ⇝ There exists a man with two heads.
2. Factive verbs (P. Kiparsky and C. Kiparsky 1970)	regret, aware, realize, know, be sorry that, be proud that etc.	Martha *regrets*/doesn't *regret* drinking John's home brew. ⇝ Martha drank John's home brew.
3. Implicative verbs (Karttunen 1971a)	manage, forget	John *managed*/didn't *manage* to open the door. ⇝ John tried to open the door.
4. Change of state verbs (Sellars 1954; Karttunen 1973)	stop, begin, continue, finish, carry on, leave, go, arrive etc.	Kissinger *continued*/didn't *continue* to rule the world. ⇝ Kissinger had been ruling the world.
5. Iteratives	again, anymore, return, repeat etc.	The flying saucer came/didn't come *again* ⇝ The flying saucer came before.
6. Verbs of judging (Fillmore 1971)	accuse, criticize	Agatha *accused*/didn't *accuse* Ian of plagiarism. ⇝ (Agatha thinks) plagiarism is bad.
7. Temporal clauses (Frege 1892)	before, while, since, after during etc.	*Before* Strawson was even born, Frege noticed/didn't notice presuppositions. ⇝ Strawson was born.
8. Cleft sentences (Halvorsen 1978; Prince 1978; Atlas and Levinson 1981)	it-clefts, pseudo-clefts	It was/wasn't Henry that kissed Rosie. ⇝ Someone kissed Rosie.

[2]Interestingly, according to Levinson (1983c), his list is based on Karttunen's handout with 31 different kinds of presupposition triggers. However, according to Karttunen (2016), neither the handout nor the list with 31 types of triggers can be found anymore.

Table 1 – *Continued from previous page*

Sub-classes	Triggers	Example with presuppositions
9. Implicit clefts with stressed constituents (Chomsky 1972; Wilson and Sperber 1979)	heavy stress on a constituent	Linguistics was/wasn't invited by CHOMSKY! ⇝ Someone invited linguistics.
10. Comparisons and contrasts (Lakoff 1971)	–	Carol is/isn't a better linguist than Barbara. ⇝ Barbara is a linguist.
11. Non-restrictive relative clauses	–	The Proto-Harrapans, who flourished 2800-2650 B.C., were/were not great temple builders ⇝ The Proto-Harrapans flourished 2800-2650 B.C.
12. Counterfactual conditionals	–	If Hannibal had only had twelve more elephants, the Romance languages would/would not this day exist. ⇝ Hannibal didn't have twelve more elephants.
13. Questions (Katz 1972; Lyons 1977)	yes/no questions, alternative questions, WH-questions	Who is the professor of linguistics at MIT?[3] ⇝ Someone is the professor of linguistics at MIT.

Note, however, that this long list given by Levinson (1983c) is disputed and revisions are required: on the one hand, there are triggers like non-restrictive relative clauses (NRRCs) whose contents are nowadays generally assumed to be only backgrounded but not presupposed (more about NRRCs later in Section 3.2 and 7.2.2). On the other hand, it also lacks triggers that have been recently proposed to have presupposed effects, such as quantifiers, names or even non-verbal triggers like gestures (see e.g. Beaver et al. 2021; Schlenker 2021b).

Nevertheless, the list with 13 (potentially even 31) sub-classes can still illustrate the diversity and controversy of presupposition triggers. Interestingly, the diversity of triggers – or the various "paradigm cases of a presumed relation between sentences" (Stalnaker 1973, p. 447) – foreshadows one of the most crucial concerns or key issues in presupposition

[3]Noticing that the presuppositions of questions, if any, are not constant under negation, Levinson (1983c) did not provide the negation test for this sub-class.

research: the heterogeneity problem, also described as *the difference problem* by Romoli (2014). During the investigation in the last decades, this problem has been observed primarily in that presuppositions of various triggers have different projection behaviors in certain contexts, at least since Karttunen's (1971) study on factive verbs:

(6) a. If I regret later that I have not told the truth, I will confess it to everyone.
$\rightsquigarrow$ The speaker had not told the truth.

b. If I discover later that I have not told the truth, I will confess it to everyone.
$\rightsquigarrow / \not\rightsquigarrow$ The speaker had not told the truth.

(Karttunen 1971b, p. 64)

As can be seen in (6), although both triggers are factive verbs and all other parts of the sentence remain the same, when embedded under the operator *if*, the projection behavior of their presuppositions varies. While (6a), or the speaker uttering (6a), still clearly presupposes that the speaker had not told the truth, in (6b), the complement can but does not have to be projective. The heterogeneity of presuppositions has been further observed between different sub-classes of triggers and in several other contexts by a great number of both theoretical and empirical investigations (see e.g. Simons 2001; Abusch 2002; Charlow 2009; Smith and Hall 2011; Romoli 2014, among others).

As this heterogeneity may also challenge the traditional semantic notion of presuppositions, theories have addressed this problem more or less by including pragmatic aspects. Based on the review in Beaver et al. (2021), roughly speaking, these theories can be split into two branches. The first branch assumes that presuppositions are triggered similarly, namely rather conventionally, and then can be canceled or lose their projectivity due to certain mechanisms, such as their felicity with regard to other implied meanings or the assertion (e.g. Gazdar 1979a,b; Van der Sandt 1988), or their at-issueness (e.g. Simons et al. 2010; Tonhauser et al. 2018), that is, their relevance regarding the Question Under Discussion (QUD, see e.g. von Stutterheim and Klein 1989; Roberts 1996). Theories from the other branch address this problem by claiming that the generating process of presuppositions varies between triggers, and therefore suggest several classification possibilities of presupposition triggers (e.g. Zeevat 1992; Abusch 2002; Kripke 2009; Abusch 2010; Tonhauser et al. 2013; Romoli 2014; Klinedinst 2016 and more).

The heterogeneity and classification issue is the first main research concern of my dissertation. Among all theories and predictions, I focus mainly on two accounts: the first with at-issueness at its center (e.g. Simons et al. 2010; Tonhauser et al. 2018), as it is one of the most recent developments of the first branch, and the other one with a soft-hard split of triggers for the second branch, firstly proposed by Abusch (2002, 2010), but in a wider sense. In particular, I will investigate whether the soft-hard split of triggers can be reflected in terms of their ability to express at-issue content, and whether the typical

soft and hard triggers can be used as anchors so that other triggers can be measured and sorted.

The second main research concern of my study focuses on the acquisition of presuppositions, including both first and second languages (L1 and L2). Addressing L1 acquisition, the question is mainly how presuppositions are acquired among children and whether or from what age on speakers are sensitive to the properties and heterogeneity of presuppositions. As for L2 acquisition, it is normally assumed that nonnative adult speakers have acquired all necessary pragmatic abilities in their mother tongue. Can they then adopt their pragmatic abilities or sensitivity automatically into their foreign language? Do they recognize presuppositions or pragmatic infelicities involving presuppositions as sensitive as native speakers?

In addition to its importance of the L1 development process and L2 teaching/learning strategies, the acquisition issue is also closely related to the first main research concern: on the one hand, L1 acquisition is a crucial aspect for trigger classification, as "if each presupposition trigger belongs either to the hard or to the soft category, how do children eventually figure out which box a specific expression should go into?" (Zehr and Schwarz 2018, p. 479). On the other hand, previous studies have also shown the importance and usefulness of L2-pragmatics and native-nonnative comparisons for pragmatic research (see e.g. Carrell 1984; Taguchi 2013; Y. Chen 2019). Thus, it is expected that aspects from both L1 and L2 acquisition can shed light on the heterogeneity and classification issue of presuppositions.

The third main research concern is about the universality of presuppositions, that is, whether presuppositions and their properties and heterogeneity are stable across languages. Take *regret* from the trigger list in Table 1 for example, the question can be then understood as follows: do *regret*-expressions in different languages, for instance as in the three sentences below in (7), have identical presuppositional effects? Moreover, if presupposition triggers can be classified, for example according to the soft-hard distinction, are *regret*-expressions similarly hard or soft in all languages?

(7) a. English: Martha *regrets*/doesn't *regret* drinking John's home brew. (Levinson 1983c, p. 181)

b. German: Martha *bereut*/*bereut* nicht, dass sie Johns Hausgebräu getrunken hat.

c. Chinese: 玛塔**后悔**/不**后悔**她喝了约翰的家酿啤酒。

Interestingly, the trigger list given in Levinson (1983c) has been translated, as part of his textbook *Pragmatics*, into several languages for a long time, while the cross-linguistic stability of presuppositions has received empirical interest only very recently and by very few studies (see e.g. Amaral and Cummins 2015; Schwarz et al. 2020; Reins et al. 2021).

Moreover, while the traditional pragmatic theories or principles are implicitly or explicitly assumed to be general or rooted in human cognition (see e.g. Grice 1989; Sperber and Wilson 1996 for their descriptions on the Cooperative Principle and the Cognitive Principle of Relevance, respectively), empirical studies show that in numerous pragmatic fields, language- and culture-specific properties are influential and crucial (see summaries in e.g. R. Chen 2010; Ameka and Terkourafi 2019; Jia and Yang 2021, among others). This contrast naturally leads to the question of to what extend presuppositions and their classification can be considered cross-linguistically stable.

In sum, my research focuses on three main research concerns and targets five main research questions as listed below in (8):

(8) a. The classification issue:
Research question 1: verifying the classification: can the soft-hard dichotomy, a classification based on the projective strength, also be reflected in terms of the ability of different triggers to express at-issue content?
Research question 2: applying the classification: choosing certain typical presuppositions or triggers as markers for certain properties, can they function like a thermometer that measures and sorts other triggers into certain subclasses?

b. The acquisition issue:
Research question 3: acquisition in first languages: are children aware of the difference between trigger types? Do children and adults respond differently once the presuppositions are forced to be at-issue? Are the different trigger types acquired at different ages?
Research question 4: acquisition in second languages: are nonnative speakers as sensitive as native speakers once the presuppositions of different triggers (or trigger types) are forced to be at-issue? If not, how can the difference be explained?

c. The universality issue:
Research question 5: cross-linguistic comparison of presuppositions: can a presupposition be triggered steadily across languages, as long as an adequate and equivalent translation of its trigger can be found? Are their properties like projection behavior or information status independent from linguistic and cultural influence?

Addressing the three main research concerns and five main research questions, my dissertation is structured as follows: first, the whole Part I aims primarily at introducing the theoretical and empirical research background and achieving an overview on current studies that are relevant to the three main issues. As the definitions of presuppositions

are various and assumptions regarding them are diverse, relevant theories from both semantic and pragmatic accounts are reviewed in Chapter 2, aiming at a very basic but not superficial question, namely: how is the presupposition defined and how are its properties like projection explained so far? After that, relevant research background to the three main issues and the key concepts of each main research question are reviewed in detail in Chapter 3, 4 and 5 respectively, while a summary can be found in Chapter 6.

In Part II, three empirical studies are presented. The first experiment (Exp.I), reported in Chapter 7, focuses on Research question 1 and 3, investigating the trigger classification from the perspective of L1-acquisition. The second experiment (Exp.II), reported in Chapter 8, adopts the design from the first experiment and explores the nonnative speakers' sensitivity to presuppositions and their heterogeneity, addressing Research question 1 and 4. The third experiment (Exp.III), reported in Chapter 9, addresses Research question 2 and 5 and examines the cross-linguistic stability of the projection behavior and classification of presuppositions. A summary of the whole experimental part can be found in Chapter 10. Since the first two experiments have already been published as journal articles, see Y. Chen et al. (2022) and Y. Chen (2022), and the third one has been accepted for publication in the Journal *Intercultural Pragmatics*, Chapter 7-9 can also be read separately. Moreover, due to the cumulative nature of this part, certain repetition or redundancy in this work (not restricted to Part II) can hardly be avoided.

In Part III, the general discussion on all three main research concerns and some methodological aspects can be found in Chapter 11, followed by a conclusion in Chapter 12, which contains a mind map of the whole dissertation in Figure 19.

Part I

Theoretical and Empirical Background

2 An Overview on Relevant Presupposition Theories

As mentioned before, presuppositions have received linguistic interest mainly since the second half of the 20th century and around the 1970s, they became a central topic (see e.g. Horn 1969; D. Langendoen and Savin 1971; Stalnaker 1972; Katz 1973; Gazdar 1978 and more reviewed below). As it may raise "substantial problems for almost all kind of (generative) linguistic theories" (Levinson 1983c, p. 167), the presupposition is one of the most well-researched phenomena at the semantics-pragmatics interface (see discussions in e.g. Stalnaker 1973; Levinson 1983c; Stalnaker 2002). However, presuppositions still remain controversial and research-demanding, not only concerning them did not initially begin with an abstract and uniform concept "but with some paradigm cases of a presumed relation between sentences" (Stalnaker 1973, p. 447), but also due to their heterogeneity and variety observed in current research, especially in experimental studies. Thus, this chapter is designed to answer a very basic but not superficial question, namely, how is the presupposition defined and how are its properties like projection explained so far? The chapter is structured as follows: first, some influential assumptions from traditional semantic accounts are reviewed in Section 2.1, together with a summary of their limitations and weaknesses. Then, in Section 2.2, the pragmatic aspects are introduced, with a focus on two branches, in which presuppositions are considered as conventional or conversational meanings.

2.1 The Traditional Semantic Accounts

According to Levinson (1983c), there are two relevant semantic accounts of presuppositions: the first one concentrates on the truth value and logic models (see e.g. Frege 1892; Russell 1905; Strawson 1950; Van Fraassen 1966; Heim 1991, among others), and the other one attempts an atomic-concept or feature-style oriented semantics (e.g. Katz and T. Langendoen 1976). As the latter one was less completely defined but more questioned and criticized (see discussions in Gazdar 1978; Levinson 1983c) and hardly discussed recently, in this section, the review on presuppositions only contains theories from the first semantic account. Their main assumptions, definitions and formal representations of presuppositions are introduced summarily in Subsection 2.1.1, while their potential weaknesses and limitations are outlined briefly in the Subsection 2.1.2.

2.1.1 Presuppositions as Truth Value Preconditions

As mentioned before, the investigation of presuppositions can be traced back to the discussion on definite expressions like *the king of France* in (5), repeated below in (9), focusing especially on their referring and definiteness: normally, the definite expression *the king of France* pre-requires the existence and uniqueness of the king.

(9) ϕ: The king of France lived in Versailles.

$\rightsquigarrow \psi$: There was one and exactly one king of France.

According to the summary given in Heim (1991), this question was firstly addressed by the theory of referential interpretation (*referentielle Deutung* in German). This theory assumes that the expression *[the* ζ*]*ξ always refers to a certain individual. In (9) for instance, the individual can be a French king like Louis XIV, and the truth value of the sentence depends then on whether he lived in Versailles or not. If there is no such a king who can be referred to in the situation, or more than one king, referential interpretation predicts that the utterance cannot express any proposition – thus no need to mention its truth value at all.

Although the prediction given by referential interpretation appears intuitive, according to Heim (1991), it has difficulties to explain certain sentence types, especially in the context of direct negation. If we assume there was no king of France, then according to referential interpretation, neither the expression (9) nor its negation (10) can express a proposition. However, the sentence (10) has two readings: while the first reading (10a) produces no proposition as expected, the second one (10b), although it might not be as natural as the first one, can be interpreted as $\neg$[[*the* ζ*]*ξ], where the existence of the exactly one king is also negated. According to this reading, the sentence (10) produces a proposition that is actually true in the given situation, which contradicts the prediction (for more discussion on this and further weaknesses, see Heim 1991).

(10) The king of France did not live in Versailles.

a. There was one and exactly one king of France who had the property *did not live in Versailles*.

b. It was not the case that the king of France lived in Versailles.

Recognizing the limitations of referential interpretation, several explanations were proposed and the truth value of assertions with failed presuppositions was discussed. Among them, the most relevant two branches are the assumptions given by Frege (1892), adopted and developed later by Strawson (1950), and by Russell (1905).

Frege (1892) differentiates between the meaning or sense ('*Sinn*') of an expression and its referent ('*Bedeutung*') by assuming that "[a] proper name (word, sign, sign combination,

expression) expresses its sense, refers to or designates its referent. By means of a sign we express its sense and designate its referent" (Frege 1948, p. 214) [4]. On the sentence level, Frege (1892) proposes that the meaning is the function of a sentence while the referent is the function of the use of a sentence – normally, the meaning of the whole sentence is a complex thought and its referent is a truth value (for more discussion, see e.g. Strawson 1950). Addressing the presupposition triggered by the definite expression [*the* ζ], Frege uses the example *the negative square root of 4* and declares explicitly that:

(11) [w]e have here the case of a compound proper name constructed from the predicate expression with the help of the singular definite article. This is at any rate permissible if the predicate applies to one and only one single object.
(Frege 1948, p. 223)[5]

However, Frege's assumption is rejected by Russell (1905). Russell abandons the concept of meaning or sense ('*Sinn*') and claims that the referent (or *donating* in Russell's words) is the critical part of sentence interpretation. With reference to presuppositions, especially the potential ambiguity in (10), Russell distinguishes between the primary and secondary occurrence of a donating phrase *[the* ζ*]*ξ. In order to clarify the difference between the two occurrences, consider the example (9) again. According to Russell, in the case of primary occurrence, ψ is considered a given property of the entity ζ, namely *being the king of France*, and the truth value depends on whether ξ: whether he lived in Versailles. In the case of the secondary occurrence, the truth value depends on whether ψ and ξ: whether there was exactly one king of France, and whether he lived in Versailles. Based on this distinction, Russell explains the truth value problem and the ambiguity of sentences like (10) as follows: in the primary occurrence of *the king of France*, in accordance with (10a), the sentence should be judged to be false: as there is no person who has the property *being the exactly one king of France*, the first condition ψ is not fulfilled, and the whole sentence cannot be true. In the secondary occurrence, the sentence is true, in accordance with (10b), as the existence of exactly one king is also negated.

In short, according to Russell (1905), a sentence ϕ in form *[the* ζ*]*ξ can always produce a complete proposition, and the truth value is bivalent: ϕ (*[the* ζ*]*ξ) is only true if there is exactly one ζ and it has the property ξ; in all other cases, ϕ is considered false. A negated sentence like (10) is predicted to be understood in the way that either the king did not

[4]In Frege (1892, p. 32): "Ein Eigenname (Wort, Zeichen, Zeichenverbindung, Ausdruck) drückt aus seinen Sinn, bedeutet oder bezeichnet seine Bedeutung. Wir drücken mit einem Zeichen dessen Sinn aus und bezeichnen mit ihm dessen Bedeutung."

[5]Frege (1892, p. 42): "Wir haben hier den Fall, daß aus einem Begriffsausdrucke ein zusammengesetzter Eigenname mit Hilfe des bestimmten Artikels im Singular gebildet wird, was jedenfalls dann erlaubt ist, wenn ein Gegenstand und nur ein einziger unter den Begriff fällt."

live in Versailles or there is no such unique king. Accordingly, ψ is not pre-required, but only a logical consequence.

In contrast to Russell (1905), Strawson (1950) adopts the assumption by Frege (1892) and criticizes Russell's confusion of meaning and referring, as an utterance can only be judged as true or false if it is used to address something. If not, that is, if there is no such suitable entity to which the definite expression can refer, the sentence is neither true nor false but has no truth value at all, as the question on its truth cannot arise. With this trivalent logic system, Strawson (1952) defines presuppositions with a focus on the truth value relation between the presupposition and the assertion:

(12) a statement S presupposes a statement S' in the sense that the truth of S' is a precondition of the truth-or-falsity of S [...]. [6] (Strawson 1952, p. 175)

(13) A statement A presupposes another statement B iff:
(a) if A is true, then B is true
(b) if A is false, then B is true
(Strawson's view summarized in Levinson 1983c, p. 175)

Following the Frege-Strawson notion of presupposition, Heim (1991), for instance, states that although interpreting definite expressions under negation in the sense of Russell's (1905) explanation is possible, such interpretation is not the most neutral one but only accessible with efforts. With a simple negation like in sentence (10), the most obvious understanding is only the partial negation, namely *[the* ζ*]*$\neg\xi$ as shown in (10a). Only if there is contradiction in the context, the secondary understanding is allowed to emerge *ad hoc*. Furthermore, Heim (1991) argues that the presupposition ψ, namely that *the* ζ expresses a set with a single-element, is neither a precondition for expressing a proposition as predicted by referential interpretation, nor a logical consequence of ϕ as claimed by Russell (1905), but a precondition for ϕ being true or false: if ψ is false, ϕ's proposition has no truth value.

In comparison to Russell's (1905) bivalent system, the Frege-Strawson notion of truth value precondition with a third truth value option – namely *no value* – is more widely accepted. Several widespread semantic definitions of presuppositions are based on this notion, for instance:

(14) A presupposition is a condition under which a sentence expressing an assertive proposition can be used (in standard uses) to state a truth or a falsehood.
(Katz 1973, p. 256)

(15) A sentence A semantically presupposes another sentence B iff:

[6] This notion has been further developed by several theorists in the following years, such as Van Fraassen (1966) and Katz (1973), among others.

a. in all situations where A is true, B is true.

b. in all situations where A is false, B is true.

(Levinson 1983c, p. 175)

(16) Seien p und q (möglicherweise partielle) Propositionen. Dann ist q eine semantische Präsupposition von p gdw. q bei allen Welt-Zeit-Paaren wahr ist, bei denen p wahr oder falsch ist.[7]
(Heim 1991, p. 494)

(17) One sentence presupposes another iff whenever the first is true or false, the second is true.
(Beaver et al. 2021)

Based on these semantic definitions and following the Frege-Strawson notion, it is also commonly accepted to represent presuppositions as domain restrictions in formal semantics. Take, for example, Kratzer and Heim (1998): they assume that in [*the* ζ], the definite article *the* is a function that applies to another function *being* ζ and returns a unique entity, namely the referent. In order to improve the prediction on empty description, where the assertion without a suitable, unique referent for [*the* ζ] should not be simply false but rather without truth value, Kratzer and Heim (1998) add a limitation of $[\![the]\!]$'s domain. Together with the interpreted function of *the*, the lexical entry of the definite article is translated in formal semantics as follows:

(18) $[\![\text{the}]\!] =$
$\lambda f : f \in D_{<e,t>}$ and there is exactly one x such that $f(x) = 1$.
the unique y such that $f(y) = 1$
(Kratzer and Heim 1998, p. 75)

In formulation (18), the first part (before the dot) is the restriction of the domain while the second part (after the dot) is the truth condition. Only when the domain restriction is fulfilled, the function is defined and the truth condition can play its roll, otherwise there is no need to consider the truth condition and the whole expression is judged as *undefined*.

Applying the formula (18) to the determinate phrase (DP) *the king of France* and the sentence (S) in (9), following Kratzer and Heim (1998), the formal representation of a sentence with a presupposition according to the semantic definitions is illustrated as follows:

(19) $[\![\text{the king of France}]\!]$
$= [\![\text{the}]\!][\![\text{NP}]\!]$

[7]The definition might be translated into English as follows: Let p and q be (possibly partial) propositions. Then q is a semantic presupposition of p iff q is true in all world-time pairs where p is true or false.

$= [\lambda f : f \in D_{<e,t>}$ and there is exactly one x such that $f(x) = 1$. the unique y such that $f(y) = 1]([\lambda z : z \in D_{<e>}.z$ is the king of France])
= the unique y such that $[\lambda z : z \in D_{<e>}.z$ is the king of France](y)= 1
defined if and only if there is exactly one x such that $[\lambda z : z \in D_{<e>}.z$ is the king of France](x)= 1
= the unique y such that y is the king of France
defined if and only if there is exactly one x such x is the king of France.

(20) ⟦S⟧
= ⟦VP⟧(⟦DP⟧)
= ⟦lived in Versailles⟧(⟦DP⟧)
= [$\lambda z.z$ lived in Versailles](⟦DP⟧)
= 1 iff ⟦DP⟧ lived in Versailles
= 1 iff the unique y such that y is (was) the king of France and lived in Versailles
defined if and only if there is exactly one x such x is the king of France

2.1.2 Potential Limitations and Weaknesses

Although defining presuppositions as preconditions for the sentence truth value and representing them as domain restrictions can (at least partially) explain their generation and projection, the semantic definitions are still under question.

Firstly, the semantic definitions directly or indirectly predict that as long as a sentence can be interpreted as true or false, the presupposition is true or fulfilled. This assumption has in particular difficulties to explain the *filtering* of presuppositions, firstly defined by Karttunen (1973). Compare for instance the two sentences in (21): unlike (21a), the sentence (21b) can be naturally asserted without requiring one and exactly one king of France:

(21) a. If the king of France lived in Versailles, the king of France was happy.
$\rightsquigarrow$ There was one and exactly one king of France.
b. If there was a king of France, the king of France was happy.
$\not\rightsquigarrow$ There was one and exactly one king of France.

Moreover, as truth value preconditions in the semantic definitions and domain restrictions in formal semantics, presuppositions are described rather as a homogeneous class with stable projection, since the operators from the family of sentences (Chierchia and McConnell-Ginet 2000) can only target the truth value conditions. Therefore, these definitions or assumptions can neither predict nor explain the various projection behaviors within the class of presuppositions. As discussed before in Introduction, this heterogeneity has been observed at least since Karttunen (1971b) on factive and semi-facitive verbs as

shown in (6), repeated below in (22) again (for more about the difference between factive and semi-factive verbs, see Chapter 9):

(22) a. If I regret later that I have not told the truth, I will confess it to everyone.
$\rightsquigarrow$ The speaker had not told the truth.

b. If I discover later that I have not told the truth, I will confess it to everyone.
$\rightsquigarrow / \not\rightsquigarrow$ The speaker had not told the truth.

(Karttunen 1971b, p. 64)

Last but not least, several semantic definitions are explicitly or implicitly based on the assumption that the truth of presuppositions is not only guaranteed by the assertion but also when the assertion is negated. In this sense, presuppositions are considered a special sub-set of entailment. This assumption is challenged by the current observations, such as in the non-monotonic context, not all presuppositions are obligatorily entailed, firstly observed by Sudo (2012) (see Section 3.3 for more details). Moreover, Schlenker (2021b) also proposes that presuppositions do not need to be semantically entailed by the expression, but are rather dependent on epistemic and contextual conditions.

Considering all limitations and weaknesses mentioned above, it is reasonable to conclude that more linguistic aspects should be taken into account for presupposition investigation. In fact, at least since Stalnaker (1972), it has been proposed that semantics are not enough and a pragmatic notion of presuppositions is needed. According to Stalnaker (1972), the semantic and pragmatic definitions do not contradict each other, but rather explain the phenomenon from different perspectives or with different ideas. Still, the pragmatic aspects are relevant and necessary in order to explain presuppositions and their behaviors; some relevant reasons mentioned in his papers are listed below:

(23) a. in conversation, it is rather the speaker who presupposes something – if an expression can trigger a presupposition semantically, then a speaker using that expression also presupposes it pragmatically, but not vice versa (Stalnaker 1972, 1977).

b. there are situations in which the failure of presuppositions does not affect the truth value of the assertion, so that "there need be no essential connection between presupposition requirements and truth value gaps" (Stalnaker 1973, p. 454);

c. false presuppositions are also allowed in conversation, especially if they are irrelevant, or used on purpose by the speaker, such as for deception or avoiding deflection (Stalnaker 1998).

In sum, mainly based on the philosophical and early linguistic research on definite expressions, presuppositions are widely accepted to be defined as truth value precondi-

tions in the traditional semantic accounts. However, with regard to the limitations and weaknesses of the semantic definitions, pragmatic aspects are considered important and necessary. Thus, in the next section, several relevant definitions and assumptions from the pragmatic perspective are reviewed.

2.2 The Evolving Pragmatic Accounts

Linguistic research aiming at a pragmatic notion or definition of presuppositions can be traced back at least to Stalnaker (1972, 1973), Kempson (1975), and Wilson (1975). Among the early assumptions, the most influential one is given by Stalnaker who, as mentioned before, argues that it is not a word or a sentence but rather a speaker who can presuppose something (Stalnaker 1972, 1973, 1977, 1998). For a pragmatic notion of presuppositions, two concepts or aspects are considered essential: the felicity of the conversation, and the constraints on the common ground (also described as mutually shared beliefs, see also the summary in Levinson 1983c). The pragmatic definitions of presuppositions are generally based on these two concepts, also the most recent one:

(24) A speaker presupposes that P at a given moment in a conversation just in case he is disposed to act, in his linguistic behavior, as if he takes the truth of P for granted, and as if he assumes that his audience recognizes that he is doing so.
(Stalnaker 1973, p. 448)

(25) An utterance A pragmatically presupposes a proposition B iff A is appropriate only if B is mutually known by participants.[8]
(Levinson 1983c, p. 205)

(26) A pragmatic presupposition associated with a sentence is a condition that a speaker would normally expect to hold in the common ground between discourse participants when that sentence is uttered.
(Beaver et al. 2021, Section 3)

Stalnaker's pragmatic notions of presuppositions inspired two relevant pragmatic accounts: the first one claims that presuppositions are triggered uniformly and conventionally (or at least consistently) and can be canceled afterwards if they are infelicitous in the context. This account can be traced back to Stalnaker's (1977) analyses of the difference between factive and semi-factive verbs and is reviewed in Subsection 2.2.1. The second one, based on Stalnaker's assumption and the Gricean theory (Grice 1989[9]), as-

[8]Levinson (1983c) points out that the requirement *mutually known* can be too strong and an improvement is needed (see also the review in 2.2.1 and 2.2.2). In the context of this section, though, it is sufficient to illustrate the basic ideas of pragmatic presuppositions.

[9]Note that the work *Logic and conversation* was firstly published in 1967 in William James Lectures, Harvard University.

sumes that there are different kinds of presupposition triggering, and some of them might be similar to the Gricean conversational implicatures. Theories belonging to this account are reviewed in Subsection 2.2.2.

2.2.1 Presuppositions as Conventional Meaning

As reported by Levinson (1983c), the discussion on treating presuppositions as conventional meaning started with two advanced theoretical sub-accounts. The first sub-account is mainly represented by Karttunen and Peters (1975). According to their theory, conventional implicatures and presuppositions can be linked through the two essential aspects mentioned before, namely the felicity of the conversation and the constraints on the common ground, as follows:

(27) As a general tendency, it is in the interest of participants in a discourse to organize their contributions in such a way that the conventional implicata of the sentence uttered are already part of the common ground at the time of utterance.
(Karttunen and Peters 1975, p. 269)

Treating presuppositions as conventional implicatures, which are always carried by the expressions and triggered consistently and context-independently (see Levinson 1983a; Grice 1989 for more about conventional implicatures), Karttunen and Peters (1975) claim that pragmatic presuppositions also neither belong to the truth conditions nor can be really canceled or removed. This claim is doubted by Levinson (1983c), as a difference in cancellation can be observed between the two. Compare the sentences below with the presupposition triggered by *manage* and with the typical conventional implicature carried by *but*:

(28) a. John didn't manage to stop - he didn't even try. (Levinson 1983c, p. 210)
b. ?? The Duke of Norfolk has three mansions, but only one car, and there is in fact no contrast between these two facts. (Levinson 1983a, p. 129)

While the sentence in (28a) sounds felicitous and the presupposition can be naturally removed via the negating context, the same does not work with the classical conventional implicature triggered by *but*. In other words, presuppositions are not as impervious as classical conventional implicatures to contextual negation.

Moreover, as mentioned before in (26) and footnote 8, presuppositions do not have to be mutually shared or already part of the common ground. It is also sufficient if such information is normally shared or can be assumed. This property has been observed at least since Karttunen (1974). Take (29) for example: even if the addressees do not know that children cannot accompany their parents to commencement exercises, the sentence

can still be properly uttered and understood, as long as there is nothing in the context that hinders the addressee from assuming it:

(29) We regret that children cannot accompany their parents to commencement exercises.
(Karttunen 1974, p. 191)

Therefore, the requirement in (27) on the common ground is also too strong. Due to its predictions of resistance against cancellation and the too strong requirement on the common ground, Karttunen and Peters's (1975) sub-account shares some difficulties with the semantic accounts and will not be further discussed here (for more about this sub-account, see the review in Levinson 1983c).

In contrast to Karttunen and Peters (1975), in the other sub-account, cancellation is possible and several cancellation mechanisms are suggested. This sub-account can be traced back at least to Gazdar (1979a,b), who distinguishes actual presuppositions from potential presuppositions of a sentence. Based on this distinction, Gazdar claims that the latter can be understood as "something given to us by the lexicon and the syntax" (Gazdar 1979b, p. 124) and consists of presuppositions of all constituents in the sentence. However, not all of them can be actually presupposed: in Gazdar's model, the implied meanings of a sentence are added in a particular order – first entailments, then conversational implicatures, and presuppositions at the end. Therefore, among all potential presuppositions, only those who do not cause infelicity with entailments and impicatures can be actually presupposed, while the others are removed. Based on this cancellation mechanism, Gazdar's prediction about the presuppositions of a sentence can be summarized as:

(30) Implicatures and entailments defeat presuppositions, so a hearer adds to his or her commitments only those presuppositions that are compatible with both implicatures and entailments. All remaining presuppositions are canceled.
(Beaver et al. 2021, Section 3)

With Gazdar's notion, the difficulties with (28a) can be simply solved: the potential presupposition of *manage p*, namely *try p*, contradicts the entailment of the second conjunction in the sentence. As entailments are added before presuppositions, the potential presupposition can not be added and therefore not be actually presupposed.

Moreover, this mechanism can also explain the difference between the two sentences in (21), repeated below as (31):

(31) a. If the king of France lived in in Versailles, the king of France was happy.
$\rightsquigarrow$ There was one and exactly one king of France.

b. If there was a king of France, the king of France was happy.

$\not\rightsquigarrow$ There was one and exactly one king of France

According to Gazdar's model, both sentences have the potential presupposition *there was one and exactly one king of France*. However, the first conjunction of sentence (31b) conversationally implies that the speaker does not know whether there was a king of France. As conversational implicatures are added before presuppositions, the potential presupposition has to be canceled in (31b) due to inconsistency. On the other hand, there is no contradiction in (31a), so the potential presupposition can naturally become actually presupposed.

The model proposed by Gazdar (1979a,b), especially the idea of explaining presuppositions (and their projection or filtering) by means of constant triggering and cancellation in certain conditions, has been adopted, developed, and improved by a number of theories. For instance, as it is unclear how Gazdar's order for adding implied meanings is determined, Van der Sandt (1988) proposes an alternative mechanism, assuming that the presuppositional content of a constituent can be actually presupposed, or in other words, be projective, only if it can be expressed before the sentence without threatening the felicity. Take again (31) for example: the projection and filtering of the presuppositions triggered by the definite expression can be explained as below:

(32) a. There was one and exactly one king of France, if the king of France lived in Versailles, the king of France was happy.

$\rightsquigarrow$ There was one and exactly one king of France.

b. ? There was one and exactly one king of France, if there was a king of France, the king of France was happy.

$\not\rightsquigarrow$ There was one and exactly one king of France

Another tool which is often used to "repair" the conventional, consistent triggering is local accommodation. According to Lewis (1979) and Heim (1982), once presuppositions are not satisfied, accommodation takes place with the aim of adjusting the global or local context and achieving a felicitous reading. Assuming that the presuppositions are conventionally triggered and then canceled, the not-projection or filtering is explained primarily with local accommodations (see also discussions in Karttunen 1974; Heim 1983; Horn 1990; von Fintel 2004, 2008). Take (22b) for example, rewritten below in (33). The question with *discover* can receive two readings, where the one with local accommodation does not presuppose the complement, at least not globally on the discourse level:

(33) If the speaker discovers later that s/he has not told the truth, s/he will confess it to everyone.

a. With local accommodation: If the speaker has not told the truth and s/he discovers later that s/he has not told the truth, s/he will confess it to everyone.

$\not\rightsquigarrow$ The speaker has not told the truth.

b. With global accommodation: The speaker has not told the truth. If s/he discovers later that s/he has not told the truth, s/he will confess it to everyone.

$\rightsquigarrow$ The speaker has not told the truth.

Besides the theories mentioned above, there are further discussions and improvements of the cancellation sub-account, such as Horton and Hirst (1988), Mercer (1992), and Marcu (1994), which cannot be all reviewed here. The last framework reviewed in this section proposes the cancellation via at-issuness. These theories are worth mentioning not only because they are, according to Beaver et al. (2021), one of the most recent development of the cancellation-based theories, but also because at-issueness, a concept rooted in the question-driven discourse model (see von Stutterheim and Klein 1989; Roberts 1996; Simons et al. 2010, among others), offers a relevant alternative perspective on the various projection behaviors of presuppositions and is therefore one of the key assumptions investigated in the empirical part of this dissertation. Thus, this account is only briefly introduced here, as a detailed review on the theoretical and empirical background is given in Section 3.2.

Describing it in a very abridged form, the question-driven discourse model assumes that every discourse aims to provide a complete or partial answer to a question, which can be either explicit or implicit – the Question Under Discussion (QUD). According to Simons et al. (2010), the QUD can be understood as a set of alternatives that relate to the current topic in discourse, and parts of meaning that address the QUD are at-issue, whereas parts that are not relevant with respect to the QUD are not-at-issue. As semantic operators target only contents that are at-issue, this results in the survival of not-at-issue material. In their model, projection is then a property of not-at-issue discourse components: the projective behavior of presuppositions or the actual presupposing is then only a consequence of presuppositions being conventionally marked as not-at-issue.

Therefore, following Simons et al. (2010), a presupposition can lose its projection or stop being actually presupposed if it can be interpreted as relevant to the QUD and becomes at-issue. Take again (33) for example: the complement of *discover*, namely that *the speaker has not told the truth*, can be presupposed if the QUD focuses on *whether discover*. However, such a presupposition can be canceled if the QUD is about whether s/he has told the truth, which makes the complement at-issue (see also discussions in Simons 2007):

(34) a. Does the speaker discover that s/he has not told the truth?
If s/he discovers later that s/he has not told the truth, s/he will confess it to everyone.

$\rightsquigarrow$ The speaker has not told the truth.

b. Has the speaker told the truth?

If s/he discovers later that s/he has not told the truth, s/he will confess it to everyone.

$\not\rightsquigarrow$ The speaker has not told the truth.

However, the assumptions of a cancellation mechanism have to deal with one question: if all presuppositions are triggered conventionally, or at least similarly and consistently, then why is the sentence with *regret* still preferably understood with global projection, as showed in (35)? Why does a repair with local accommodation or an at-issue reading not take place? In other words, why is a local reading without global projection of presuppositions, no matter due to which cancellation mechanism, more easily and naturally available for *discover* than for *regret*?

(35) If the speaker regrets later that s/he has not told the truth, s/he will confess it to everyone.

Default reading (with global projection): The speaker has not told the truth. If s/he regrets later that s/he has not told the truth, s/he will confess it to everyone.

$\rightsquigarrow$ The speaker has not told the truth.

A possible answer to this question is offered by theories assuming that (at least some) presuppositions are not conventionally, but rather conversationally triggered. These theories will be reviewed in the next subsection.

2.2.2 Presuppositions as Conversational Implicatures

In contrast to those theories that take presuppositions for conventional meaning with consistent triggering, and explain their various projection behavior with cancellation mechanisms, the other relevant pragmatic account claims that the triggering of presuppositions is not always the same. That is, some presuppositions are not conventional but triggered similarly to Gricean conversational implicatures. This assumption can be traced back at least to Simons (2001), who observes that some presuppositions share certain typical properties with conversational implicatures, such as contextual defeasibility and nondetachability (for more about conversational implicatures, see e.g. Grice 1989; Horn 2004).

Let us start with the contextual defeasibility, which, according to Simons (2001), is shared particularly between some presuppositions and generalized conversational implicatures (GCIs). According to Grice (1989), although the triggering of GCIs is default, robust, and less dependent on context, but it can still be defeated via context, as can be seen in the examples (36) and (37). Simons proposes that such contextual defeasibility can also be observed with respect to some presuppositions, particularly in contexts

where the speaker explicitly ignores the potential presupposition which normally emerges – the explicit ignorance context. Interestingly, not all presuppositions are defeasible in this explicit ignorance context, some of them still remain and cause contradiction with the assertion. This difference is shown in Simons (2001, p. 433), repeated below in (38). Simons explains this difference by assuming that presuppositions of triggers like change of state predicates are conversationally triggered, while expressions like *again* trigger presuppositions conventionally.

(36) a. *Some* students pass the exam.

$\rightsquigarrow$ *Not all* students pass the exam.

b. George has *three* children.[10]

$\rightsquigarrow$ George has *exactly three* children.

(37) a. Some students pass the exam, and maybe all of them do.

$\not\rightsquigarrow$ Not all students pass the exam.

b. George has three children, and may have more for all I know. (Simons 2001, p. 434)

$\not\rightsquigarrow$ George has exactly three children.

(38) a. I have no idea whether Jane ever smoked, but she hasn't *stopped* smoking.[11]

$\not\rightsquigarrow$ Jane has smoked before.

b. # I don't know if Jane ever rented "Manhattan" before, but perhaps she' s renting it *again*.

$\rightsquigarrow$ Jane rented "Manhattan" before.

Nondetachability describes that conversational implicatures are less dependent on each single linguistic form. If the phrase is replaced with synonyms or similar expressions, the same implicature can still be triggered if the context remains the same. Simons argues that similar to conversational implicatures, some presuppositions can also be triggered by different expressions if the context is not changed, see (39b).

(39) a. Jane: Do you want to go out for a drink?
Julia: I have to finish writing my SALT paper/I need to finish my SALT paper/My SALT paper needs to get finished tonight.

$\rightsquigarrow$ Julia cannot go out for a drink.

[10] It is disputed whether numerals are triggers of GCIs or scalar implicatures at all (see e.g. discussions in Y. T. Huang et al. 2013; Y. Chen 2019). However, as it is not the main topic here and numerals were used in Simons (2001) to illustrate the contextual defeasibility, an example with numerals is still given here.

[11] For a natural reading of this sentence, Simons (2001) suggests to imagine a context in which Jane does not show any typical symptom of quitting smoking.

b. Jane didn't stop laughing/didn't quit laughing/didn't cease laughing.

$\rightsquigarrow$ Jane laughed before.

(Simons 2001, p. 435, sightly modified)

However, Simons also notices that in contrast to the contextual defeasibility, the nondetachability of presuppositions is more dependent on whether there are common synonyms of the trigger in a certain language than on whether the presupposition is conversationally or conventionally triggered, although there might be some overlaps. I will discuss nondetachability of presuppositions in more detail in Chapter 5.

Addressing the question of whether presuppositions can be suspended, a property similar to the contextual defeasibility discussed in Simons (2001), Abusch (2002, 2010) differentiates between two sub-classes of presupposition triggers: soft and hard. A contrast between these two trigger types can be seen in (40):

(40) John will either attend the first meeting, or miss it.

a. hard: #/? And he will either attend the second meeting *too*, or miss the second meeting *too*.

b. soft: And he will either *continue* attending meetings, or *continue* missing them.

(Abusch 2002, p. 4, slightly modified)

According to Abusch (2002, 2010), triggers like *too, also, even, again* and *it*-clefts are classified as hard, as they generate semantic presuppositions which are stronger and less suspendible. On the other hand, triggers like *continue, stop, win* and *know* are classified as soft, as their presuppositions arise pragmatically through their lexical alternatives.

One step further, some researchers propose that some presuppositions, especially those generated by soft triggers, are similar or even equivalent to scalar implicatures (SIs) or indirect scalar implicatures (ISIs) (see e.g. Chemla 2008, 2009; Romoli 2014; Romoli and Schwarz 2015). According to Horn (1976, 1984) and Grice (1989), among others, SIs and ISIs are special kinds of GCIs which are triggered by expressions that belong to a linguistic scale. Take the scale *<some, all>* for example: a SI can be triggered by using the weak term *some*, which implies the negation of the proposition with the stronger term, that is, *not all*, as shown above in (36). An ISI, on the other hand, is triggered by the negation of the strong term and implies the weak term, as shown in example (41) below. Comparing for instance (42) with (41), it seems possible to consider the presupposition *participate* as an alternative of the trigger *win* in an achievement scale. From this perspective, presuppositions of soft triggers can be defeated or canceled via context because they are also scalar-based conversational implicatures, so that principles for SIs/ISIs can be extended to them.

(41) *Not all* students pass the exam.
⇝ *Some* students pass the exam.

(42) The duck did not win the game.
⇝ The duck participated in the game.

However, note that Romoli (2014) claims that with scalar terms, the alternatives are symmetric, while this is not the case with soft triggers. For instance, *all* and *some* are alternatives to each other, while *participate* is an alternative of *win* but not vice versa, meaning that soft presupposition triggers are only 'scalar' under negation and *participate* does not behave like *some* in affirmative sentences.

In a similar vein but sightly different from the theories mentioned above, Sauerland (2008) and Yatsushiro (2008a,b) propose that it is rather the presuppositions, not the triggers, which need to be classified. They claim that in contrast to lexical presuppositions, the triggering process of implied presuppositions is analogical to that of SIs. The distinction between lexical and implied presuppositions can be traced back to Heim (1991) and her investigation on indefinite articles. According to Heim (1991), an indefinite article like '*ein*' in German or *a* in English can trigger two presuppositions: the existence presupposition and the anti-uniqueness presupposition. Compare the two sentences below:

(43) I interviewed a biological sister of the victim.
   a. existence presupposition: There is one biological sister of the victim.
   b. anti-uniqueness presupposition: There is more then one sister of the victim.

(44) # I interviewed a biological father of the victim. (Yatsushiro 2008b, p. 665)
   a. existence presupposition: There is one biological father of the victim.
   b. anti-uniqueness presupposition: There is more then one biological father of the victim.

Unlike (43), the sentence (44) is quite unnatural. According to our world knowledge, a person can have more than one biological sister, but only one biological father. Therefore, the definite expression, namely *the father*, should be used in this case. However, as the definite article is not used, the hearer is then allowed to imply that it is not the case that there is one and exactly one father of the victim, which sounds therefore odd and unusual. Considering these observations, Heim (1991) proposes a new maxim for presuppositions that is analogical to Grice's first Sub-maxim of Quantity ('make your contribution as informative as is required', Grice 1989, p. 26). This maxim is cited below in (45a), which can be translated as in (45b). Following this maxim, the anti-uniqueness presupposition of the indefinite article is implied scalar-like because of the uniqueness presupposition of the definite article, in contrast to the existence presupposition which is semantically required.

(45) a. Präsupponiere in deinem Beitrag so viel wie möglich! (Heim 1991, p. 515)

b. Make your contribution presuppose as much as possible! (Sauerland 2008, Section 2)

The scalar-based assumptions on presuppositions or their triggers have been investigated experimentally. On the one hand, Yatsushiro (2008a,b) for instance shows by means of a comparison between adults and children that the implied presuppositions, such as the anti-uniqueness of the word *every* in German, are indeed acquired later than the existence presupposition. Moreover, Romoli and Schwarz (2015) observe in their experiment that the literal meaning of both presuppositions and ISIs is interpreted faster than their pragmatic enrichment. However, on the other hand, these SI-analogical explanations are not without challenge. Bill et al. (2016) for instance compare the comprehension of presuppositions, SIs and ISIs among adults and children. Their results show that for both children and adults, there is a clear difference between the understanding of presuppositions of soft triggers, SIs and ISIs, and the acquisition process of presuppositions differs from that of SIs/ISIs, too (for more details on these studies, see the review in Section 4.1).

Although they rely on different criteria, theories from this account generally propose a classification of presuppositions or their triggers. Among these, the terms of soft and hard triggers (Abusch 2002, 2010) as well as the two-way split have been widely discussed and investigated in the last two decades. As they also belong to the key assumptions investigated in my experimental studies, more details about the soft-hard classification are reviewed in Section 3.1.

In this chapter, regarding the questions on what presuppositions are and how their projection, filtering, or local reading can be explained, some (but of course not all) relevant definitions and assumptions from both semantic and pragmatic accounts were reviewed.

Starting with the research on definite expressions, the traditional semantic theories widely accept that presuppositions can be conceptualized as preconditions for a sentence's truth value and thus as a special sub-set of entailments whose truth can also be guaranteed by the negated assertion. Correspondingly, they are represented as domain restrictions in formal semantics (see Subsection 2.1.1). However, as summarized in Subsection 2.1.2, these semantic definitions and representations have their limitations: they have difficulties to explain several observations, such as the filtering of presuppositions or their various projection behaviors, and they face challenges like presuppositions being non-obligatorily entailed in certain context (see e.g. Sudo 2012).

As a consequence, the importance of pragmatic aspects of presuppositions has been increasingly recognized since Stalnaker (1972, 1973), with regard to two essential notions: the felicity of the conversation and the constraints on the common ground. Inspired by Stalnaker, two relevant pragmatic accounts were reviewed, which can be summarized as

follows: the first account assumes a uniform, conventional triggering of presuppositions, while the filtering and projection variations are mainly explained with cancellation or repair under certain conditions or according to certain rules (see Subsection 2.2.1). In contrast, the second account assumes that presuppositions are triggered differently. That is, while some of them are stored conventionally or semantically, some presuppositions are triggered or implied conversationally. This assumption requires a distinction of presuppositions or their triggers (see Subsection 2.2.2).

Both accounts considering pragmatic aspects offer several relevant assumptions regarding the heterogeneous projection behaviors of presuppositions. As not all assumptions can be reviewed and examined in my work in detail, my study mainly concentrates on one assumption from each account in particular: the first one focusing on at-issueness, taken from the first account, and the second one proposing a soft-hard classification, taken from the second account. As the heterogeneity of presuppositions is the first main research concern of my study, these two most relevant assumptions for my research will be introduced in detail in the next chapter.

3 The Heterogeneity of Presuppositions

Considering the various behaviors and differences within the category of presuppositions, partially already reviewed and discussed in Chapter 2, see (22), (38) and (40) for example, it is widely assumed that "[t]he zoo of presupposition triggers should have been constructed with separate cages for different species" (Karttunen 2016, pp. 706–707). The heterogeneity of presuppositions and controversies about trigger classification are the first main research concern of my work. As this issue is more complex than one chapter can cover, this chapter offers only an overview on assumptions that are relevant for my experiments. I will start with the soft-hard split of triggers, the main assumption originally given by the theoretical account which assumes that some presuppositions are pragmatically or conversationally triggered while the others are not. The second relevant assumption introduced in detail in this chapter is based on the notion of at-issuness, an assumption that explains the projection and cancellation of presuppositions with respect to the Question Under Discussion (QUD, see e.g. von Stutterheim and Klein 1989; Roberts 1996; Simons et al. 2010). In the last section, some other classification possibilities that might be interesting for the discussions in Part II and III are briefly reviewed.[12]

3.1 The Classical Soft-Hard Classification

As mentioned before, the idea to classify presupposition triggers based on their different triggering processes can be traced back at least to Simons (2001), who observes that some presuppositions share certain typical properties with conversational implicatures (for more, see the review in Subsection 2.2.2). The terms *soft* and *hard* trigger, on the other hand, are firstly suggested by Abusch (2002): she distinguishes soft triggers from hard triggers by claiming that while hard triggers introduce presuppositions semantically, the presuppositions of soft triggers are generated pragmatically through lexical alternatives. Take the typical soft trigger *win* for example: its lexical alternative is *lose*, and the triggering process of the presupposition *participate* can be summarized in three steps:

(46) Abusch's (2002) pragmatic triggering of presuppositions:

a. Step 1: Using a soft trigger like *win*, the speaker introduces an alternative set C into the discourse: {*win, lose*}

[12]Note that the review in this chapter is later summarized as part of the research background in Chapter 7, which has been published as Y. Chen et al. (2022).

b. Step 2: The speaker pragmatically assumes that some alternatives in the set C are true: *win* $p \vee$ *lose* $p = 1$

c. Step 3: The overlap of these alternatives is pragmatically presupposed:
win $p \vee$ *lose* $p = 1$
$\Rightarrow$ *participate in p and be No.1* $\vee$ *participate in p and not be No.1*$=1$
$\Rightarrow$ *participate in* $p = 1$

Furthermore, Abusch (2002, 2010) claims that besides achievement phrases like *win*, change of state verbs like *stop, wake up, freeze over* and (semi-)factive verbs like *know, discover, be right* are also soft triggers with pragmatic presuppositions, and that they are different from hard triggers like *too, also, even, again* and clefts. Interestingly, the mechanism proposed by Abusch is in fact also influenced by QUD theories, not only in that both conceptualize an utterance as the contribution of alternatives to the discourse, but also in that they both assume that at least one alternative in the set is true.

In order to distinguish soft from hard triggers, Abusch (2010) adopts the explicit ignorance context from Simons (2001), see for instance Simons' examples in (38) above, or below in (47) with a contrast between the soft trigger *win* and the hard trigger *too*. In this context, the speaker explicitly ignores the truth of the presupposition. According to Abusch, triggers that are acceptable in this context can be classified as soft, while the hard triggers can cause infelicity:

(47) Test with explicit ignorance context:

a. I don't know whether Pikachu participated in the game, but if he won the game, he was happy.

b. # I don't know whether anyone else was ill, but if Pikachu was ill too, his friends were worried.

One widely accepted explanation for this felicity-infelicity contrast is that while in (47b), the presupposition of the hard trigger *too* is accommodated globally, projects up to the highest extent and then contradicts the proposition, the soft trigger *win* in (47a) receives a local reading through the local accommodation, so that it is interpreted under the *if*-embedding (for more about local accommodations, see the review in Subsection 2.2.1). The default reading of the two sentences can be described as follows:

(48) a. default understanding of (47a) with a local reading: I don't know whether Pikachu participated in the game, but if *he participated* and won the game, he was happy.

b. default understanding of (47b) with a global reading: # *Someone else was ill,* and I don't know whether anyone else was ill, but if Pikachu was ill too, his friends were worried.

As a consequence of these observations, the following questions arise: why is the local reading more available for soft triggers, and why are hard triggers impervious to this strategy? According to the mechanism proposed by Abusch (2002, 2010), soft triggers can be repaired automatically with local accommodation in the explicit ignorance context because the second step in the process, namely assuming that some alternatives in the set C are true, is in fact optional, and the presupposition of *win* is generated pragmatically. Therefore, it does not obligatorily project and can be easily suspended if needed. In contrast, the presupposition of *too* is stored lexically and triggered semantically, so its projection strength is stronger and it can hardly be suspended.

Another context in which the difference of presuppositions' projection behavior can be illustrated is the scope of quantifiers like *some*, *at least two* or *less than three* (see e.g. Charlow 2009). Adapting the summarized contrasts from Romoli (2014), the difference between triggers like *also* and *stop* can be illustrated as following:

(49) a. Some of these students *also* smoke[s] $Marlboros_{Focus}$.

b. More/Less than three of these students *also* smoke $Marlboros_{Focus}$.

$\rightsquigarrow$ Each of these students smokes something other than Marlboro.

(50) a. Some of these students *stopped* smoking.

b. More/Less than three of these students *stopped* smoking.

$\not\rightsquigarrow$ Each of these students used to smoke.

(Romoli 2014, p. 176, slightly modified)

Based on the assumption of different projection behaviors under quantifiers, weak triggers like change of state verbs are distinguished from strong triggers like factive verbs (see e.g. Chemla 2008). Moreover, a uniform approach for presuppositions, at least those of soft or weak triggers, and scalar terms is proposed (see e.g. Chemla 2008, 2009; Romoli 2014; Romoli and Schwarz 2015 and Subsection 2.2.2 for more detail). However, this weak-strong distinction is empirically challenged: although Chemla (2009) observes in his first experiment that once embedded by the quantifier *no*, the universal presuppositions of change of state verbs like *stop* and *continue* were less robust than presuppositions of factive verbs like *know* and *be aware*, this difference disappeared in his second experiment, in which the rating scale had been changed from binary to graded. The uniform approach, on the other hand, is challenged by experimental findings as well. Besides the difference in language acquisition observed by Bill et al. (2016) (reviewed in detail in Section 4.1), presuppositions and scalar terms also differ in their reaction times: while a sentence with false scalar imlicatures is rejected with more reaction time, as illustrated in Bott and Noveck (2004), Chemla and Bott (2013) observe in their study that there was no such delay with respect to presuppositions which do not globally project. Additionally, Bill et

al. (2018) show in their experiment that once the presupposition triggers were prosodically stressed, participants were more likely to interpret them with presuppositions than in the written form. In contrast, with scalar terms, the prosodic stress increased the preference of the literal reading. Therefore, although the weak-strong distinction suggested as part of this sub-account is also two-way split and more or less similar to Abusch (2002, 2010), it will not be further discussed.

Note, however, that the soft-hard distinction of presupposition triggers is also controversial. Theoretically, Abrusán (2016) for instance claims that a classification of triggers is not needed because all presuppositions are "fundamentally the same type" (p.168) and their difference is caused by the interaction between triggers and other effects like focus, anaphoricity, context and the generating process. In this process, not only the trigger is influential but also the relevance of the information with respect to the main point of the sentence (see also Abrusán 2011). This assumption is rather in line with the at-issueness approach, which will be reviewed in the next section. Briefly summarized, according to Abrusán (2011, 2016), contents can be considered presupposed if they are, by default or contextually, not paid attention to: while the (potential) presuppositions of triggers like *discover, win, manage* can be paid attention to in certain contexts if needed, the presuppositions of additive particles or *it*-clefts can hardly play such a role due to their lexical or syntactic properties.

The observations in empirical studies on the soft-hard distinction are also controversial. Regarding global accommodation, Spenader (2002) supports quantitatively that the difference between presuppositions may depend on trigger types. Focusing on local accommodation on the other hand, Jayez et al. (2015) investigate the possibility of a local reading of the typical hard trigger *too* in French with the explicit ignorance test. In their experiment, two discourse structures were used: a contrast one with *but*, and an explanation one with *because*. The critical items with *too* were compared to control item without *too*. For illustration, one example set from their test is given below. Participants were asked to rate the naturalness of the following items, given a context in which the two persons, Paul and Mary, should not meet each other:

(51) a. I don't know whether Paul will go to the party but, if Mary goes ∅/too, it will be embarrassing.

b. I don' t know whether Paul will go to the party because, if Mary goes ∅/too, it will be embarrassing.

(Jayez et al. 2015, Section 1.3.1, slightly modified)

Results show that in both *because*- and *but*-structures, the presence of *too* increased the naturalness. According to Jayez et al. (2015), this indicates that even for a hard trigger like *too*, a local reading is still possible. Together with the observations in their second

experiment on *regret* and clefts, Jayez et al. doubt the two-way split based on the availability of local readings or certain lexical properties of triggers but instead argue for the critical influence of contexts.

Adopting the covered box task from Y. T. Huang et al. (2013), Zehr et al. (2017) test the local reading possibility of another typical hard trigger, *again*, with a picture-matching task. In line with Jayez et al. (2015), their experiments also show that non-global readings are possible for *again*, especially in *neither-nor* sentences, even though they are not preferred. Moreover, they also argue that the difference with respect to the local reading possibility is not two-way split, but rather gradual. This assumption is also supported by Amaral and Cummins (2015), who consider the interaction between lexical meaning and contextual information as the most salient predictor, too (for more about this study, see Section 5.3).

In short, the experimental studies so far have illustrated that, contrary to original predictions, there is no absolute prohibition of a local reading for typical hard triggers, as its availability is rather context-dependent. Another crucial challenge is that instead of a clear two-way split which can support the soft-hard distinction, the difference between triggers was observed rather gradient.

However, despite the controversies with regard to the details, what the theoretical and empirical studies have in common – and what is most relevant for my study as well – is that first, there are indeed some presuppositions which can receive a local reading more easily than others and therefore do not always globally project. Second, although in interaction with other contextual effects, triggers do play a role in this process and more or less influence the availability of a local reading. Therefore, I will still use the term soft and hard trigger in my study, but in a wider sense – that is, soft is used to refer to those triggers whose presupposition is generally considered as more likely to be locally accommodated, or can be canceled or suspended (more) effortlessly, or assumed to be implied instead of semantically triggered, such as achievement verbs like *win* and *manage*, or semi-factive verbs like *discover*. Hard triggers, on the other hand, are expressions with presuppositions that normally strongly and globally project, or can hardly – although not impossibly – receive a local reading, or are classically assumed to be semantically, conventionally or at least consistently triggered, such as *too*, *again*, and *it*-clefts.

3.2 The At-Issueness Gradient

As shortly mentioned before in Subsection 2.2.1, at-issueness is a concept from the question-driven discourse model, in which the discourse is understood as the answer to an explicit or implicit question, the Question Under Discussion (QUD), mainly following Klein and Von Stutterheim (1987) ('Quaestio') and Roberts (1996). According to Simons

et al. (2010), the at-issuness of a sentence can be defined based on their relevance to the QUD as follows:

(52) Relevance to the QUD (Simons et al. 2010, p. 316):

a. An assertion is relevant to a QUD iff it contextually entails a partial or complete answer to the QUD.

b. A question is relevant to a QUD iff it has an answer which contextually entails a partial or complete answer to the QUD.

(53) A proposition p is **at-issue** relative to a question Q iff *whether* p is relevant to Q (Simons et al. 2010, p. 317, sightly modified).[13]

From the perspective of at-issueness, presuppositions are considered contents that are linguistically and conventionally marked as not-at-issue (Simons et al. 2010). Their not-at-issueness can be illustrated for instance with the following examples from Antomo (2016, p. 40), in which the embedded clauses provide, at least logically, an answer to the QUD. However, once embedded under a factive verb like *ignore* so that the complement becomes presupposed, it cannot suitably target the QUD anymore and the answer is odd. However, there is no such infelicity with non-factive verbs like *believe*:

(54) Q: Where is Homer?

A: # Marge ignores that he's at Moe's.

A′: Marge believes that he's at Moe's.

Based on this nature, Aravind and Hackl (2017) propose the Not-At-Issueness Constraint on presuppositions:

(55) Not-At-Issueness Constraint:
Presuppositions cannot be used to directly target the Question Under Discussion. (Aravind and Hackl 2017, p. 51)

The at-issueness notion also provides an explanation for projection by taking it as a property of discourse components: while entailment-canceling operators like negation or question can target only at-issue contents, the not-at-issue components are ignored by these operators and can thus project (Simons et al. 2010). In line with this assumption, it has been observed that some other not-at-issue but not-pre-required contents can also survive once they are embedded under operators from the family of sentences (Chierchia

[13]Note that in order to take the speaker intention into account as well, this definition is revised in Simons et al. (2010, p. 323). According to the revised definition, not only the linguistic markers but also the intention of the speaker can influence the at-issueness. However, they also claim that the speaker intention itself is restrained by the relevance with regard to QUD, too. Thus, to briefly and clearly introduce the at-issueness account and their predictions on presuppositions and their projection, the simplified definition should be sufficient.

and McConnell-Ginet 2000) and project, too, such as the content of non-restrictive-relative clauses (NRRCs), expressive contents or nominal appositives (for more, see e.g. Potts 2005):

(56) a. NRRC: Togebi, who was ill, likes cake.
S?: Does Togebi, who was ill, like cake?
$\rightsquigarrow$ Togebi was ill.

b. expressive content: The damn Meowth stole the cake.
if S: If the damn Meowth stole the cake, we have to make a new one.
$\rightsquigarrow$ The speaker does not like Meowth.

c. nominal appositive: Pikachu, a professional baker, brings a new cake.
perhaps S: Perhaps Pikachu, a professional baker, brings a new cake.
$\rightsquigarrow$ Pikachu is a professional baker.

Thus, from the aspect of at-issueness, presuppositions are considered only a subset of not-at-issue and thus projective contents. In contrast to other projective sub-sets, presuppositions require that their content should be given, or at least can be assumed as given in discourse[14], while other non-presupposed not-at-issue contents like NRRCs should rather be new and novel (for more about the information status of NRRCs, see e.g. A. Holler 2005; Roberts et al. 2009; for further discussion on differences between presuppositions and non-presupposed projective contents, see also e.g. Venhuizen et al. 2013, 2015).

Furthermore, with regard to the explanation for projection in the context of the notion of at-issueness, a positive correlation is predicted, that is, the more not-at-issue a content is, the more possible it can be ignored by the operators, and thus, the more projective it is. This correlation has been investigated experimentally at least since Xue and Onea (2011). Xue and Onea (2011) conducted two experiments in German: in their first experiment, the projection strength was measured. Ratings confirm that different presupposition triggers have different probabilities to project. This result patterns with the predictions of the soft-hard approach in that the presuppositions of *know* and *find out* were less projective than those of *too* and *again*. However, note that the results of their experiments were more cascading than a sharp two-way split (highest projection probability to lowest: *again* > *too* > *find out* > *know*).

In their second experiment, Xue and Onea (2011) measured the degree of at-issueness of the above-mentioned presuppositions by testing whether a sentence employing them could be denied directly. That is, participants were asked to reject sentences containing

[14]Note however that Abbott (2000) argues that some presupposing constructions can convey new information.

presuppositions by choosing an answer out of three options, see examples in (57). By doing so, Xue and Onea (2011) assume that only the *no*-answer indicates that the presupposition is targeted directly. Therefore, the proportion of the *no*-ratings illustrates the at-issueness level of the presupposition:

(57) Tina has just found out that Max is on vacation.

a. Yes, and Max is not on vacation at all.

b. Yes, but Max is not on vacation at all.

c. No, Max is not on vacation at all.

(Xue and Onea 2011, p. 179)

Ratings show that in 73% of all cases, *know* was negated directly with '*No,...*', followed by *find out* (50%), *too* (26%) and *again* (22%). Eventually, Xue and Onea (2011) conclude that both experiments show a clear positive correlation between not-at-issueness and projection, namely the more not-at-issue, the more projective.

However, there are two problems with this conclusion: firstly, Snider (2017a) argues that tests relying on the possibility of a direct deny are not a suitable method to measure (not-)at-issueness. Instead, he claims that it is rather anaphoric availability that is tested with this method. Second, there is a fundamental problem with the experiment design: whereas the two answers with *yes* lead to inconsistency in the case of *know* and *find out*, the same problem does not arise for the two other triggers (for more details and discussion, see Antomo 2015).

More recently, Tonhauser et al. (2018) also investigate the correlation between projection and (not-)at-issueness empirically. In their first study, they test not only presupposition triggers but also other kinds of projective contents such as NRRCs and nominal appositives in their first sub-experiment, and diverse factive predicates in their second sub-experiment. In order to assess the projectivity and at-issuness, Tonhauser et al. (2018) designed two polar questions as shown below: the *'certain that' question* measures the projectivity, and the *'asking whether' question* measures the at-issueness. A *yes*-answer to the *'certain that' question* indicates that the context is projective, and a *yes* to the *'asking whether' question* means that the content is judged as at-issue:

(58) Patrick asks: Was Martha's new car, a BMW, expensive?

a. 'certain that' question (projectivity): Is Patrick certain that Martha's new car is a BMW?

b. 'asking whether' question (at-issueness): Is Patrick asking whether Martha's new car is a BMW?

(Tonhauser et al. 2018, p. 502)

In their first dual experiment, they observe that the presuppositions of expressions like *be annoyed, be amused* and possessive NPs were highly projective and also highly not-at-issue, while the presuppositions triggered by *discover, reveal* and *confess* were less projective and more at-issue. Moreover, they illustrate that the difference between triggers in their projectivity and not-at-issueness was rather a gradient than a two-way distinction.

The at-issueness of the expressions evaluated in their first study were double checked with the *Are you sure?*-diagnostic in their second study, one example is adopted below. After the first speaker has uttered a sentence with a trigger of projective contents, the second speaker always asks *Are you sure?* The first speaker then gives an answer saying that s/he is sure about the (potentially) projective content. Participants were asked to judge whether the first speaker actually answers the question by choosing between *yes*, which indicates that the content is at-issue, and *no*, indicating it is not-at-issue.

(59) Sandra: Shirley is aware that Raul was drinking chamomile tea.
Carl: Are you sure?
Sandra: Yes, I am sure that Raul was drinking chamomile tea.
(Tonhauser et al. 2018, p. 519)

Results show again that the presuppositions of triggers like *be annoyed, be amused* and possessive NPs were more not-at-issue than those of triggers like *discover, reveal* and *confess*, and the participants' rejection rates were also rather gradient than two-way split, confirming the observations of the first two experiments. However, when the *Are you sure?*-diagnostic was used, the relationship between projectivity and not-at-issueness seemed not as linear as in the first (dual) one.

In sum, despite their different methods and measurement for projection and at-issueness, both Xue and Onea (2011) and Tonhauser et al. (2018) illustrate experimentally that the heterogeneous projection behavior of presuppositions, especially their various projection strength, is correlated with their at-issueness status. In contrast to the soft-hard split, which assumes that there are two sub-classes of presupposition triggers, a gradient or graded continuum between presuppositions is rather expected from the at-issueness perspective. The insight from not-at-issueness and the prediction of a graded continuum of projection doubtlessly shed light on the investigation on the classification issue of presuppositions, providing a reasonable alternative explanation for their heterogeneity problem. However, one question that remains in this account is: why the presuppositions of some triggers are more impervious to an at-issue reading than others? In other words, the at-issueness account in fact expands and relocates the question of projection heterogeneity of presuppositions to the question of at-issueness heterogeneity of all projective contents. Therefore, it is reasonable to consider investigating the soft-hard distinction of presuppo-

sition triggers in the broader context of at-issueness, together with other projective but not presupposed contents.

3.3 Some Other Classifications

As mentioned above, there is a great amount of research on the heterogeneity problem of presuppositions with a focus on the trigger classification. Besides the two accounts reviewed before, presupposition triggers are also classified with respect to other properties like the similarity between presuppositions and anaphora (see e.g. Zeevat 1992; Kripke 2009), the obligatory vs. non-obligatory entailment relation between assertions and presuppositions (see e.g. Klinedinst 2016), the obligatory vs. optional repair regarding the failure (e.g. Glanzberg 2005), the strong contextual felicity and obligatory local effect (Tonhauser et al. 2013), the obligatory exhaustivity implicature in case of missing triggers (Bade 2016), and more. Naturally, they cannot all be outlined. Thus, only the first two of them are reviewed in this section, but only very briefly, as they might be interesting for the discussion later in the experimental part.

The idea of explaining the heterogeneity of presuppositions by indicating several sub-classes of triggers goes back at least to Kripke (2009) (orally presented in 1990) and Zeevat (1992), who classify triggers based on the similarity between presuppositions and anaphora.

Take Zeevat (1992) and his distinction of resolution and lexical triggers for instance: the initial observation for his classification is that some triggers are more anaphora-alike than the others. That is, these phrases generally tend to say something new about entities that have already been introduced in the context before, just like an anaphor. Those triggers, such as the definite article, factive *when*- and *after*-clauses and clefts, are classified as resolution triggers. Another relevant sub-class is that of lexical triggers, whose presupposition is claimed to be a lexical and logical precondition of a certain action or state, such as *regret* or *stop*. According to Zeevat, these preconditions have to be satisfied, as people can hardly regret something that is not the case or stop something that never took place (cf. the discussion and review in Section 2.2). Zeevat (1992) also mentions a third group of triggers, including words like *too, also, again*. Zeevat distinguishes them from the other two types by two properties: first, they are more anaphoric than the first group as they can refer to antecedents from contexts that are normally inaccessible for anaphora; second, their presuppositions can hardly be accommodated. Thus, in the experimental studies afterwards, this type is sometimes merged with the resolution sub-class, as they are both more anaphoric than the lexical triggers (see for instance Amaral and Cummins 2015).

The last classification I want to briefly mention in this chapter is the one based on the obligatory vs. non-obligatory entailment relation between assertions and presuppositions, as the initial observation of this account might challenge the semantic definitions of presuppositions fundamentally.

As discussed before, according to the traditional semantic account, presuppositions are allowed to be understood as a special subset of entailments, as their truth is not only guaranteed by the utterance, but also remains when the sentence is embedded under operators from the family of sentences (Chierchia and McConnell-Ginet 2000). This is also the basic assumption for numerous semantic definitions (see the review in Section 2.1). However, this claim is challenged by Sudo (2012), who observes that in a non-monotonic environment, such as *exactly one*, some presuppositions are not obligatorily entailed. To illustrate the difference between the two types of presuppositions, Sudo compares the truth value conditions of *herself* (namely with the presupposition of gender information) with *stop* (with the presupposition *used to do something*) in the non-monotonic context. This contrast is summarized in Zehr et al. (2016, p. 321) as follows:

(60) Exactly one student stopped using Mac.

  a. Exactly one student used Mac and does not use Mac now.

  b. Exactly one student does not use Mac now.

(61) Exactly one student criticized herself (…namely Mary).

  a. Exactly one student is a female who was self-critical.

  b. Exactly one student was self-critical.

According to Sudo (2012), for triggers like *stop*, the truth value condition with the presupposition like (60a) is preferred, suggesting that the presupposition is entailed. On the contrary, with presuppositions triggered by expressions like *herself*, the truth value condition like (61b), in which the presupposition is removed, matches speaker's intuition better. In contrast, readings like (61a) can also suit situations where there is more than one self-criticizing student, as long as there is only one female among them, which conflicts with the intuited interpretation.

As Sudo assumes that this absence of entailment can be changed by means of entailment-canceling operators like negation, he does not suggest a classification based on this observation. However, Klinedinst (2016) proposes this entailment contrast as another perspective to explain the trigger distinction. According to Klinedinst (2016), the suspension or a local reading is generally available for some triggers, such as Abusch's (2002) soft triggers, among others, because their presuppositions are obligatorily entailed by the assertions. On the other hand, assertions with hard triggers do not, or at least not

mandatorily, entail the presuppositions in a non-monotonic context. As a consequence, local readings seem less available with respect to hard triggers.

Empirically, this perspective is investigated by Zehr and Schwarz (2016, 2018). Using a covered box design, Zehr and Schwarz (2016) compared the presuppositions of *stop* with *also*, embedded in a monotonic context with *at least one* and a non-monotonic context with *exactly one*. Observing a clear contrast in the picture-choice preference and reaction time between these two triggers, they claim that their observation is in line with Sudo's (2012) and Klinedinst's (2016) prediction but also notice that the ratings on *stop* were more mixed than expected. In Zehr and Schwarz (2018), these two triggers were investigated with an *exactly* x-test again, but served rather as a baseline for other triggers in the experiment. The ratings are interpreted as being rather in favor of Sudo's (2012) and Klinedinst's (2016) prediction, too. Moreover, Zehr and Schwarz (2018) also manage to classify expressions like *return*, *again* and *back*, of which the former two were rated similar to *also* and therefore assigned to the hard sub-class, and the latter patterned with *stop* and were considered as rather soft.

In sum, besides the soft-hard split of presupposition triggers and the assumption about a correlation between projection strength and not-at-issueness, there are also plenty of other classification possibilities based on observations on various properties, such as anaphoric readings or obligatory entailment. These classifications can also offer relevant aspects to the research on the heterogeneity of presuppositions with interesting criteria for the distinction of their triggers.

In this chapter, two assumptions that more or less represent the two major lines for presupposition heterogeneity are reviewed in detail, as they are relevant for the classification issue of my study, followed by a brief summary of some other interesting classifications. Comparing the hypotheses of the two main assumptions, one may conclude that while the first one, which is based primarily on the availability or difficulty of local readings (see Section 3.1), aims at a clear, two-way split of presupposition triggers, the other one, which is derived mainly from the correlation between projection and not-at-issueness, predicts a gradient or continuum of projection strength (see Section 3.2). The question that follows then is which one is more capable or suitable to describe and explain the heterogeneity issue of presuppositions, or whether these two assumptions can be combined with regard to certain properties. With respect to this question, the two theories will be empirically investigated in the experiments in Part II from another important perspective of presupposition research, namely the insight into language acquisition, the second main research concern of my study.

4 The Acquisition of Presuppositions

The second main research concern of this work addresses the acquisition of presuppositions in both first and second languages (L1 and L2). In comparison to other related phenomena, such as scalar implicatures, on which the empirical research on acquisition has mainly been conducted after Noveck (2001), studies on presuppositions in first language acquisition can be traced back at least to the 1970s on factive verbs and the factivity of their complements (see e.g. Harris 1975; Macnamara et al. 1976; Johnson and Maratsos 1977; Hopmann and Maratsos 1978). As presuppositions and their triggers are quite numerous and diverse, this chapter cannot review all experiments on the acquisition of every trigger or trigger set of the last half century. Therefore, only some recent empirical studies on presuppositions in L1 acquisition that are relevant for the design of the experiments in Part II are reviewed in Section 4.1. In contrast, there are far less experiments investigating presuppositions in second language acquisition. Therefore, in Section 4.2, I will start with some experiments on related phenomena in L2 acquisition and then come to the few experiments on presuppositions. A summary of this chapter can be found in Table 3.

4.1 Presuppositions in First Language Acquisition

As a phenomenon at the semantic-pragmatic interface, how is the derivation of presuppositions developed during the L1 acquisition? Since when are children aware of presuppositions and their certain properties, which distinguish them from assertions and perhaps also from each other? Regarding those questions, most experiments investigate presupposition acquisition in L1 by comparing children at different development stages with adults – that is, monolingual children from different age groups are tested and their behaviors are compared with native adult controls. Among them, the development stages that receive the most attention in empirical studies can be roughly split into two groups: the preschool age from 3 to 5 or 6 years old, and the early school age from 6 to 9 years old. In this section, the former will be reviewed more detailed and later only briefly, as my experiment on L1 acquisition focuses on preschool children.[15]

With regard to the younger children at preschool age, most research focuses on whether they are aware of presuppositions as a whole class and whether they can interpret

[15]Thus, part of the review in this section is summarized later in Chapter 7, which has been published as Y. Chen et al. (2022).

them adult-like. However, the observations in the experiments are not consistent. On the one hand, some studies show that children over 3 years old are able to recognize and understand presuppositions. Syrett et al. (2010) for instance compared children of 3, 4 and 5 years of age with adult controls using a request-action task. In this task, two different objects sharing some properties, such as color or shape, were presented to the participants. Then, a puppet asked for something with the definite expression *the x one*. The participants' task was to decide whether they can accept the request by giving one of the two objects to the puppet based on the this description, or have to reject the request as there is no object that fits the description, contradicting the existence presupposition of *the*, or both objects fit the description, contradicting the uniqueness presupposition. Examples adopted from their control items are illustrated below in Table 2. Syrett et al. (2010) report that even 3-year-old children rejected requests with failed presuppositions with a high accuracy of about 90%, far above chance level[16], indicating that they understand the definite article *the* with its presuppositions in an adult-like way.

Table 2: Items adopted from Syrett et al. (2010, p. 13) with and without presupposition failures, slightly modified.

Adjective in request	Stimuli	Expected action
Red	Red poker chip and a white poker chip	Accept
Green	Purple yo-yo and yellow yo-yo	Reject due to the existence presupposition
Red	Pictures of a red square and a red circle	Reject due to the uniqueness presupposition

Similar conclusions are also made by Berger and Höhle (2012), who tested German children at 3 and 4 years of age with the presupposition trigger *'auch' (also)* and the exhaustivity marker *'nur' (only)* using a story-reward task. In their experiments, children first watched a story told by two experimenters: one (E1) introduces an animal that has two things to do, for instance a lion who should eat a banana and an apple, and the other one (E2) plays the role of this animal. After having disappeared with the two fruits for a while, E2 shows up again and the two experimenters have the following conversation:

(62) E1: Oh, there's the lion again. I'll ask him what he did: Lion, you've surely eaten the BANANA!
E2: Guess what? I've also eaten the APPLE/I've eaten the APPLE/I've only

[16]In their first experiment, this accuracy decreased in case of the critical items in which the relative gradable adjectives were involved. As gradable adjectives do not belong to the main topics of my study, this decrease will not be further discussed.

eaten the APPLE.

(Berger and Höhle 2012, pp. 374/398, slightly modified)

After that, participants had to decide whether the lion has done all his tasks and should be rewarded. Results show that both 3- and 4-year-old children almost always rewarded the animals after having heard answers with *also*, significantly more often than regarding answers without *also*, indicating that they can access and understand the presupposed meaning of this additive particle. Analogically, the reward proportion decreased significantly with respect to *only* (lower than 10% in both groups), showing that the children are able to interpret *only* with its exhaustive meaning. Berger and Höhle (2012) also notice that their observations differ from those of previous studies in Dutch or Japanese (see e.g. Bergsma 2006; Matsuoka et al. 2006). Instead of a language specific difference, they assume that the children's competent performance is due to the improvement in experiment design: firstly, the animals (or E2 in fact) disappear for a while so that the presupposition is neither confirmed nor denied, but can be assumed by the participants. Moreover, based on the story, the alternatives to the focus are also highly salient, which facilitate the interpretation. Therefore, Berger and Höhle (2012) argue that already at 3 years of age, children can access and understand the presuppositions triggered by *also* as properly as the exhaustiveness asserted by *only*.

On the other hand, there are also studies that illustrate 3-year-old children's difficulty with presuppositions, showing that young children's awareness of presuppositions is restricted, which is for instance demonstrated in the experiment conducted by Dudley et al. (2015). In their study, 3-year-old children's understanding of factivity is investigated, that is, if children can understand that only complements of some cognitive verbs can be considered presupposed or guaranteed, while those of others cannot. Choosing the factive verb *know* and the non-factive verb *think* as contrast, Dudley et al. (2015) tested whether children are aware of the different truth conditions or different levels of speaker commitment of their complement clauses in both affirmative and negated environments. Designing a hidden toy task, similar to the tasks used in Moore et al. (1989) and Moore and Davidge (1989), participants were told that a toy was hidden in one of the two boxes in front of them. After hearing sentences like (63), they were asked to decide whether the toy is in the red or in the blue box.

(63) a. no negation: Lambchop thinks/knows that it's in the red/blue box.

b. negation in matrix: Lambchop doesn't think/know that it's in the red/blue box.

c. embedded negation: Lambchop thinks/knows that it's not in the red/blue box.

(Dudley et al. 2015, Section 4.2, slightly modified)

Results show that children performed adult-alike on all three kinds of *think*-items. With respect to those employing *know*, their choices were similar only if the negation was not in the matrix clauses. In those sentences with negation in the matrix clause, only 40% of the children were able to recognize the projection of the presupposition and chose the correct box. These results show that although some 3-year-old children might be aware of the presuppositions triggered by factive verbs, this knowledge is restricted as more of them still struggle with projection through negation. However, note that in this experiment design, at the time of utterance, the presupposition is neither given nor can be naturally assumed as part of the Common Ground but needs to be accommodated. This requirement might cause the incapability among young children.

A similar restriction is also observed among older preschool children: Aravind and Hackl (2017) combined the story-reward design with the Truth-Value Judgment Task (TVJT, Crain and McKee 1985) and tested the acquisition of presuppositions triggered by the factive verbs *forget* and *remember* from another perspective: at-issueness. In their experiments, children between 4 and 6 years of age were compared with adults in order to investigate whether they are sensitive to the Not-At-Issueness Constraint, mentioned in (55) above and repeated in (64) below.

(64) Not-At-Issueness Constraint:
Presuppositions cannot be used to directly target the Question Under Discussion.
(Aravind and Hackl 2017, p. 51)

Their study consists of two experiments. In both experiments, the items contain a story, a comprehension question that also implicitly highlights the Question Under Discussion (QUD), and a sentence uttered by a puppet that needs to be judged as true or false. The crucial difference between these two experiments is how the implicit QUD is targeted. Examples from both experiments are adopted below:

(65) Material used in the first experiment:
Story: Billy was supposed to help Farmer Mary around her farm. He helped by feeding the chickens and milking the cow. The next day, Farmer Mary asked, "How did you help?" He said, "I milked the cow, but I'm not sure what else I did!" He didn't tell her about feeding the chickens!
Comprehension: How many stickers will Billy get? (One)
Puppet's sentence: Billy forgot that he fed the chickens. (True)
(Aravind and Hackl 2017, p. 53)

(66) Material used in the second experiment:
Story: Today, Billy and Jane were supposed to help Farmer Mary around her farm. Billy fed the chickens, but Jane was sleeping instead. The next day, Farmer Mary asked them, "Did you feed the chickens?". But both of them said, "It was so long

ago, we're not sure anymore!"
Comprehension: Who should get a sticker? (Billy)
Puppet's sentence: Billy forgot that he fed the chickens. (True)
(Aravind and Hackl 2017, p. 55)

In the first experiment, as can be seen in (65), the QUD (implied by the comprehension question) is about how many jobs done by Billy he did or did not mention to Mary, which determines how many rewards he can get. This question is (partially) answered with the puppet's assertion. In this experiment, children and adults performed similarly with a high accuracy regarding the ratings of the puppet's sentences. In the second experiment, however, the implied QUD of (66) is about who fed the chickens. In this case, it is the presupposition of the puppet's sentence that contains relevant information and answers the QUD. With this change, differences were observed between both groups: while adults' rating accuracy of the puppet's sentences remained at a high level, children's accuracy decreased considerably and was below chance level.

Combining these observations, Aravind and Hackl (2017) propose that on the one hand, children can understand presuppositions and are sensitive to the Not-At-Issueness Constraint of presuppositions. On the other hand, regarding the divergence in the second experiment, Aravind and Hackl argue that adults can assume a more appropriate QUD as a repair tool once there is incongruence between question and answer, while children at preschool age are not able to use this tool. Still, note that the QUDs in their experiments are only implied and the design with two kinds of judgment tasks can be confusing for preschool children. Moreover, the test conditions vary between groups: while the stories were presented auditorily with visual support to the children, adults received only a written version of the experiment without any visual input.

Overall, the experiments described above illustrate that when a suitable task with proper contexts is used, preschool children older than 3 years of age are able to understand presuppositions and their requirements, although with certain restrictions or incapability.

Besides the acquisition or awareness of presuppositions as a whole, another question that has been frequently investigated from the perspective of language acquisition is the difference within the category of presuppositions. That is, how are the different sub-classes of presuppositions or presupposition triggers, if any, acquired among children, and when are they aware of these distinctions. Interestingly, according to my research, previous studies that have investigated the heterogeneity of presuppositions or their triggers mainly test older children at early school age. Yatsushiro (2008a) for instance investigated the acquisition of lexical and implicated presuppositions (for more about this distinction, see e.g. Sauerland 2008 and the review in Subsection 2.2.2) in German among 6-, 7-, 8- and 9-year-old children. Two triggers were used: the first one was *'jede-' (every)*, with

the existence presupposition that is assumed to be lexical as well as the anti-uniqueness presupposition that is assumed to be implicated, as it is triggered via the maxim *Maximize Presupposition* (see 45a and 45b, and Heim 1991; Sauerland 2008). The second trigger was *'beid-' (both)*, as its duality presupposition is considered lexical. In her experiments, a sentence-picture match task was used. Children and adult controls were asked to decide whether a sentence like (67) can suitably describe the given pictures with no, one, or three target objects.

(67) a. Every uncle of mine is also sitting on a chair.
(In picture: there is no uncle, the existence presupposition is violated.)

b. Every mother of mine is sitting on a chair here.
(In picture: there is only one mother, the anti-uniqueness presupposition is violated.)

c. Both stuffed animals are on the shelf.
(In picture: there are three stuffed animals, the duality presupposition is violated.)

(Yatsushiro 2008a, Section 3.3, summarized and slightly modified)

Results show that once the lexical presuppositions were violated, the utterance was rejected in over 90% of all cases, even by children at 6 years of age. However, when it came to the implicated presuppositions, even 9-year-old children rejected them in only 65% of all cases (6-year-olds in only 34%), whereas adults rejected them much more robustly (90%). This difference supports the idea that some presuppositions – in this case, the lexical ones – are acquired at the latest by 6 years of age, which is in line with the observations with regard to children of preschool age in Syrett et al. (2010). More importantly, the varying sensitivity towards implicated presuppositions between the four children groups illustrates that the different kinds or sub-classes of presuppositions might be acquired at different ages: the more the pragmatic process involves, the more uncertainty or tolerance among children can be observed.

Another study that is concerned with a prediction related to trigger classification is Bill et al. (2016). With a focus on acquisition, they investigate the assumption that some presuppositions (mainly these of soft triggers) are similar to scalar implicatures (see Chemla 2008; Romoli 2014 and the review in Section 2.2.2 and 3.1). With preschool children (between 4 and 5 years old) in one group and children at early school age (7 years old) in the other, Bill et al. (2016) assessed their interpretation of presuppositions triggered by *win*, scalar implicatures triggered by *some* and indirect scalar implicatures triggered by *not all*, and compared them with adult controls. Adopting the covered box paradigm (Y. T. Huang et al. 2013), the visual pictures in their test were consistent only with the literal meaning, that is, contradicting the presupposition or implicature, so that

the choice of the covered, invisible picture can indicate that the participant derives the presupposition or implicature. A set of examples with a description of the visible pictures is presented below:

(68) a. Presupposition: The bear didn't win the race.
(Visible picture: The bear didn't participate in the race.)

b. Scalar implicature: Some of the lions got balloons.
(Visible picture: All of the lions got balloons.)

c. Indirect scalar implicature: Not all of the rabbits brought balls.
(Visible picture: None of the rabbits brought balls.)

(Bill et al. 2016, p. 63, summarized and slightly modified)

Results show that among both children and adults, sentences with presuppositions and implicatures are interpreted differently: while both groups of children chose the covered pictures more often with presuppositions than with implicatures, it was exactly the opposite among adults. This contrast supports the idea that presuppositions are acquired differently from implicatures and are more easily accessible for children. Additionally, Bill et al. (2016) explain children's preference for the covered picture, that is, their derivation of presuppositions, by assuming that they do not have an adult-like ability to accommodate. Another relevant observation in this experiment is that under the presupposition condition, the older children chose the covered picture significantly less often than the young children, indicating a crucial stage of presupposition acquisition before age 7.

Combining the findings of these two experiments, one can conclude that acquisition has to be a crucial aspect of the heterogeneity research on presuppositions, especially in the context of trigger classification. Considering that the crucial stage of acquisition appears to be before age 7, that acquisition can be delayed in cases of presuppositions that are assumed to be more pragmatic, and that the preschool children are able to understand presuppositions in general, it is possible and reasonable to investigate their awareness of the soft-hard split, or in other words: whether and in how far are preschool children between 4 and 6 years of age aware of the difference between trigger types?

4.2 Presuppositions in Second Language Acquisition

In comparison to acquisition in the first language, there are far less empirical studies on presuppositions in second language (L2) acquisition. In fact, pragmatics have received attention in L2 acquisition much later than traditional grammatical forms: according to the review in Taguchi (2011), attention has been paid to L2-pragmatics or pragmatic abilities in L2 only after the occurrence of communicative competence models with a focus on the acquisition of the social use of a foreign language in the 1990s (see e.g. Bachman

1990; Bachman and Palmer 1996). Although the amount of studies on pragmatics in L2 has increased in the last two decades, most of them still focus rather on teaching and examining strategies of foreign languages (for more, see reviews in e.g. Jung 2002; Taguchi 2011; Taghizadeh 2017; Culpeper et al. 2018). As for the pragmatic phenomena, based on the review in Taghizadeh (2017) and my literature research, there are more studies on implicatures (including both conversational and conventional) in L2 acquisition than on presuppositions. Although studies on first language acquisition seriously challenge the assumption that some presuppositions are implicatures (see the review in the last section), I will still briefly review two empirical studies on implicature comprehension in L2. Besides the potential similarity, the most crucial reasons are that these two studies offer some relevant observations that might be extended to pragmatic skills in L2 in general, and that there are very few empirical studies on presuppositions from the perspective of L2 acquisition.

The first one is conducted by Taguchi (2013) and investigates conventional implicatures. Taguchi (2013) compared both accuracy and reaction time of Japanese learners of English with that of English native speakers. The comprehension of several kinds of conventional implicature, including indirect request and indirect refusal, were tested in the experiments by means of a multiple choice task. Examples of the items are adopted and presented below.

(69) a. Requests

Tom: Oh, hi Sally. How are you?

Sally: So, I heard the boss just gave you a nice raise. Do you mind if I ask you how much you got this year?

Question: What is Sally telling Tom?

a) She wants to know how much raise Tom got.

b) She is very pleased with her boss.

c) She wants to know if Tom's life is good.

d) She doesn't care that Tom got a raise.

b. Refusals

Susan: I'm having a party this Saturday, and it should be fun. I hope you can come.

Dave: Oh, Susan, I already have plans on Saturday.

Question: Which of the following is correct?

a) Dave doesn't think Susan's party is exciting.

b) Dave is going to attend Susan's party.

c) Dave doesn't know his plans for Saturday.

d) Dave can't come to Susan's party.

(Taguchi 2013, Section 3.3.2)

Among all findings, the most interesting one for my study is that the Japanese learners of English understood indirect refusal easier and faster than indirect request. Taguchi (2013) explains this observation by the influence from L1, that is, it is common in both Japanese and English to refuse something indirectly by giving an explanation or excuse. But on the other hand, there is no such similar form of indirect request with *Do you mind if...?* in Japanese, which leads to a slowdown and higher cognitive cost in comprehension. In short, based on her experiments, Taguchi (2013) proposes that for nonnative speakers, linguistic forms and meanings that have counterparts in their mother tongue can be comprehended faster and more precisely in the foreign language, while there might be more difficulties or uncertainty with forms that that exist only in their L2.

Another experiment I want to shortly outline here is concerned with conversational, scalar implicatures and has been conducted in the context of my master's thesis; the paper based on it has been published, see Y. Chen (2019). In this experiment, native German speakers and Chinese learners of German were asked to judge the felicity of German sentences with scalar terms like *some* (the weak term in the scale <all, some>), *it can be* (the weak term in the scale <must, it can be>), and numerals. The critical items are logically correct but pragmatically odd if the scalar implicature is enriched, see for instance the target sentence in (70). In this design, a lower rating indicates that participants interpret the sentence pragmatically and derive the implicature. On the other hand, a higher rating is preferred if the literal meaning is more accessible for the participant.

(70) Context: From his journey, Mr. Lange has brought his sisters silk handkerchiefs and his brothers ties.
Target sentence: Some Mr. Lange's siblings have received souvenir.
(Y. Chen 2019, p. 126)

Results show that nonnative speakers rated sentences like (70) with typical scalar terms like *some* and *it can be* significantly higher than German native speakers, but there was no difference between the two groups with regard to numerals. Based on this observation, the exact meaning of numerals is assumed to be semantic, which is in line with the assumption based on comparisons between children and adults in Y. T. Huang et al. (2013). More importantly, these results show that in comparison to native speakers, nonnative speakers might have a different sensibility regarding pragmatic infelicity: analogical to children's pragmatic tolerance observed by Katsos and Bishop (2011), this difference possibly illustrates nonnative speakers' pragmatic tolerance, which might be a relevant aspect for investigating phenomena at the semantics-pragmatics interface, too.

As mentioned before, in comparison to implicatures, there are far less studies on presuppositions in L2 acquisition, although empirical research with a comparison between native and nonnative speakers can be traced back to Carrell (1984). In her study, Carrell compared factive predicates like *be thoughtful* with implicative predicates like *remember*, two sub-classes which are firstly defined by P. Kiparsky and C. Kiparsky (1970) and Karttunen (1971a). According to Carrell, both types of predicates can trigger presupposed and implied meanings, but exactly reversed: the implied meaning triggered by implicative predicates is presupposed by the factive, and vice versa. Among both type of predicates, half of them were semantically positive, such as *be thoughtful* and *remember*, and the other half were semantically negative, such as *be thoughtless* and *forget*. Two example sets of her materials are presented below in (71), together with the four types of inferences that she examined.

(71) a. John remembered to let the dog out.
The dog is out. (Implication, true)
The dog is in. (Implication, false)
The dog is supposed to be out. (Presupposition, true)
The dog is supposed to be in. (Presupposition, false)

b. It was thoughtless of John to let the dog out.
The dog is out. (Presupposition, true)
The dog is in. (Presupposition, false)
The dog is supposed to be out. (Implication, false)
The dog is supposed to be in. (Implication, true)

(Carrell 1984, p. 9)

The participants in her experiment consisted of three groups: 17 native speakers for control, 30 advanced English learners (average TOEFL scores 585), and 13 high-intermediate learners (average TOEFL scores 500). However, the mother tongues of the nonnative speakers were not homogeneous: their first languages included Chinese, Korean, Japanese, and more. During the test, the sentences were orally read by the experimenter and the participants had to rate the truth value of the inferences only based on the utterance.

Results show that although both groups of nonnative speakers were able to understand the presuppositions and implications in general, in comparison to the English native speakers, their accuracy decreased with respect to all inference types, and as expected, the group of high-intermediate learners had the highest error rate. This observation illustrates that inferences related to pragmatic skills like presuppositions need to be acquired in L2, and nonnative speakers at advanced and high-intermediate levels are still in this

process. Moreover, all three groups comprehended implications better than presuppositions, and semantically positive words better than negative ones. This tendency with a significant interaction was more obvious in nonnative speakers than in native speakers: in general, sentences with presuppositions triggered by semantic negative expressions caused an error rate over 50% among both groups of nonnative speakers. Carrell interprets this interaction as evidence for the distinction between presupposed and implicated inferences (see e.g. Clark 1977; Carrell 1978) and highlights the importance of native-nonnative comparison for research on pragmatic phenomena and their acquisition.

After Carrell (1984), there had hardly been any research on presuppositions in L2 acquisition, until more recently, Al-Zubeiry (2020) compared Saudi English speakers with British English speakers. Analyzing 30 newspaper opinion articles in English, Al-Zubeiry observes that triggers considered lexical (such as factive verbs or change of state verbs) and existential (such as definite expressions) were used more often than structural triggers (like clefts or wh-questions) in both groups. As for the comparison between the two groups, although there were some differences in frequency, these were statistically not significant. Thus, he concludes that the use of presupposition triggers in English is similar in Saudi and British speakers. However, according to Al-Zubeiry, this study mainly contributes to foreign language teaching strategies, for instance for writing and speaking skills – in other words, we may conclude that there is a lack of investigations on presuppositions from the perspective of L2 acquisition since Carrell (1984).

To summarize, in Chapter 4, several empirical studies on presupposition acquisition in both L1 and L2 are reviewed in Section 4.1 and 4.2, respectively, as their test design or main findings are important, informative, or interesting for my experiments in Part II. In Table 3, the relevant studies with their test methods, main findings, and remarks that are relevant for my experiments are listed.

Regarding L1 acquisition, some experiments show that preschool children are already aware of presuppositions and their not-at-issueness, while others show that their awareness is still restricted (see e.g. Syrett et al. 2010; Berger and Höhle 2012; Dudley et al. 2015; Aravind and Hackl 2017). Thus, these observations indicate that preschool age should also be a critical stage for investigating presuppositions from the perspective of L1 acquisition. However, focusing on differences between different categories of presuppositions, previous studies have mainly investigated children at early school age, showing that there may be some differences in the acquisition of lexical and implicated presuppositions on the one hand, and between presuppositions of soft triggers and scalar implicatures on the other (see e.g. Yatsushiro 2008a; Bill et al. 2016). An interesting question is then whether we can observe similar tendencies at the preschool stage: with children who can understand presuppositions but still have difficulties dealing with certain properties, we may be able

to investigate and compare the two main assumptions addressing the classification issue (see Chapter 3).

In contrast to L1 acquisition, there are only very few experiments on presuppositions from the perspective of L2 acquisition, although studies on related pragmatic features, such as Taguchi (2013) and Y. Chen (2019), show that this perspective can be both interesting and informative. Therefore, in my study, I will also take insights from L2 acquisition into account. However, in order to investigate presuppositions and their heterogeneity by means of L2 acquisition, there is one critical question that needs to be assessed firstly: how universal are presuppositions? In other words, are they and their heterogeneity stable across different languages, and if so, to what extent? The universality of presuppositions is the third main research concern of my study, and thus, some relevant theoretical assumptions and empirical cross-linguistic studies will be reviewed in the next chapter.

Table 3: Summary of empirical studies on presuppositions in first and second language acquisition.

Studies	Task	Triggers	Main findings	Remarks
L1 acquisition: preschool age				
Syrett et al. (2010)	request-action task	*the*	3-year-old children are aware of both existence and uniqueness presuppositions and their failures.	- Potential influence from the word *one*? - Observations are in line with Yatsushiro (2008a).
Berger and Höhle (2012)	story-reward task	*also*	3-year-old children can understand presuppositions triggered by *also*.	- Observations differ from previous studies in Dutch and Japanese. - Good performance may be due to improved design, such as the assumable state of the presupposition and the highly salient alternative to the focus.
Dudley et al. (2015)	hidden toy task	*know*	Although 3-year-old children might be aware of presuppositions, they still struggle with projection through negation.	- Weakness in design: the content of presuppositions is not part of Common Ground at the time of utterance, accommodation is required.

Table 3 – *Continued from previous page*

Studies	Task	Triggers	Main findings	Remarks
Aravind and Hackl (2017)	story-reward task with TVJT	*remember, forget*	Children between 4 and 6 years of age are sensitive to the Not-At-Issueness Constraint of presuppositions, but not able to use repair tools such as conforming a new QUD like adults.	- Tasks are complex. - QUDs are only implied. - Test procedure differs: audio and visual stimuli for children, written items for adults.
L1 acquisition: early school age				
Yatsushiro (2008a)	sentence-picture match	*every, both*	Children are capable of understanding lexical presuppositions already at 6 years of age, while implicated presuppositions cause difficulties even for some 9-year-old children.	- Observations are in line with Syrett et al. (2010) - Acquisition should be a relevant perspective for the heterogeneity and classification issue of presuppositions.
Bill et al. (2016)	covered box paradigm	*win*	Presuppositions and implicatures are acquired differently. In comparison to children at 4/5 years of age, 7-year-old children interpret presuppositions more adult-like.	- The assumption that some presuppositions are triggered SI-like is seriously challenged. - The stage before 7 years of age is presumably crucial for presupposition acquisition.
L2 acquisition				
Taguchi (2013)	multiple choice	conventional implicature	Japanese learners of English can comprehend indirect refusal better and faster than indirect request with *Do you mind if...?*	- L1 can affect the acquisition of pragmatics in L2. - Linguistic forms and meanings that have counterparts in L1 can be understood easier and faster in L2.

Table 3 – *Continued from previous page*

Studies	Task	Triggers	Main findings	Remarks
Y. Chen (2019)	felicity judgment task	scalar implicature	Chinese learners of German are more tolerant than native speakers towards logically correct but pragmatically infelicitous scalar terms, but not towards numerals.	- Nonnative speakers are more tolerant towards pragmatic violations than native speakers. - Nonnative speakers' pragmatic tolerance might be a relevant aspect for investigating phenomena at the semantics-pragmatics interface.
Carrell (1984)	TVJT	factive and implicative predicates	In comparison to implication and semantic positive words, presuppositions and semantic negative words can cause more comprehension difficulty, especially for nonnative speakers.	- Presuppositions need to be acquired in L2 - Nonnative speakers at advanced or high-intermediate levels are presumably a crucial group for investigations on pragmatics in L2.

5 Cross-linguistic Studies on Presuppositions

The last main research concern of my study addresses the universality of presuppositions, that is, whether presuppositions and their properties (including heterogeneity) are stable across languages. However, in comparison to other linguistic areas like syntax, there are less cross-linguistic studies or even typology research on pragmatic phenomena. This might due to the reason that it is generally assumed that the pragmatic enrichment mainly arises because of universal principles based on human rationality: the Gricean/neo-Gricean account (Horn 1984; Grice 1989; Levinson 2000; Horn 2004) for instance takes the Cooperative Principle (CP) as fundamental and formulates it as "a rough general principle which participants will be expected (ceteris paribus) to observe" (Grice 1989, p. 26[17]). As for the post-Gricean account, Relevance Theory (see e.g. Sperber and Wilson 1996; Wilson and Sperber 2004, among others) claims for instance a universal tendency by emphasizing *human cognition* in their First Principle of Relevance (also known as the Cognitive Principle of Relevance):

(72) Human cognition tends to be geared to the maximization of relevance.
(Sperber and Wilson 1996, p. 260 and Wilson and Sperber 2004, p. 610)

As both accounts hardly predict anything about the universality of presuppositions in particular on the one hand, and there are only few experiments that compare presuppositions between languages on the other, in this chapter, my review will start with a brief summary of theories and predictions on universal pragmatics that may directly or indirectly be related to presuppositions in 5.1, followed by some empirical studies on pragmatic phenomena which could be inspiring in 5.2. The experiments investigating presuppositions with language comparisons are outlined in 5.3.

5.1 Theoretical Predictions Related to Presuppositions

Let us start with the Gricean and neo-Gricean account: although the four maxims of CP and their sub-maxims are disputed and modified in neo-Gricean theories (see e.g. Horn 1984; Levinson 2000, and the discussion about conversational implicatures below in 5.2), the CP itself and its generality are widely accepted and adopted – at least there has been no explicit discussion or prediction on language specific CP in these classical (neo-)Gricean

[17]Note that the work *Logic and conversation* was firstly published in 1967 in William James Lectures, Harvard University.

works. Among them, the most relevant implicit clue to presuppositions is given by Brown and Levinson (1987). Brown and Levinson (1987) assume a universal politeness based on the assumption of the universality of face and rationality. They consider face to be "something that is emotionally invested, and that can be lost, maintained, or enhanced" (p.61). Among the indirect ways for doing face threaten actions, presuppositions are mentioned as a strategy that can give clues for a pragmatic interpretation containing implicatures, namely by violating the Gricean maxim of Relation/Relevance: "Be relevant" (Grice 1989, p. 27). Brown and Levinson (1987, p. 217) give the following example:

(73) I washed the car *again* today.
$\rightsquigarrow$ I had washed the car before.

According to Brown and Levinson's (1987) interpretation, the presupposition seemingly avoids the maxim of Relation/Relevance, which therefore gives the hearer a clue that the sentence should be interpreted in a broader context with events that have happened before: assuming that relevance can be achieved by imagining that the speaker and the hearer have agreed to share the work of car washing before, and they are counting now how many times each of them did it. This sentence and presupposition should then be interpreted with a critical implicature, and the face threaten action is done indirectly.

Additionally, Brown and Levinson (1987) also consider presuppositions as manipulation strategies for positive politeness. Note however that in this sense, they use the loose, wide and pragmatic definition of presupposition, that is, the speaker "presumes that it is mutually taken for granted" (p.122). With presuppositions in this sense, speakers can claim common point of view, attitudes, knowledge, and thus achieve positive politeness. Some examples given in the text (Brown and Levinson 1987, pp. 122–124) are summarized below:

(74) Presuppose knowledge of hearer's wants and attitudes
   a. Wouldn't you like a drink?
   b. Don't you think it's marvellous!?
      $\rightsquigarrow$ The hearer wants a drink/thinks it is marvellous.

(75) Presuppose hearer's values are the same as speaker's values
   a. tall man
      $\rightsquigarrow$ The speaker and the hearer share the same understanding of *tall*.

(76) Presuppose hearer's knowledge
   a. Well I was watching *High Life* last night and...
   b. Oh, *this* is lovely! (on walking into a house)
      $\rightsquigarrow$ The hearer knows what *High Life* and *this* refer to.

Although Brown and Levinson (1987) do not say anything specific about whether presuppositions are universal or language specific, considering that they assume the universality of politeness, face, rationality and Gricean principles, it should be rational to conclude that they also imply the universality of presuppositions. However, interestingly, politeness is one of the areas where the pragmatic universality is extremely seriously challenged. Some experiments on this topic will be reviewed in the next section, together with studies that challenge other maxims of CP.

As for Relevance Theory, Sperber and Wilson (1996) adopt the Gricean notion of Relation/Relevance, but doubt other claims including the need of CP. They propose that relevance is the center of pragmatic enrichment as "the expectations of relevance raised by an utterance are precise and predictable enough to guide the hearer toward the speaker's meaning" (Wilson and Sperber 2004, p. 607). In their work, the differences between presuppositions and assertions are only shortly discussed (see e.g. Sperber and Wilson 1996, Section 4.5). Based on the distinction between foreground and background implications, presuppositions are assumed to be analytically implied by background implications. For illustration, they use a sentence pair, adopted below in (77), as an example: according to Strawson (1964), if there is no King of France, only (77a) will be considered a sentence with no truth value because of the presupposition failure, while (77b) can be simply false. Using Relevance Theory, Sperber and Wilson argue that this difference is due to the reason that only the background implications of sentence (77a) imply that the King of France did something, while the background implications of (77b), on the contrary, imply that something happened with regard to the exhibition.

(77) a. The King of France visited the EXHIBITION.

b. The exhibition was visited by the King of FRANCE.

(Sperber and Wilson 1996, p. 214)

This idea can be seen as being related to the at-issueness theories, reviewed before in Section 3.2. But the most relevant prediction here is that nothing language- or culture-specific is mentioned or involved in this analysis. Combining this analysis with their universality assumption of relevance, expressed in (72), we are allowed to assume that Relevance Theory also supports the idea of the universality of presuppositions.

However, there are two things that need to be noted. First, addressing the distinction between focus and presuppositions, Sperber and Wilson (1996, p. 216) claim that their effects can be explained with the "interaction between syntax, stress assignment and the principle of relevance", of which the former two can vary between languages and/or so-

cial groups[18]. Second, Sperber and Wilson already notice the non-universality of another pragmatic phenomenon: speech acts. In a short note, they mention that the form of commitments can be particular and culturally defined, therefore, speech acts can be different in non-Western languages and they "have no doubt that a cross-cultural study of such speech acts would confirm their cultural specificity and institutional nature" (p.290).

In addition to those implicitly addressed in the Gricean accounts and Relevance Theory, a third indirect inference can be concluded from the nondetachability of presuppositions, discussed in Simons (2001) (see also the review in Section 2.2.2). As shown in (39b), repeated below again in (78), Simons notices that presuppositions are less dependent on a special linguistic form, but can be triggered if the expression is replaced by words or phrases with a similar meaning.

(78) Jane didn't stop laughing/didn't quit laughing/didn't cease laughing.
$\rightsquigarrow$ Jane laughed before. (Simons 2001, p. 435, sightly modified)

Extending this nondetachability, we might be allowed to expect that if a trigger has a counterpart in other languages with a similar semantic meaning and perhaps also similar syntactic features, then the generation of their presuppositions and the projection behavior should also be stable across languages. An example of the implicit (and unverified) adoption of this assumption could be the translations of Levinson's (1983) textbook *Pragmatics*, in which his list of presupposition triggers is also translated to various languages, such as German and Italian.

In short, theoretically, we are allowed to assume that both Gricean accounts and Relevance Theory more or less predict the universality of pragmatic strategies and hence also imply the universality of presuppositions, although in Relevancy Theory, some language or cultural specific pragmatic cases are also noticed. This universality might also be supported by the nondetachability of presuppositions, proposed by Simons (2001), and already implicitly assumed in the translation of Levinson's presupposition trigger list.

5.2 Empirical Cross-linguistic Studies on Related Pragmatic Phenomena

Among all implicit predictions that might be related to presuppositions, the most direct one is perhaps from the theory of universal politeness (Brown and Levinson 1987). Therefore, the review of experiments on related phenomena starts with cross-linguistic studies

[18]Sperber and Wilson 1996, p. 213:"The fact that contrastive stress is a natural highlighting device need not prevent it from being more costly to use in some circumstances than in others, just as pointing, another natural highlighting device, may have greater social costs attached to it in some circumstances than in others. This suggests, then, an interesting approach to cross-linguistic variation in stress patterns[...]."

on politeness, followed by research on other principles and features, including some maxims of CP proposed by the Gricean accounts, and the difference in speech acts, which has already been noticed by Relevance Theory. Note that due to the great amount of research works, this section is only a brief summary of some selected studies.

Since the 1990s, politeness has already been a favored topic in comparisons between Western and non-Western languages with the focus on whether the concept and behavior of politeness is the same around the world. Among others, challenge started with regard to the Japanese language: Ide (1989) and Matsumoto (1989) firstly doubt the universality of Brown and Levinson's (1987) politeness because of two observations: the use of honorifics in Japanese and the greeting phrase *yoroshiku onegaishimasu* that can be imposing and polite at the same time. As both observations can hardly be explained with Brown and Levinson's (1987) theory, which is based mainly on Western society, *'discernment'* is proposed as a critical concept for politeness in Japanese, which should at least be an alternative or complementary to Brown and Levinson's (1987) framework. Experimentally, Ohashi (2008) for instance investigates the o-rei (Japanese way of thanking) during gift-season at the end of the year. Analyzing data collected from telephone conversations, Ohashi claims that most theories on politeness or speech acts about 'thanking' in English indeed cannot explain or correctly predict o-rei conversations, which can contain several turns of 'acknowledging debt/benefit–denigrating credit' in Japanese.

Besides Japanese, another language to which the traditional pragmatic maxims are often applied and with which the limitations of pragmatic strategies of politeness are illustrated, is Chinese. Research on this can be traced back at least to Gu (1990) and Mao (1994): while Gu (1990) implies that Brown and Levinson (1987) might fail to analyze the normative function of politeness in Chinese society, Mao's (1994) work directly criticizes Brown and Levinson's (1987) universal 'face' model as empirically inadequate. Analyzing the Chinese concepts of 'face' (*mianzi* and *lian*), he claims that "[t]he social and moral connotations evidenced in *mianzi* and *lian* lie well beyond the semantic boundary marked by negative and positive face" (Mao 1994, p. 483), which was proposed by Brown and Levinson. Instead, he suggests a relative construction of 'face' with cultural variations. In terms of experiments, Chinese also belongs to the languages in which the so-called East-West divide has most frequently been investigated and observed. Various studies on linguistic features that are associated with politeness have been conducted, with focuses ranging from typical hearer-speaker events like refusal (Liao and Bresnahan 1996), offering and acceptance of gifts (Zhu et al. 2000), request (Rue and G. Q. Zhang 2008) and compliment responses (Tang and G. Q. Zhang 2009) to multi-party conversation (Xia and Lan 2019), and from informal conversation on telephone (H. Sun 2004) to communications via posts on social media like Twitter and Weibo (Li et al. 2020). Due to the great amount

of studies and the rather indirect relation between these features and presuppositions, the details are not outlined here. For a detailed summary of these studies, see reviews in e.g. R. Chen (2010) and Jia and Yang (2021). In short, these empirical studies on Chinese – not limited to those mentioned above – support language and cultural specific pragmatic strategies, propose (im)politeness based on Chinese society (see e.g. also Ran and Zhao 2018; X. Chen 2019), and doubt a universal model.

Besides the East-West divide, inadequacy of Brown and Levinson's universal politeness model is also observed in African languages. Nwoye (1992) and De Kadt (1998) for instance examine the validity of Brown and Levinson's conception of 'face' in Igbo (Nigeria) and Zulu (Southern Africa) societies, respectively. They observe that with a collectivism nature, individual face is less relevant in African communities, in contrast to the Anglo society. Take the Igbo society for example: Nwoye (1992) observes that there is a clear distinction between 'group face' and 'individual face', of which the former is ranked higher than the latter, and the Igbo politeness also clearly differs from the politeness supposed in Brown and Levinson (1987) (for a more detailed review of these and further studies, see Ameka and Terkourafi 2019).

In sum, comparing politeness between languages, the empirical studies show that although there could be similarities between pragmatic strategies across languages, the differences are dominant and the cultural and social impacts need to be considered.

As for other pragmatic principles from the Gricean accounts, it is also observed that not all of them can be adopted with regard to non-Western languages. In fact, already half a century ago, Keenan (1976) observes that Malagasy speakers have different conversational maxims from those used by Western speakers. Keenan therefore doubts the universality of these Gricean and neo-Gricean principles and argues that (p.67) "Gricean analysis retains usefulness but within the perspective of a comparative typology in which locally valid systems may differ strikingly in what is marked and unmarked. An ethnographic base and ethnological comparison are required." More recently, Ameka and Terkourafi (2019) argue that the Gricean CP in general results from Western, especially Anglo-interactional influence and therefore the maxims cannot be granted to be cultural-neutral or universal. Take the maxim of Manner and its four sub-maxims, see (79), for example: Ameka and Terkourafi propose that they cannot be obeyed in African cultures, as the African society requires exactly the opposite. Based on observations in languages like Akan (Obeng 1994, 1999, 2003), Yuroba (Ayodele 2016), several West African languages including Akan, Ewe, Ga, Hausa, Moore and more (Ameka 2006) and the use of English in Malawi (Kondowe et al. 2014), Ameka and Terkourafi (2019) summarize the "African ways of words", adopted below in (80), as opposed to the Gricean maxims:

(79) Gricean maxim of Manner: Be perspicuous.
Sub-maxim 1: Avoid obscurity of expression.
Sub-maxim 2: Avoid ambiguity.
Sub-maxim 3: Be brief. (Avoid unnecessary prolixity.)
Sub-maxim 4: Be orderly.
(Grice 1989, p. 27)

(80) Ameka and Terkourafi's "African ways of words"
Maxim of Manner: Be opaque.
Sub-maxim 1: Be obscure [Use veiled speech].
Sub-maxim 2: Be ambiguous.
Sub-maxim 3: Be long-winded.
Sub-maxim 4: Be circuitous.
(Ameka and Terkourafi 2019, Section 3, slightly modified)

Besides the maxim of Manner, the universality of the maxim of Quantity is also controversial: being one of the most relevant phenomena in the Gricean and neo-Gricean account, conversational implicatures might also be language specific. According to Grice (1989), the triggering of conversational implicatures is mainly explained with the maxim of Quantity, first sub-maxim: "Make your contribution as informative as is required (for the current purposes of the exchange)" (Grice 1989, p. 26). In the neo-Gricean account, the analysis of conversational implicatures is improved by means of different modified maxims, such as the hearer-based Q-principle and speaker-based I-principle proposed by Horn (1984), and the Q-, I- and M-heuristics proposed by Levinson (2000). Despite the difference in details, note that both Gricean and neo-Gricean accounts predict a distinction between two kinds of conversational implicatures: the generalized conversational implicature (GCI)[19], whose triggering is considered rather default, robust, and less dependent on context, and the particularized conversational implicature (PCI), which is more context-dependent (for more about GCI and PCI, see e.g. Grice 1989; Horn 2004). For example, expressing a sentence like (81) below can lead to both GCIs and PCIs: while the GCIs are stable in most contexts, the PCIs vary. Assuming that the CP and the maxims of Quantity are general, such computation of conversational implicatures and their distinctions are predicted to be stable across languages, too.

(81) Some balloons are yellow.
a. Context 1: the addressee wants a not-yellow balloon and the speaker knows it.

[19]Note that in this sense, scalar implicatures, which are assumed to be similar to some presuppositions (see the review and discussion in Section 2.2.2 and 3.1), build a subset of GCI.

GCI: Not all balloons are yellow.

PCI: There are balloons for you.

b. Context 2: the speaker and addressee are preparing a party for a friend who loves yellow.

GCI: Not all balloons are yellow.

PCI: The friend should be happy with the party.

Empirically, both stability and variation are observed. On the one hand, Thalmann et al. (2021) illustrate that the distinction of GCIs and PCIs can be stable across languages: in their experiments, implicatures are tested with indirect lies in German and Chinese. That is, while the sentences are literally true, the GCIs and PCIs triggered by these sentences are false. German and Chinese participants were asked to sort these sentences with false implicatures into three categories: lie, deception, or truth. Results show that in both languages, false GCIs were interpreted rather verbal and classified as lies, while false PCIs were considered more compatible with deceptions, the non-verbal option. This parallel rating can be seen as an evidence for some universal behaviors of conversational implicatures. On the other hand, Stateva et al. (2019) observe in their experiments that the GCI trigger *some* in English is understood differently from its counterparts in French, Slovenian and German. In their first experiment, participants first read contexts where a proportion of subjects did something, and then they were asked to rate sentences with *few, some, half, almost* and *most*. While in other non-English languages, the word *some* was rated as acceptable when 3% to 50% of subjects did something, in English, this was extended to a range from 3% to 80%. With a picture-choice test in their second experiment, they further illustrate that the difference between *some* in English and its counterparts in other languages is not due to semantic or lexical factors, but rather due to different pragmatic understanding mechanism.

In addition to (im)politeness, CP and their maxims (including conversational implicatures), the last area of research that I would like to shortly mention here is that on speech acts[20]. Cross-linguistic investigations on speech acts can be traced back at least to Finnegan (1969), who compares performative utterances in Limba with English, proposing that Austin's (1962) model also fits the speech acts in Limba. However, afterwards, more investigations show that speech acts are not universal but also dependent on language-cultural specific features. Besides studies on speech acts associated with po-

[20] According to the review in Ameka and Terkourafi (2019), among cross-linguistic pragmatic studies, the fourth favored theme, i.e. after (im)politeness, conversational strategies and speech acts, is conversation analysis, as several conversation analytic studies across languages and cultures have been executed recently by the Max Planck Institute in Nijmegen with both quantitative and qualitative analyses of phenomena in conversation, such as other-initiated repair (Dingemanse et al. 2014; Dingemanse and Enfield 2015) and turn-taking (J. Holler et al. 2016). As their connection with presuppositions is more indirect, the studies will not be further discussed here. For more, see e.g. Rossi et al. (2020).

liteness, especially those concerned with the East-West divide already mentioned above, differences have also been observed among Western languages. Wierzbicka (1985) for instance compares Polish with English. She observes that for certain illocutions like advise or offer, interrogative and conditional forms are used more often in English, while imperatives are preferred in Polish. More recently, Ameka (2017) also shows that in Ewe, an African language, imperative is routinely used for some requests for help, which reduces the possibility of an unconditional universal theory of speech acts.

In sum, empirical studies (not limited to those reviewed in this section) with their observations on cross-linguistic variation illustrate that pragmatic principles and features are not doubtlessly universal or general like traditional theories assume, as they are more or less based on the Anglo-American communities. In fact, language and cultural specific pragmatic strategies are crucial for research, and always need to be taken into account. Therefore, the cross-linguistic stability of presuppositions cannot just be assumed but experimental investigation is needed, especially between Western and non-Western languages.

5.3 Empirical Cross-linguistic Studies on Presuppositions

Although all empirical studies reviewed so far are doubtlessly relevant for universal pragmatics, they offer hardly any direct evidence for or against the universality of presuppositions. In fact, in terms of empirical investigations, the universality of presuppositions has not been of interest for a long time. According to my research, there is only a very small amount of cross-linguistic studies on presuppositions, and I will review all three studies I found in this section.[21]

Let us start with Amaral and Cummins (2015), who compare the projection behavior of presuppositions in Spanish with that in English, which has been examined in their previous study Cummins et al. (2012). In their experiments, two kinds of triggers were tested: resolution triggers and lexical triggers in Zeevat's (1992) sense. For the resolution triggers, they merge the first and third groups of Zeevat's classification, as they all share the anaphoric property and refer to entities that have already been introduced in the context, such as *too*, *also* and *again* in English. The other group contains lexical triggers such as *stop* and *continue*, which encode their preconditions logically and lexically (for more about this distinction, see the review in Section 3.3). In total, 8 triggers of these

[21]Note that besides these three studies reviewed here, there are also some studies that investigate presuppositions triggered by gestures (see e.g. Schlenker 2019; Tieu et al. 2019) or in sign languages (see e.g. Schlenker 2021a). According to the summary in Schlenker (2021b), in comparison to presuppositions in spoken languages, these studies, to some extent, prove the cross-linguistic validity of certain triggering processes or rules. As my studies only target spoken languages, while gestures and sign languages do not belong to the main issues of my research, these experiments will not be further discussed.

two types were tested with question-answer pairs in their experiments, based on the assumption that presuppositions are backgrounded and therefore can hardly be denied directly. Using an acceptability judgment task, Amaral and Cummins (2015) prepared four kinds of answers based on a 2×2 design: with 'yes' or 'no', and with or without denying the presupposition triggered by the question. An example from their test is adopted below:

(82) Question
¿Sigue siendo Victoria la directora del departamento?
'Does Victoria continue to be the director of the department?'
$\rightsquigarrow$ Victoria was the director before.

a. Answer 1: with 'yes', without denial of presupposition
Sí, Victoria sigue siendo la directora del departamento.
'Yes, Victoria continues to be the director of the department.'

b. Answer 2: with 'yes', with denial of presupposition
Sí, aunque antes Victoria no era la directora.
'Yes, although Victoria was not the director before.'

c. Answer 3: with 'no', without denial of presupposition
No, Victoria ya no es la directora.
'No, Victoria isn't the director anymore.'

d. Answer 4: with 'no', with denial of presupposition
No, porque Victoria no era la directora.
'No, because Victoria was not the director before.'

(Amaral and Cummins 2015, p. 165, sightly modified)

According to Amaral and Cummins (2015), results show that in both Spanish and English, answer types 1 and 3 were rated better than 2 and 4, indicating that in both languages, presuppositions are backgrounded and can hardly be addressed directly and naturally. Additionally, with regard to lexical triggers like *continue*, participants clearly favored type 4 with 'no' over 2 with 'yes', while there was no such tendency with respect to resolution triggers. This distinction between triggers was also stable across languages. They therefore propose that "the differences between presupposition triggers are rooted in general logical or cognitive principles, rather than arising as a language-specific feature of English (p.158)", supporting the universality of presuppositions. Note however, that a potential weakness of these studies is that, as argued by Snider (2017a,b), the probability or possibility for a direct denial might not illustrate the backgoundedness but rather the anaphoric availability. With this in mind, the similarity in both studies nevertheless

supports that at least the difference in anaphoric properties of presuppositions is stable in Spanish and English.

Another relevant experimental work that should be reviewed here is Schwarz et al. (2020). In their study, presuppositions of emotive and cognitive factive verbs in Italian are compared with that in English, which has been investigated before in their previous study Djärv et al. (2018). Similar to Amaral and Cummins (2015), in their studies, the presuppositions were also tested with question-answer pairs and an acceptability judgment. Using the same design, items with four triggers, namely *be happy* and *appreciate* for emotive and *be aware* and *realize* for cognitive factive verbs, were tested first in English in Djärv et al. (2018) and then translated into Italian in Schwarz et al. (2020). A set of examples in English is presented below:

(83) Question
Is Anna aware/happy that Ryan is coming to the wedding?
Does Anna realize/appreciate that Ryan is coming to the wedding?
$\rightsquigarrow$ Ryan is coming to the wedding.
Answer
a. 'yes' with denial: Yes, although he isn't.
b. 'no' with denial: No, because he isn't.
(Djärv et al. 2018, p. 373, sightly modified)

In Djärv et al. (2018), it is observed that once the presupposition was denied, the difference between emotive and cognitive verbs was visible mainly in 'yes' responses: the 'yes' response was clearly more acceptable for emotive verbs than for cognitive ones, but there was no clear difference between them in 'no'-answers. In Schwarz et al. (2020), results show that the ratings for the 'yes'-answers in Italian are similar to those in English: 'yes'-answers were clearly more acceptable for emotive triggers than for cognitive triggers in both languages. However, Schwarz et al. (2020) also observe a significant interaction with regard to the 'no'-answers: in comparison to their English counterparts, 'no'-answers were judged significantly less appropriate for cognitive triggers in Italian, illustrating some potential differences of presuppositions between languages.

Last but not least, recently, Reins et al. (2021) compared the use of presuppositions between English and Russian speakers by using presuppositions as one of four possibilities for indirect lying. In contrast to direct lies, in indirect lies, the false information is not given directly in the assertion, but implied by pragmatic enrichment such as presuppositions, GCIs and PCIs, or by non-verbal actions. One example of an indirect lie by means of these four variations that was used in the test is cited below:

(84) Last year, Emma's mother passed away, leaving her a large amount of money behind. With that money, Emma bought herself a nice and big house. Surprisingly, shortly after, Emma also won the lottery, receiving another large amount of money. Today Emma is working in her new house's front yard when her acquaintance Clara, who Emma had not talked to in a while, incidentally walks by on the footway. Emma does not want Clara to know that she bought her house with money she inherited from her mother. That is why, when Clara starts a conversation and asks, 'How did you afford a place like this?' ⋯

a. Presupposition: Emma answers: 'I admit that I bought the house with money I won in the lottery!'

b. GCI: Emma answers: 'Well, last year I won the lottery and bought this place!'

c. PCI: Emma answers: 'Well, winning the lottery allows people to buy the houses they've always dreamt of!'

d. Emma takes out her smartphone and shows Clara a photograph of herself posing with a giant lottery check in front of her newly purchased house.

Clara comes to believe that Emma won the lottery and that she used the money from the lottery win to buy her new house.

(Reins et al. 2021, Section 4.3.1, sightly modified)

After reading the items, participants were asked to evaluate them from three perspectives: how morally correct is the behavior, whether it is considered misleading, and whether it is considered lying.

There are several interesting observations in this experiment, but here, I only mention those involving presuppositions. Addressing presuppositions, Reins et al. (2021) only mention with respect to the rating as lies that there was hardly any difference between the groups regarding presuppositions. With respect to the morality ratings, the figure shows that the mean judgments on presuppositions are also very similar between languages, although there is a significant difference of the grand mean of all four answers between groups. With respect to the misleading ratings, a group difference was also observed and all four deception types were considered less misleading by the Russian participants than by the English participants. This difference can also be observed for presuppositions in the figure, however, without any information about its significance. In short, Reins et al.'s (2021) experiment shows that in the context of indirect lies, presuppositions are considered stable across languages with respect to the lying rating and perhaps also with respect to the morality judgment, but there might be a (at least numerical) difference of presuppositions' misleading effect between Russian and English.

To conclude, in Chapter 5, theoretical and empirical research on the universality of presuppositions has been reviewed. While the pragmatic principles are traditionally

considered general, rooted in human cognition, and cultural-neutral (see Section 5.1), experiments on politeness, conversational strategies and speech acts, among others, reveal the necessity and importance of social and language specific perspectives, as shown in Section 5.2. As for presuppositions, slightly different from what had been assumed so far, cross-linguistic experiments, although only a few, illustrate not only similarities but also some potential differences in their behavior (see Section 5.3). Therefore, the universality of their projection and the stability of their projection strength should not be taken for granted either, but needs to be experimentally investigated. Additionally, note that the present experiments on presuppositions compare mainly European languages. Considering the observations on related phenomena, especially the East-West divide, a comparison of presuppositions in Western and non-Western languages is reasonable and necessary.

6 Summary of Part I

In this part, starting with the development of definitions of presuppositions, the theoretical and empirical background of presuppositions has been reviewed with a focus on three main research concerns: first, the heterogeneity of presuppositions, which leads to the classification issue; second, the acquisition issue, including both first and foreign languages and third, the universality of presuppositions. This chapter is a short summary of the relevant research reviewed in this chapter, with a list of the main research questions in (85) for my experiments in Part II.

Regarding the development from semantic towards pragmatic definitions of presuppositions, the most classical definition from the purely semantic perspective, as reviewed in Section 2.1, assumes that they are preconditions for the truth value of an utterance in a trivalent logic system: if the presupposition is satisfied, the utterance can be *true* or *false*; if this is not the case, then the utterance is considered *undefined* (see e.g. Frege 1892; Strawson 1950, 1952; Katz 1973; Heim 1991, among others). However, this uniform, strong definition fails to explain several observations, such as the filtering problem and the heterogeneous projection behaviors as shown in (21) and (22). Moreover, this assumption in fact treats presuppositions as a special subset of entailments, whose truth is guaranteed (and not only guaranteed) by the truth of the utterance. This assumption is challenged by recent studies like Sudo (2012), Klinedinst (2016), and Schlenker (2021b). In order to resolve these difficulties, pragmatic aspects have received increasing attention. As outlined in Section 2.2, these theories can be roughly split into two accounts: the first one assumes a conventional or at least a consistent and uniform triggering of presuppositions and proposes diverse cancellation or repair strategies to explain the various projection behaviors (see Subsection 2.2.1). The second account claims that some – but not all – presuppositions can be considered conversational implicatures, or are at least triggered like conversational implicatures, while the others are rather lexical, semantic or conventional. Thus, this account proposes that presuppositions are not triggered uniformly and aims at a split or a classification of presuppositions (see Subsection 2.2.2).

Addressing the first main research concern, namely the heterogeneity of presuppositions and the classification issue, among all assumptions and criteria, two of them are more relevant for my experiments and thus have been reviewed more detailed in Chapter 3. First, the soft-hard distinction of triggers (see e.g. Simons 2001; Abusch 2002, 2010 and the review in Section 3.1) has been chosen from the second pragmatic account. Generally

speaking, this distinction is based primarily on a two-way split of projection strength or of the difficulty of local readings. The second one, which is also one of the most recent development of first pragmatic account, proposes the correlation between projection and not-at-issueness and assumes a gradient of projection strength (e.g. Simons et al. 2010; Xue and Onea 2011; Tonhauser et al. 2018). This assumption also targets a broader context including also non-presupposed projective meaning, such as non-restrictive relative clauses, as reviewed in Section 3.2. Combining these two assumptions, their relation will be empirically investigated in Part II, that is, whether the soft-hard split can be reflected in terms of at-issueness. Moreover, focusing on the two-way classification, the possibility of measuring the *softness* of some triggers by using some typical soft and hard triggers as anchors is also investigated. Note that the term *soft* and *hard triggers* are used in a wider sense in this work than originally proposed by Abusch (2002, 2010): instead of a pragmatic-semantic triggering process, *soft* in this work refers to triggers whose presuppositions are considered, depending on theories, weaker or less often projective, more defeasible or suspendible, more sensible to certain contextual influence or more easily to lose projectivity and receive a local reading. *Hard triggers*, on the other hand, are expressions whose presuppositions are in general stable, highly projective and reluctant to these effects.

As for the second main research concern, acquisition in both first and second languages (L1 and L2) are taken into account. Comparing the current empirical studies on presuppositions in L1 acquisition, it can be seen that first, children over 3 years of age are able to understand presuppositions and their properties, if the experiment design is appropriate, although with limitations or difficulties. Second, the experiments on the heterogeneity or the difference between presuppositions (or their triggers) have mainly been conducted with children of early school age, although studies prove that the development before 7 years of age is of great interest, too. Considering these observations and the potential weaknesses or possible improvements in experiment design (see Section 4.1 and the summary in Table 3), it is possible and reasonable to investigate preschool children's awareness of the difference between presuppositions by means of at-issueness. As for the acquisition of presuppositions in L2, as reviewed in Section 4.2, this area has hardly been experimentally studied so far, and there are only very few experiments on presuppositions with native-nonnative comparisons. Still, studies on related phenomena show that on the one hand, pragmatic inferences need to be learned in L2 and this process can be influenced by similarities and differences between L1 and L2 (Taguchi 2013). On the other hand, nonnative speakers might be more tolerant towards pragmatic infelicities with respect to scalar implicatures (Y. Chen 2019). Thus, it is expected that such nonnative speakers'

pragmatic tolerance can also be observed with respect to presuppositions and reveal more about the mechanism behind the heterogeneity of presuppositions.

Last but not least, the third main research concern is about the universality of presuppositions. While pragmatic principles and features are traditionally and theoretically considered general or rooted in human cognition, as discussed in Section 5.1, empirical studies seriously challenge this assumption. Differences have been observed especially in research areas like (im)politeness, conversational strategies and speech acts, and mainly by means of comparisons between Western and non-Western languages or societies, as shown for example in the East-West divide (see Section 5.2). As for presuppositions, there are less cross-linguistic experiments that have investigated their universality – to the best of my knowledge, only three. In these three studies, both similarities and potential differences between languages have been observed, although in all of them, the comparison takes place only between European languages (see Section 5.3). Thus, the stability of presuppositions between languages, especially between Western and non-Western languages, cannot be simply assumed, but needs to be empirically investigated.

In sum, based on the theoretical and empirical background outlined in this part, this study mainly addresses the above-mentioned three main research concerns with the following five main research questions in the experimental part:

(85) a. The classification issue:
Research question 1: verifying the classification: can the soft-hard dichotomy, a classification based on the projective strength, also be reflected in terms of the ability of different triggers to express at-issue content?
Research question 2: applying the classification: choosing certain typical presuppositions or triggers as markers for certain properties, can they function like a thermometer that measures and sorts other triggers into certain subclasses?

b. The acquisition issue:
Research question 3: acquisition in first languages: are children aware of the difference between trigger types? Do children and adults respond differently once the presuppositions are forced to be at-issue? Are the different trigger types acquired at different ages?
Research question 4: acquisition in second languages: are nonnative speakers as sensitive as native speakers once the presuppositions of different triggers (or trigger types) are forced to be at-issue? If not, how can the difference be explained?

c. The universality issue:
Research question 5: cross-linguistic comparison of presuppositions: can a

presupposition be triggered steadily across languages, as long as an adequate and equivalent translation of its trigger can be found? Are their properties like projection behavior or information status independent from linguistic and cultural influence?

Part II

Empirical Studies

7 Exp.I: Trigger Classification and At-issueness in L1 Acquisition[22]

7.1 Introduction

In communication, there is a vast amount of information that we take for granted. Some of these are requirements triggered by specific linguistic expressions or constructions – we call these language-based assumptions presuppositions. By uttering (86), for example, the speaker presupposes that Peter has a dog – it is a requirement of the expression *Peter's dog* that Peter's dog exists.

(86) Peter's dog is ill.

The occurrence of presupposed content is linked to specific expressions or structures, so-called presupposition triggers. Whereas in (86), the presupposition is triggered by the possessive NP, in (87a), the cleft construction is responsible for the presupposition. Here, the speaker presupposes that someone smoked. If this precondition is not fulfilled, the utterance is odd or, depending on the theory, cannot be interpreted truth-conditionally. The same holds for example (87b): By uttering the sentence containing the change of state verb *stop*, the speaker presupposes that the duck used to smoke.

(87) a. It was the duck who smoked.
   $\rightsquigarrow$ Someone smoked.
   b. The duck stopped smoking.
   $\rightsquigarrow$ The duck used to smoke.

Other words and structures that trigger presuppositions include factive verbs like *discover*, implicative verbs like *manage*, or additive particles like *too*, among others (see, for instance, Levinson 1983c for a more extensive overview). Such pre-required information is abundant in natural language and abundant is the body of literature on the topic. Traditionally, presuppositions are defined as those inferences of an utterance that survive embedding, i.e., that project out of the scope of an entailment-canceling operator. This

[22]This chapter is joint work with Maik Thalmann and Mailin Antomo and has been published as a journal article in *Journal of Pragmatics* with the title *Presupposition triggers and (not-)at-issueness: Insights from language acquisition into the soft-hard distinction* (Y. Chen et al. 2022). This is the accepted version of the paper, for the published version, see https://doi.org/10.1016/j.pragma.2022.06.014.

can be seen in the following example, where the entailed but not the presupposed content of (87b) is negated or questioned:

(88) a. ¬S: The duck did not stop smoking.

b. S?: Did the duck stop smoking?

(88a) - (88b) $\rightsquigarrow$ The duck used to smoke. Presupposition of (87b)

(88a) - (88b) $\not\rightsquigarrow$ The duck stopped smoking Assertion of (87b)

Besides negated environments, the persistence of presuppositions has been observed in conditionalization (if S) and modalization (perhaps S), the so-called family of sentences (Chierchia and McConnell-Ginet 2000), while entailed meaning is modified.

Although all presuppositions have the ability to project out of the scope of an entailment-canceling operator, numerous studies conclude that in certain contexts not all presupposition triggers behave the same way. Compare (88) with (89), adopted from Simons (2001): Whereas in (88) it is presupposed that the subject used to smoke, (89) shows that the same presupposition does not hold, at least not globally (since it is interpreted under the question operator), given the speaker and the addressee meet for the first time:

(89) I notice that you keep chewing on your pencil. Have you recently stopped smoking? (Simons 2001, p. 432)

$\not\rightsquigarrow$ The addressee used to smoke.

However, such a local reading, which we will have more to say about later, is not equally available for all presuppositions. Abusch, for instance, argues that presuppositions are differently projective and illustrates the difference with the two sentences below (Abusch 2002, p. 3):

(90) a. After the first meeting, John will either *continue* missing meetings, or *continue* attending them.

b. # After the first meeting, John will either miss the second meeting *too*, or attend the second meeting *too*.

Without any context or embedding, both *continue* and *too* can trigger the presupposition that John missed or attended the first meeting. However, although both presuppositions arise, the one triggered by *too* in (90b) causes infelicity, while the presupposition of *continue* in (90a) receives a local reading and does not project globally. That is, an utterance like (90a) neither requires that John missed the first meeting nor does it require that he attended the first meeting. While explanations for this heterogeneous pattern differ, various researchers, such as Abusch (2002), seek to account for it by subclassifying

presupposition triggers into two classes, hard and soft, such that soft triggers like *continue* are more easily read locally, while hard triggers like *too* resist non-global interpretations.

The observation that presuppositions show divisive behavior with respect to such local readings will be at the heart of our experiments, which will seek to answer the following questions: Is there a need for a soft-hard dichotomy? And if so, are soft and hard triggers acquired at the same time?

To investigate these questions, we exploit another property of presuppositions: their reluctance to express at-issue content, as described by Simons et al. (2010) and Tonhauser et al. (2018). The basic observation is that presupposed content cannot be used to felicitously target the Question Under Discussion (QUD). Thus, more precisely, the main issue of our study is to investigate if the soft-hard dichotomy is reflected in terms of the ability of different triggers to express at-issue content and, furthermore, whether preschool children are aware of these differences. For this aim, we will first review the relevant theoretical concepts on trigger classifications, especially concerning the soft-hard dichotomy in Section 7.2.1 and (not-)at-issueness in Section 7.2.2, and then give a brief overview of the main experiments on presupposition acquisition in Section 7.2.3, before we present our study in Section 7.3. The main findings and their theoretical implications are discussed in the general discussion.

7.2 Research Background

7.2.1 Soft and Hard Triggers

Consider again examples (90a) and (90b) in the introduction, which illustrate the heterogeneous behavior of presuppositions: The presupposition of *too* projects globally in (90b) while that kind of projection is absent in (90a). Another classical context that also illustrates this difference between triggers is provided by epistemically deficient contexts, first discussed by Simons (2001), where the speaker's ignorance regarding presupposed meaning components is made explicit. Consider example (91) below:

(91) a. **Soft trigger: *win***
I don't know whether the duck participated in a race, but if she won, she is probably drunk now.

b. **Hard trigger: *too***
I don't know whether anybody else was ill, # but if the duck was ill too, she needed rest.

In (91a), global accommodation is blocked because it would result in an incoherent discourse. That is, assuming a common ground in which the presupposition holds

globally clashes with the speaker's assertion, as shown in (92a). However, without accommodation, there is a similar crash: *win* presupposes participation, but the speaker specifically mentions that she does not know whether the presupposition holds. Since global accommodation is unavailable, the commonly assumed rescue strategy to explain the acceptability of cases like (91a) is a local variant of global accommodation, as in (92b).

(92) a. # **The duck participated in a race**, and I don't know whether the duck participated in a race, but if she won, she is probably drunk now.

b. I don't know whether the duck participated in a race, but if **she participated** and won, she is probably drunk now.

The difference is clear: While global accommodation causes infelicity, under the local reading, the presupposed material is interpreted in the scope of the conditional, which avoids a clash with the assertion. However, this rescue strategy seems unavailable for *too* in (91b), which hence projects globally, and leads to infelicity. What causes this difference?

As hinted at in the introduction, numerous analyses have been proposed so that we cannot review them all in detail. In the following, we will focus on the main ones, as well as those that were crucial for the selection of the triggers we included in our experiment.

The first proposal in this regard goes back to at least Kripke (2009) (presented orally in 1990) and Zeevat (1992), who separate triggers based on their similarity between their presuppositions and anaphora. In updated terminology, hard triggers, e.g., *too* and *again*, are strictly anaphoric with respect to previously established material and are, one may reason, less easily interpreted locally. Soft triggers, by contrast, come with weaker anaphoricity requirements, and thus appear to receive a local reading more easily.

More recently, approaches have championed other aspects than anaphoricity, though these newer accounts are similar in spirit: They argue for a qualitative difference between the triggering mechanisms of soft and hard presuppositions. This in turn is used to account for the difference in local readings.

One family of analyses proposes that some presuppositions are triggered pragmatically or conversationally like implicatures (see e.g., Simons 2001; Abusch 2002; Simons 2007; Chemla 2008; Abusch 2010; Romoli 2014). Abusch (2002), for instance, differentiates between soft and hard triggers by suggesting that hard triggers like *too* lead to semantic presuppositions and that soft triggers like *continue* cause pragmatic presuppositions. While the former arise semantically and are therefore not defeasible, the presuppositions of soft triggers involve a set of alternatives, much like with scalar implicatures.

In (90a), for example, the utterance of *continue* activates the alternative *stop*. According to Abusch, the speaker optionally and pragmatically presupposes that one of the alternatives in the set is true, with the best candidate being the disjunction of *continue*

x and *stop* x. As the two alternatives have an overlap, which is also the presupposition of the sentence, namely *used to do* x, its truth can be considered pragmatically taken for granted. However, as this presupposition comes about pragmatically, Abusch argues, it is weaker than the presuppositions of hard triggers. Using the explicit ignorance context as shown in (91), Abusch (2010) classifies triggers like *it*-clefts, *too* and *again* as hard triggers, whereas triggers like *discover* and *win* are viewed as soft ones.

Slightly different to the semantic-pragmatic opposition above, another perspective was proposed recently to explain why the presuppositions of some triggers do not easily give rise to local readings: entailment (see e.g. Sudo 2012; Klinedinst 2016; Zehr and Schwarz 2016, 2018).[23] They suggest that a local reading is only available if the presupposed content is part of the entailments of the assertion. On this account, local accommodation is generally available for soft triggers because the assertions always entail their presuppositions. With hard triggers, the presuppositions are not obligatorily entailed and thus the availability of local readings is limited. Since, in monotonic contexts, entailments and presuppositions are not distinguishable, the two can only be separated when embedded in a non-monotonic environment like *exactly one*. Sudo (2012) illustrates the difference between the presuppositions of *herself* (namely the gender information) and *stop*, summarized in Zehr et al. (2016, p. 321), as follows:

(93) Exactly one student stopped using Mac.

  a. Exactly one student used Mac and does not use Mac now.

  b. Exactly one student does not use Mac now.

(94) Exactly one student criticized herself (…namely Mary).

  a. Exactly one student is a female who was self-critical.

  b. Exactly one student was self-critical.

For *stop*, the preferred truth conditions include the presupposition, (93a), suggesting that the presupposition is entailed. With *herself*, on the other hand, the reading without the presupposition, (94b), aligns better with speaker intuitions, as (94a) can also hold in situations where there are more than one self-criticizing students, as long as there is only one female among them, contrary to the intuited interpretation. It is assumed that soft triggers generally pattern with *stop* and hard triggers with *herself*.

However, the idea of capturing the varying potential for local accommodation via a subclassification of presupposition triggers is not uncontroversial. Abrusán (2016), for

[23]Note, however, while Klinedinst (2016) and Zehr and Schwarz (2016, 2018) assume that there are some correlations between the entailment relation and the soft-hard split, Sudo (2012) claims that the non-entailed status of some presuppositions can be changed by semantic operators like negation and therefore should not be understood as an indicator of the soft-hard split. This controversial issue is not the main research point of this paper and will not be discussed further.

instance, claims that a trigger type distinction is not needed as all presuppositions are "fundamentally the same type" (p.168). That is, presuppositions are contents which are "not necessarily about the event time of the predicate" (p.178) and therefore do not belong to the main point of the utterance (this idea is related to the at-issueness approach, which will be discussed in Section 7.2.2). In other words, all triggers share the same triggering mechanism, and their presuppositions hold as long as they are, defaultly or contextually, not paid attention to. That is, presuppositions vary with respect to how easily they can be focused contextually to convey a (secondary) main point of the sentence: While the complements of triggers like *discover* can be easily focused in certain contexts and, in these cases, do not encode old information, the presuppositions of additive particles or *it*-clefts cannot easily be used this way because of their lexical/syntactic properties.

With this in mind, it is important to at least survey the experimental data in regard to the availability of local readings, which the above accounts at least indirectly predict to be restricted to a subset of triggers (though, as shown, the approach-internal reasons differ). While Spenader (2002) quantitatively supports that the difference in global accommodation may depend on trigger type, Jayez et al. (2015) observed that, even for a hard trigger like *too*, a local reading is still possible. Therefore, they argue that the different possibilities of local accommodation might be better explained with the interaction between context and trigger, rather than based on trigger type, which is more compatible with Abrusán (2016).

In line with Jayez et al., Zehr et al. (2017) also found that local readings are possible for *again* in *neither-nor* sentences, even though they are dispreferred. Additionally, Amaral and Cummins (2015) observe that the differences between triggers are rather gradient than categorical, and also argue for an interaction between lexical meaning and contextual information as the most salient predictor. On the other hand, Zehr et al. (2016) and Zehr and Schwarz (2018) add corroborating evidence for a trigger classification in their experiments. Using the *exactly-one*-test suggested by Sudo (2012) with the covered box design, they show that there is a clear split between triggers like *stop* and triggers like *also* and *again*. However, they also point out that one of the key questions that needs to be answered by the accounts supporting a trigger classification concerns language acquisition: If there are different types of presupposition triggers, then how are these types and their differences acquired by children?

In sum, the debate can be roughly summarized as follows: As a point of departure, we can observe that there are some contexts in which some presuppositions do not project globally but rather receive a local reading. This observation has led to a discussion about the underlying cause of this asymmetry. Although the details are controversial, it is generally accepted that there are properties that are not shared by all triggers, and

the local accommodation is more available for some of the triggers. These triggers are therefore considered 'soft', no matter which property is (assumed to be) responsible for this softness or whether it can be determined using lexical semantics alone. Note, however, that the aim of our experiment is not to investigate any particular approach or theory on the soft-hard distinction, but rather to contribute to this debate by exploiting another property of presuppositions, namely their reluctance to express at-issue content, as well as potential discrepancies when it comes to language acquisition.

7.2.2 Presuppositions and (Not-)At-Issueness

As we have seen, local accommodation yields an interpretation of presupposed meaning components within the scope of semantic operators. However, local accommodation, at least in the epistemically deficient contexts discussed above, is not available for the presuppositions of hard triggers. What we have left unanswered so far is why by means of local accommodation, (at least some of) the originally presupposed contents can become foregrounded and lose their projectivity.

A clue to answer this lies with non-restrictive relative clauses (NRRCs). Like expressives and nominal appositives (for more, see e.g., Potts 2005), these are also interpreted outside the scope of operators, despite not being presupposition triggers. As can be seen in (95), the content of a NRRC is interpreted outside the scope of negation, whereas the proposition of the main clause is negated:

(95) It is not the case that the duck, who was ill, likes cake.
$\rightsquigarrow$ The duck was ill. (content of the NRRC)
$\not\rightsquigarrow$ The duck likes cake. (content of the main clause)

Now, if projection is not a property uniquely associated with presupposed content, how can we explain their projective behavior? By appealing to at-issueness, i.e., taking projection to be a property resulting from the way information is structured in discourse. Though the details of at-issueness shall not be discussed here, an important feature of QUD-based approaches is captured in (96):

(96) **At-issueness** (slightly modified from Simons et al. 2010)
A proposition p is at-issue relative to the Question Under Discussion (QUD) Q iff $?p$ (*whether* p) is relevant to Q; where $?p$ is deemed relevant when it provides a partial or exhaustive answer to Q.[24]

[24]This definition is actually revised in Simons et al. (2010, p. 323) to rely on speaker intentionality in order to deal with more complex cases. As these will not feature here, the simpler version of at-issueness will suffice.

In this model, projection is a property of discourse components (von Stutterheim and Klein 1989; Roberts 1996): Semantic operators target only at-issue content and leave untouched not-at-issue components, which project on account of them not being relevant to the QUD. And while NRRCs are merely not-at-issue, presuppositions come with the added requirement that their content has to be old information relative to the common ground.[25,26]

Thus, what is important for us is that presuppositions cannot be used felicitously to answer the QUD *a priori*. This can be seen in the following examples from Antomo (2016, p. 40), where the content of the embedded clause answers the QUD. When embedded under a factive predicate such as *ignore*, which presupposes its complement, the utterance is odd. However, a non-presuppositional control such as *believe* is fine:[27]

(97) Q: Where is Homer?
A: # Marge ignores that he's at Moe's.
A′: Marge believes that he's at Moe's.

As a corollary of the above, Aravind and Hackl (2017), following early lecture notes by Irene Heim, propose the Not-At-Issueness constraint (which, of course, also applies to NRRCs):

(98) **Not-At-Issueness Constraint** (Aravind and Hackl 2017, p. 51)
Presuppositions cannot be used to directly target the Question Under Discussion.

The Not-At-Issueness Constraint holds also for NRRCs, even if they are not presupposed, as can be seen in (99). This has led to the conclusion that NRRCs are conventionally marked as not-at-issue (see, for instance, A. Holler 2005, Beaver 2012, and Antomo et al. 2021), which is why we use them as a baseline for not-at-issueness in our experiment.

(99) Q: Where is Jill?
A: # Jill, who is in Berlin, lost something on the train.

With this in hand, we can answer the question from above: Since local accommodation makes presuppositions interact with semantic operators, and since not-at-issue

[25]Note that we adopt the common ground view on presuppositions even if, as observed by amongst others Abbott (2000), presupposing constructions can convey new information. What is crucial is that presuppositions can (and often do so) express old information, whereas NRRCs are not adequate to contain aforementioned content.

[26]Another property that is claimed to be distinct for not-at-issue material is that they can only be rejected indirectly by using special discourse strategies like 'Hey, wait a minute!' (as originally described by Shanon 1976 and von Fintel 2008 for presuppositions; see Roberts et al. 2009, Xue and Onea 2011 and Antomo 2015 for more information und experimental investigations).

[27]Note, however, that the presupposed complement clause embedded by *ignore* violates the Common Ground constraint, since its proposition clearly is informative.

material is unaffected by such operators, local accommodation causes presupposed material to become at-issue at the cost of projection.

We find empirical support for this claim in two studies: Xue and Onea (2011) observe that highly projective (and thus not-at-issue) presuppositions are difficult to reject directly, which, in their view, is a measure for not-at-issueness (see also Tonhauser 2012). If we adopt the latter, their results confirm a correlation between projection and (not-)at-issueness.[28]

In the same vein, Tonhauser et al. (2018) show that not-at-issue content triggered by NRRCs, *be annoyed*, and possessive noun phrases is highly projective, while the presuppositions of triggers like *stop* and *discover* are less projective. However, contrary to what is typically assumed in the literature, they argue for a continuum rather than a discrete dichotomy between soft and hard presupposition triggers (as well as at-issueness) and base this on the amount of projective variability between the different lexical triggers.

Now we are in a position to present the main rationale for the experiment we will present later. Standardly, presuppositions make not-at-issue contributions, which are odd when they target the QUD directly, see (97). Secondly, presuppositions project, that is, they are impervious to the influence of semantic operators in the basic case, see (88). Since we also find this kind of projection for non-presupposed, not-at-issue material as well—see (99)—it stands to reason that it is not-at-issueness which causes projection. However, as we have seen, some presupposition triggers, namely those we called soft triggers, can be locally accommodated, a process which cancels their projective behavior, like (91a). In combination with the suggested link between at-issueness and non-projection, we hence assume that locally accommodated presuppositions are at-issue. In contrast, local accommodation is not easily available for hard presupposition triggers, i.e., they cannot be used in an at-issue way and thus remain projective, see (91b). Since most of the studies focused on the projection-side rather than the at-issueness side of local accommodation, we want to investigate if the soft-hard dichotomy (or a continuum) is reflected in the potential for at-issue contributions of presupposition triggers.

7.2.3 Presuppositions in Language Acquisition

For the investigation of the heterogeneity of presuppositions, another insightful perspective comes from language acquisition. Compared to implicatures, especially scalar ones, the acquisition of presuppositions is much less explored. This part gives a short overview of some relevant empirical research that focuses on presuppositions in language acquisition and serves as a background for our own experiment.

[28]Note that Snider (2017a) argues that direct deniability is not a suitable method to measure (not-)at-issueness and shows this diagnostic instead targets anaphoric availability. To avoid this problem, we will apply question-answer pairs in our study.

Investigating the acquisition of presuppositions in general, evidence from 3-year-old children is given by Dudley et al. (2015). This experiment employs a design that is characteristic for studying the acquisition of presuppositions and in the context of which we want to highlight a central problem. In this study, the authors investigate whether 3-year-old children are aware of the difference between *know* and *think* in terms of selecting a complement that is presupposed (*know*) versus one that is not (*think*). Children and adults were told that a toy was hidden in one of the two boxes in front of them. After hearing sentences like (100), they needed to decide whether the toy was in the red or the blue box.

(100) a. no negation: Lambchop thinks/knows that it's in the red/blue box.
b. negation in matrix: Lambchop doesn't think/know that it's in the red/blue box.
c. embedded negation: Lambchop thinks/knows that it's not in the red/blue box

While children and adults performed alike on all three kinds of *think*-items, children made adult-like choices only on *know*-trials with embedded or without negation. When *know* came with negation in the matrix clause, only 40% of the children recognized it and chose the correct box. These results show on the one hand that even some 3-year-old children are aware of the difference between *know* and *think* by recognizing that *know p* presupposes the truth of *p*. On the other hand, young children struggle with projection through negation. Thus, when investigating the acquisition of presuppositions, we have to keep in mind that projection out of embedding is acquired later. Furthermore, as Aravind and Hackl (2017) point out, the study conducted by Dudley et al. (2015) violates the Common Ground constraint. At the time of utterance, the presupposition of *know* was clearly not given as a part of the Common Ground but needed to be accommodated globally. Thus, besides presupposition triggers and projection, a third factor (accommodation) is included, weakening the explanatory force of the results.

Turning to local accommodation, Bill et al. (2016) compare the comprehension of presuppositions and scalar implicatures among both adults and children (split into two groups: 4 and 5 years old in the one group and 7 years old in the other one). The results show (i) that presuppositions, in contrast to implicatures, are easily accessible for both young children and adults, and (ii) that young children do not have an adult-like ability to accommodate locally.

Together, the studies show that when testing whether children have already acquired the presuppositional properties of a trigger, accommodation must not feature. If it does, the acquisition of accommodation is tested. On the other hand, if the presupposition is already established in the context, it is not clear whether children understand it because of the trigger itself or the linguistic context. How is this tension to be resolved?

Aravind and Hackl (2017) investigate the acquisition of presuppositions by using another property of presupposed content, which we discussed in the previous section, namely their inability to express at-issue content. By doing so, they also produce one of the rare studies that investigate the acquisition of at-issueness. In their study, Aravind and Hackl (2017) tested 4- to 6-year-old children as well as adults using a Truth-Value Judgment Task (Crain and McKee 1985) paradigm to investigate whether children are sensitive to the Not-At-Issueness Constraint – see (98). For this goal, they tested the complements of the factive verbs *forget* and *remember*. One example, adopted from Aravind and Hackl (2017, p. 55), where the implicit QUD is answered by a presupposition, is shown in (101). After answering a question about which protagonist should be rewarded, which tested both the comprehension of the story and highlighted the implicit QUD – who fed the chickens –, participants were asked to judge a sentence uttered by a puppet as either true or false.

(101) Today, Billy and Jane were supposed to help Farmer Mary around her farm. Billy fed the chickens, but Jane was sleeping instead. The next day, Farmer Mary asked them, "Did you feed the chickens?". But both of them said, "It was so long ago, we're not sure anymore!"

Comprehension: Who should get a sticker? (Billy)

Puppet's sentence: Billy forgot that he fed the chickens. (True)

They found children and adults to perform similarly when the contextual restrictions of presuppositions are met, whereas children's accuracy declines once this is no longer the case, that is, when the implicit QUD is answered via the presupposition in the puppet's sentence. In their experiments, adults performed at ceiling and, so Aravind and Hackl (2017) reason, assume a more appropriate QUD to repair the incongruence between question and answer, while children performed below chance level. From this, the authors conclude that children (much like adults) are sensitive to the Not-At-Issueness Constraint, but are unable to apply the same repair strategy (i.e., changing QUD) as the adults.

As an interim summary, we can conclude that being not-at-issue is an important necessary (although not sufficient) property that can be exploited to investigate various aspects of presuppositions. Aravind and Hackl (2017) exemplify how this property can be leveraged to study the acquisition of presuppositions; we will add to this by focusing on the soft-hard distinction.

All studies just discussed provide important data on the interface of language acquisition, presuppositions and at-issueness. However, there are a number of points our study attempts to improve upon. First, in previous work, participants were often asked to give binary judgments, and it is possible that some nuance between children and adults was

lost. From the vast literature on the acquisition of scalar implicatures, it is well known that children, much more so than adults, are affected by the study design, such as the available response categories. With binary response options, children do not judge implicature violations to be severe enough to reject the utterance on a binary rating scale – in stark contrast to adults (see Katsos and Bishop 2011 who introduce the notion of children's pragmatic tolerance for this purpose). Second, we include NRRCs as controls, which provides us a baseline for conventionally encoded not-at-issueness. Third, there are weak points in the design of previous investigations we would like to improve upon. In some studies, the presuppositions are not given in context in violation of the Common Ground constraint. In others, for instance in Aravind and Hackl (2017), where the Common Ground constraint is fulfilled, the QUD is implicit or indirect, which may also influence comprehension and response behavior. Therefore, in our experiment, we will present the presuppositions in a context with an explicit QUD. Last but not least, hard and soft triggers have been distinguished on the ground that their presuppositions have different projective behavior. We hypothesize that the presuppositions associated with hard and soft triggers can be distinguished on other grounds: their tendency to (be judged to) felicitously answer a QUD.

Taking stock, hard and soft triggers differ along with a number of important dimensions, and most importantly for us, in the ability for local accommodation. Further, there is a large amount of overlap between different theories regarding our experimental predictions. That is, while *again*, for instance, is typically seen as hard or highly projective, cognitive factive verbs like *discover* or achievement verbs like *win* are usually judged as soft (and thus predicted to be more readily interpreted locally). With this in mind, we will use the triggers below in our study, which we take to be typical instances of soft and hard triggers:

(102) a. **Soft Triggers**
entdecken ('discover'), *gewinnen* ('win'), *schaffen* ('manage to')

b. **Hard Triggers**
auch ('also', 'too'), *wieder* ('again'), *it*-clefts

Our experiment, presented in the next section, aims at answering the following questions:

(103) a. Heterogeneity of triggers: Is the soft-hard dichotomy correlated with the tendency of different triggers to express at-issue content? Is there varying sensitivity to the Not-At-Issueness Constraint between triggers (possibly beyond the soft-hard dichotomy)?

b. Acquisition of presuppositions: Do children and adults respond differently to violations of the Not-At-Issueness Constraint? Are soft and hard triggers acquired at different times?

7.3 Experiment

7.3.1 Design

To answer the research questions above, we devised a modified acceptability judgment task including both children and adult controls as participants. The offline rating experiment comprised a number of short video vignettes consisting of stories to adequately set up the contextual parameters before the target utterances were played back and a rating on a 5-point Likert scale was requested. The entire experiment was programmed using the open source software OnExp (version 1.3.1) and featured a $3 \times 2 \times 2$ design with the factors TRIGGER (NRRC vs. hard trigger vs. soft trigger; between-item and within-subject), ISSUENESS (NRRC/Presupposition at-issue vs. Assertion at-issue; within-item and within-subject), and AGE (children vs. adults; within-item and between-subject).

7.3.2 Participants

As Dudley et al. (2015) observed that only about half of the three-year-old children in their study understood *know* in a consistently adult-like way but Yatsushiro (2008) demonstrated that six-year-old children were able to reject utterances with violated lexical presuppositions over 90% of the time, we chose preschool children between four and six years of age as our critical group and compared them with adult controls. Therefore, the two participant groups of our test are: 23 4-to-6 year old Kindergarten children (16 female, mean age: 5.29 ± 0.78), and 33 adult controls (25 female, mean age: 25.72 ± 10.3). All were monolingual German native speakers from Göttingen and the surrounding area and were naïve as to the purpose of the experiment.[29] Children received a small gift for their participation, adults were not compensated.

7.3.3 Materials

One experimental session, about 30 to 35 minutes in length, involved at least 30 critical items, of which one half was such that the target utterance contained a trigger whose presuppositional content addressed the QUD and thus was at-issue and in violation of the Not-At-Issueness Constraint (see (96) and (98) above), and the other half was the baseline condition where no violation occurred in that the presupposition was supported by the

[29] Due to restrictions and protective measures undertaken in preschools as a consequence of COVID-19, testing more children was rendered impossible.

context and backgrounded information. Note that the speaker of the target utterance is thus always cooperative regarding the content (the QUD is answered in all items), though perhaps not with respect to how that content is expressed.

A further subdivision of the critical items instantiated a second factor: TRIGGER. Here, there were 10 items per soft and hard presupposition trigger as well as 10 NRRCs, where the latter served as a baseline for this factor by being conventionally not-at-issue but not presuppositional. Relative clauses were doubly marked as being non-restrictive: for one, the antecedent DP always referred to a unique and thus unambiguously identified single individual in the discourse domain (cf. Jacobsson 1994) and the relative clause itself contained *übrigens* ('by the way') which is only compatible with a non-restrictive interpretation (see A. Holler 2013, p. 276). As an anonymous reviewer pointed out, there is good reason to believe that *übrigens* is, by itself, a conventionally encoded marker for not-at-issueness. Therefore, the items involving NRRCs are doubly marked as not-at-issue. A list of the soft and hard triggers we used can be found in (102), repeated here with their frequency in (104) below.

(104) a. **Soft Triggers**

entdecken ('discover', 3 items), *gewinnen* ('win', 3 items), *schaffen* ('manage to', 4 items)

b. **Hard Triggers**

auch ('also', 'too', 3 items), *wieder* ('again', 3 items), *it*-clefts[30] (4 items)

All items were structured the same way. First, a short story featuring the protagonists panda, duck and frog – see Figure 1 – was told. All stories presented set-ups that were very familiar to even the youngest children and typically involved friends in Kindergarten engaged in some everyday activity.

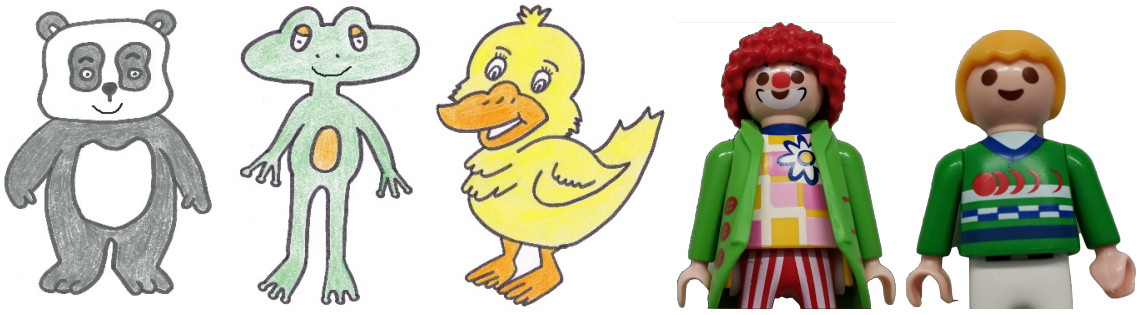


Figure 1: The protagonists of the stories – the panda, the frog, and the duck – as well as the clown and Peter (from left to right). Pictures used in the introductory section of the experiment.

[30]For the acquisition of *it*-clefts, see Aravind et al. (2016), Tieu and Kriz (2017), and Aravind et al. (2018).

Following that, a clown, introduced as being inattentive and watching the story together with the participants, would pose a question concerning the plot (the QUD) and Peter, a child described as sometimes having trouble with language, would utter the target sentence; see the two rightmost characters in Figure 1. The examples below, translated from German, showcase one example item per trigger in both the NRRC/Presupposition at-issue and the Assertion at-issue variant.

(105) **NRRC:**

The duck has a cough. She gets some cough syrup. Then she goes to bed in order to rest. The clown did not pay attention again and asks:

a. Assertion at-issue:
"Where is the duck now?"[31]

b. NRRC at-issue:
"What did the duck get?"

Little Peter responds, "The duck, who got some cough syrup by the way, is in bed now."

(106) **Soft trigger:** *gewinnen* ('win')

The panda, the duck, and the frog are having a drawing competition. All like the duck's drawing the most. The duck receives a crown as a prize. The clown did not pay attention again and asks:

a. Assertion at-issue:
"Did the duck do the best at the competition?"

b. Presupposition at-issue:
"Did the duck participate in the competition?"

Little Peter responds, "The duck won the competition."

(107) **Hard trigger:** *wieder* ('again')

The panda is a member of a soccer club. Recently, he has been playing very well and scored many goals in the past weeks. Yesterday there was a football game and he scored a goal. The clown did not pay attention again and asks:

a. Assertion at-issue:
"Did the panda score a goal yesterday?"

[31]The items further included the German modal particle *denn* in the question establishing the QUD, which exerts a number of restrictions on the interpretation on the embedding question. The first, discourse anaphoricity, signals that the answer to the question is to be found in the immediate utterance context. This ensures that the participants only consider the salient discourse moves or the preceding context and restrict their reliance on world knowledge (cf. Theiler 2021).

b. Presupposition at-issue:

"Did the panda score a goal for the first time yesterday?"

Little Peter responds, "The panda scored a goal again yesterday."

Because of the inherent differences between the triggers, in some items the QUD was expressed using wh-questions, while in others polar questions had to be used. Where possible, we tried to balance between the two question types within the different conditions of the TRIGGER factor. We will return to this issue in the results section.

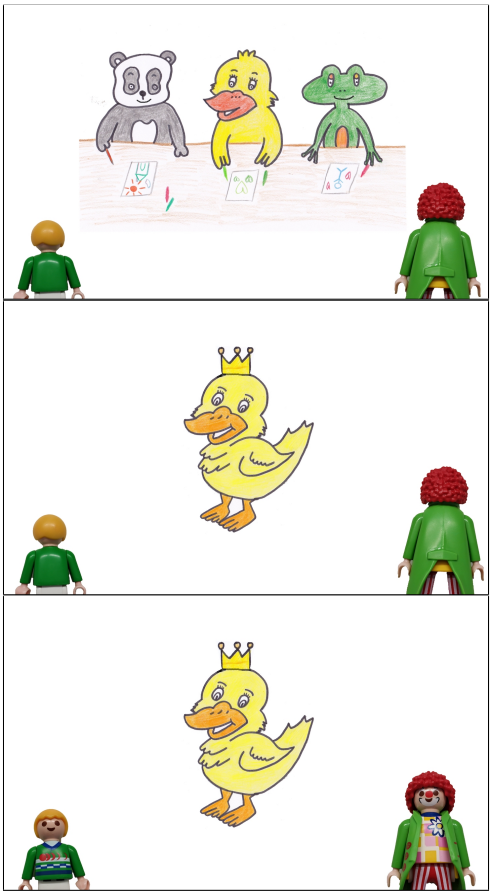
In addition to the critical items, adults, but not the children in order not to overwhelm them, were presented with 10 fillers in total, made up of 5 partial, non-exhaustive and 5 over-informative answers employing conjunction. One example for each is contained in the appendix.

On top of the audio, which was recorded in a sound-proof room and digitized at a 44.1 kHz sampling rate in a 16 bit mono format, all items contained a sequence of hand-drawn stills to lend visual support to the context and to encourage the children in particular to pay attention throughout the entire session. The contexts in the NRRC/Presupposition at-issue and the Assertion at-issue variant of each item as well as the target utterance by Peter used the same recording to minimize variance between both conditions and to avoid effects induced by prosody (cf. Stevens et al. 2017). In effect, they only differed with respect to the QUD posed by the clown.

7.3.4 Procedure

Experimental sessions for adults and children took place exactly the same way. In a quiet room and accompanied by an experimenter, the participant was sat in front of a computer running the experiment. First, all protagonists were introduced with a combination of recordings and hand-drawn pictures. Then the inattentive clown and Peter were introduced that same way. Both were present in the item videos with their faces turned towards the story characters and were said to be "watching" the stories along with the participants. Because of the clown's short attention span, he would ask questions (the QUD) at the end of the video and Peter would try to answer them with the target utterance. At this final stage, both Peter and the clown were shown facing away from the story and towards the participants. A complete vignette for the item in (106) is shown in Figure 2. It was the participant's task to judge whether Peter's response was felicitous with regard to the QUD; that is, if Peter's answer matched up with the clown's inquisitive goal – exact wording of the task: *Wie gut passt Peters Antwort zu der Frage vom Clown?* ('To what extent does Peter's answer match up with the clown's question?'). For this judgment, we used the 5-point Likert scale in Figure 3, which is based on Antomo et al.

(2018), who employed that same methodology and scale in the context of investigating reflexes of deceptive utterances in language learners in a similar age group.



The panda, the duck, and the frog are having a drawing competition. All like the duck' s drawing the most.

The duck receives a crown as a prize.

The clown did not pay attention again and asks,
Assertion at-issue: "Did the duck do the best at the competition?"
Presupposition at-issue: "Did the duck participate in the competition?"
Little Peter responds, "The duck won the competition."

Figure 2: Images accompanying the item in (106) in both the Assertion at-issue and the Presupposition at-issue condition.

Figure 3: Likert-type scale from 1 (red) to 5 (green).

Following the instructions, there were three warm-up trials without presupposition triggers or NRRCs. Of the three, two were inadequate with respect to the clown's QUD. The three warm-up items as well as the fillers for the adult participants are included in the appendix. With the children, extreme care was taken to familiarize them with the smiley-coded scale and warm-up trials could be repeated optionally. After the end of this introductory phase, the items described in the Materials section were presented in randomized order with a balanced mix of at-issue and non at-issue items (both across trigger type and relative to the total number of items). To do this, two lists were created such that all participants would see each critical item but only in one variant. At this

stage, no repetitions were possible. Upon completion of all trials, children were given a small gift.

7.3.5 Predictions

The general prediction for our experiment is the following: In the NRRC/Presupposition at-issue condition, Peter's utterance is infelicitous, and we thus expect ratings at the lower end of the scale. The inverse holds for the Assertion at-issue condition.[32] Considering that children display what is called pragmatic tolerance, observed by Katsos and Bishop (2011), we expect the difference between the two conditions to be larger with adults than with children.

Turning to the two types of presupposition triggers, in accordance with Xue and Onea (2011) and the theoretical accounts we reviewed in Section 7.2.1, we predict that soft triggers should be rated better than hard triggers in the Presupposition at-issue condition. As for NRRCs, our baseline trigger, we expect that they pattern with hard triggers, since their not-at-issue status is conventional.

7.3.6 Results

The data was analysed using R (version 4.0.0, R Core Team 2018). Specifically, we fitted a linear mixed model (LMM) using the *lme4* package Bates, Mächler, et al. (2015), adding as fixed effects Age, Trigger, and Issueness, as well their interactions. All predictors were sum-coded. The random effects structure consisted of by-participant random intercepts and random slopes for trigger type and at-issueness manipulations, and by-item random intercepts with random slopes for the factor Issueness.[33] Though the most maximal random effects structure would also include by-item random slopes for age groups, and interactions between random slopes for by-item and by-participant random effects (cf. Barr et al. 2013), these models failed to reach convergence.

In order to generate p-values from the linear mixed model, we employed likelihood ratios between the fully specified model and restrictive models, which left out one parameter at a time. To achieve this, the `mixed` function from the *afex* package (Singmann et al. 2020) was used. An overview of the results can be found in Table 4.

For the pairwise comparison between at-issue NRRCs and hard triggers, we used the *emmeans* package (Lenth 2021) and compared the estimated marginal means from the LMM using Satterthwaite's method for the effective degrees of freedom. In addition, we also computed by-group type III tests on the estimated marginal means of the linear

[32]Note, that this is in contrast to Aravind and Hackl (2017), who found adults and children to adopt a different QUD from the one implicit in the context. Because the QUD is given explicitly in our experiment, however, we do not expect this strategy to be very applicable here.

[33]Y ∼ Age*Issueness*Trigger + (Issueness + Trigger | Participant) + (Issueness | Item).

Table 4: LMM overview. p-values based on likelihood ratios.

Model Parameter	df	χ^2	df_χ	p
AGE	25	0.1	1	0.75
ISSUENESS	25	69.05	1	< 0.001
TRIGGER	24	22.32	2	< 0.001
AGE×ISSUENESS	25	57.66	1	< 0.001
AGE×TRIGGER	24	15.15	2	< 0.001
ISSUENESS×TRIGGER	24	33.59	2	< 0.001
AGE×ISSUENESS×TRIGGER	24	34.59	2	< 0.001

model. Because the NRRC condition served as baseline for general effects, but is not involved when it comes to the hard-soft distinction, they were excluded from this analysis. This gives us insight into whether the effects, in particular the interaction of ISSUENESS and TRIGGER, are also present within each level of grouping factor. These calculations were also performed with the *emmeans* package (using the `joint_tests` function) and are reported in Table 5.

Table 5: Inspection of the two-way interaction for each group to the exlusion of appositive relative clauses in the TRIGGER condition.

	Model Parameter	dfs	F	p-value
Adults	TRIGGER	1, 59.71	32.01	< 0.001
	ISSUENESS	1, 67.61	195.24	< 0.001
	TRIGGER×ISSUENESS	1, 60.96	94.97	< 0.001
Children	TRIGGER	1, 71.48	1.13	0.29
	ISSUENESS	1, 92.90	8.15	< 0.01
	TRIGGER×ISSUENESS	1, 121.72	4.17	< 0.05

Consistent with our predictions, there was a main effect for the factor ISSUENESS such that in the NRRC/Presupposition at-issue condition, Peter's utterances are judged as ill-suited. Additionally, there was a main effect of TRIGGER, indicating that, indeed, the triggers are judged heterogeneously. This, at least for the adults and the two kinds of presupposition triggers, is further confirmed by the supplementary group-wise analysis, which revealed a main effect for TRIGGER. As for AGE, there was no significant effect, chiefly because group means for children and adults are very similar, even though composed of very different condition means on the one hand and resulting from more centrally-clustered rating behavior on behalf of the children on the other.

Turning towards interactions now, AGE featured significantly in the interaction with the factor ISSUE, which was, again, as predicted. Developmental stage thus interacts with how violations of at-issueness are perceived: children have a less averse reaction to this type of violation than more mature participants. Additionally, AGE and TRIGGER

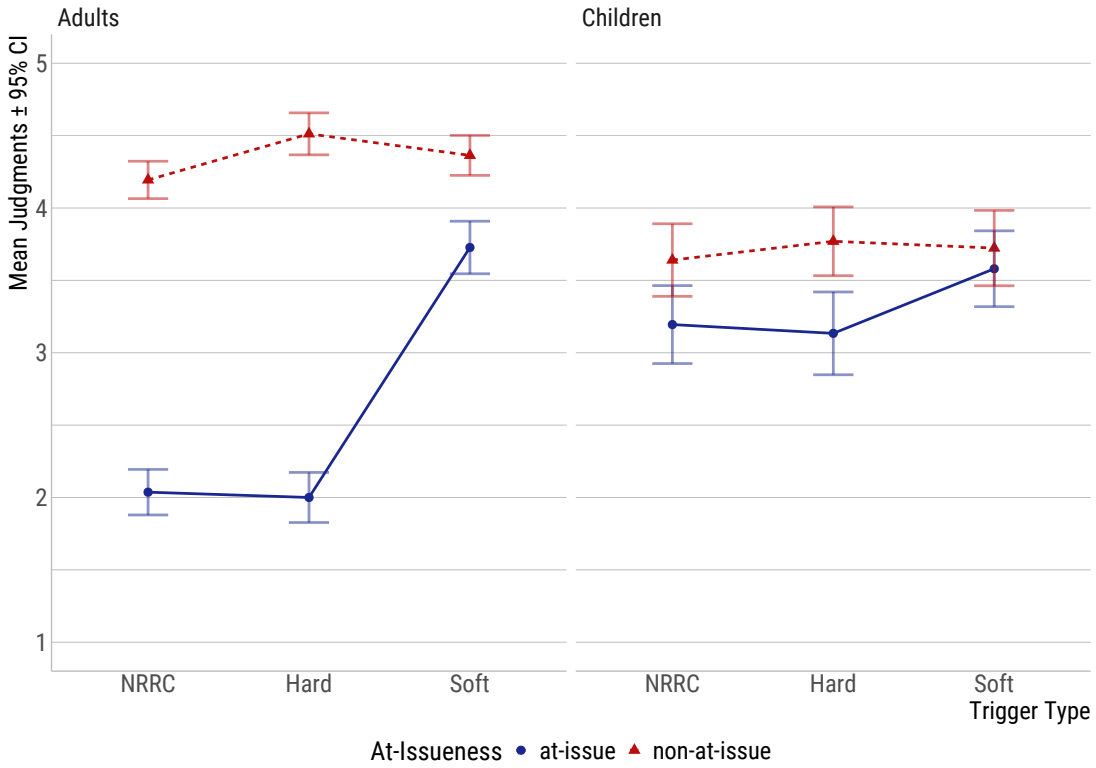


Figure 4: Mean judgments ± 95% confidence interval for the factors Age, Trigger and Issueness.

featured in a significant interaction. This can be viewed as the corollary of the previous: children and adults respond differently to the selection of triggers presented here. As can be seen in Figure 4, while adults judge soft triggers, particularly when the Not-At-Issueness Constraint is violated, quite differently from the other triggers, children are only tendential in this respect.

Finally, there was a significant interaction between Trigger and Issueness, as well as three-way interaction between Age, Trigger, and Issueness which substantiates the above interpretation. These interactions are once more indicative of the fact that triggers are differently suitable for violations of the backgroundedness constraint. Note also, that the two-way interaction is not mainly driven by the adults, but, as the effects for the estimated marginal means show, also by the children (though these effects are much less strong numerically). Further, unlike the adults, children do not make an across-the-board distinction between soft and hard triggers that is strong enough for a main effect of the Trigger manipulation, although at-issueness variation does affect the rating behavior differently depending on the trigger choice, mirroring the adults.

As for the *post hoc* result, at-issue NRRCs and hard triggers did not differ significantly with adult participants ($\Delta M = 0.002$, $\text{CI}_{95\%} = [-0.32, 0.48]$, $t(45.3) = 0.02, p = 0.98$). This motivates the exclusion of the NRRC condition for the results in Table 5.

Finally, one interpretation of the results might be that children did not understand the experimental task on account of the small numerical difference between conditions coupled with the general avoidance of extreme ratings on either end. The analysis shows, however, that this is not the case, as demonstrated by the within-group analyses which reveals a significant main effect for ISSUENESS, which cannot be explained unless a grasp of the Not-At-Issueness Constraint is assumed.

A last note regarding the disparity introduced by the way the QUD was posed: as can be seen from Figure 22 in the appendix, there was no effect of question type, i.e., between polar and wh-questions, on the acceptability judgments in either age group. Though follow-up experiments may shed light on how the semantic constraints induced by the question giving the QUD might affect the at-issueness constraint for different triggers, we will not return to this issue here.

7.4 General Discussion

Our experiment sought to answer the research questions in (103), repeated below.

(108) a. Heterogeneity of triggers: Can the soft-hard dichotomy be reflected in terms of the ability of different triggers to express at-issue content? Is there varying sensitivity to the Not-At-Issueness Constraint between triggers (possibly beyond the soft-hard dichotomy)?

b. Acquisition of presuppositions: Do children and adults respond differently to violations of the Not-At-Issueness Constraint? Are soft and hard triggers acquired at different times?

Regarding the first two questions in (108a), our results show that the triggers that are generally regarded as soft indeed behave differently from hard triggers in that they are much more suitable for expressing at-issue content. In other words, slightly different from the view defended by Aravind and Hackl (2017), our results show that the Not-At-Issueness Constraint in (98) is not an inviolable constraint after all. At least the presuppositions of soft triggers seem to be suitable, though not perfectly acceptable, for answering the QUD. Hard triggers, on the other hand, pattern with our conventionally not-at-issue, doubly marked appositive controls – the difference between the two is neither significant among adults nor with children. Thus, we can conclude, that the observed differences between soft and hard triggers in the domain of projection are reflected in the ability to express at-issue content.

The varying sensitivity to the Not-At-Issueness Constraint can be explained with local accommodation at center. Recall that in our experimental materials, presupposition triggers occurred in matrix contexts, which are not often discussed for global versus local

accommodation because the difference between the two is not easy to see there. However, since our experiment did not test projection but at-issueness – by manipulating whether presupposed material addressed the QUD or not – we can rely on a different effect of local accommodation to discriminate between the two types: at-issueness. A possible solution to explain our data is then this: Even in unembedded contexts, soft triggers allow for local accommodation, while hard triggers resist this strategy. Now, given our results, it seems that participants make use of local accommodation not only for reasons of projection but also to repair otherwise ensuing discourse failures. That is, participants treated the response in (109a) like (109b) to turn the backgrounded presupposition into an at-issue contribution.

(109) "Did the duck participate in the competition?" (at-issue)

a. Little Peter responds, "The duck won the competition."

b. Little Peter responds, "The duck participated in the competition and won (it)."

As we explained at the outset, global accommodation alone is not sufficient because it does not address the mismatch between addressing an at-issue question with non-at-issue meaning contributions. Now, one may be tempted to argue that our clown was presented as inattentive, which complicates the common ground assumptions. Supposing that the common ground is the set of proportions that all participants in a conversation jointly take for granted (for the purpose of the conversation), then arguably the presuppositions were not entailed by the common ground. Alternatively, if we allow for some flexibility and take the common ground to contain information participants can reasonably be expected to know, then it is much more likely that the presuppositions are entailed, despite the clown's continued inattentiveness. The first case seems to us rather less likely, because on this understanding, we would expect many more confirmatory utterances to definitively settle the propositions contained in the common ground, contrary to our intuition. In the second case, we would expect global accommodation to apply in order to settle some of the remaining uncertainty – which should not be difficult given the clear contextual setup for each trigger. However, regardless of the position one favors, global accommodation does not address the cause for the oddity of the discourse in the absence of a pertinent rescue strategy, namely the fact that the presupposition was used to address the QUD, in violation of the Not-At-Issueness constraint.

To look at a concrete analysis, consider the operator-based approach to local accommodation (Beaver and Krahmer 2001). Here, a covert operator is responsible for local accommodation under the assumption of a trivalent logic where the two classical truth values are supplemented with an additional one which indicates undefinedness. If the

operator is inserted into Logical Form (LF), it collapses the third, undefined truth-value, which results from presupposition failure, with the falsity conditions. In effect, in the presence of the operator, presuppositions are turned into assertions. As is the case with assertions, locally accommodated presuppositions are affected by entailment canceling operators, and do not project. But this is not the only consequence: Even though we standardly see local accommodation diagnosed via projection, turning presuppositions into assertions is also expected to come with consequences for their at-issueness status, because not-at-issue content typically projects. On this account, we expect for local accommodation to lead to higher acceptability of (originally) presupposed materials that is used to target the QUD in a given discourse.

Now, if we assume this operator-based analysis, an asymmetry between root contexts and embedded ones is actually unexpected. Much like the grammatical approach to scalar implicatures, where an operator is assumed to be responsible for upper-bounded inferences and which can apply freely at embedded levels as well as at the root, the same holds for the operator that leads to local accommodation. Hence, even though we cannot easily diagnose projection in environments without any entailment-canceling operators, the changed at-issueness status is diagnosable in question-answer pairs.

Regarding the difference between the hard and soft triggers, this perspective could, in principle, be used to argue for several of the approaches. First, note that using local accommodation in the way described above is a discourse-level feature and thus potentially favors approaches that champion features of the context as the deciding factors between the trigger classes. That is, both Abusch (2002, 2010) and Abrusán (2011, 2016) proposals are relevant here, as they both observe that some triggers are clearly more sensitive to contextual changes and correctly predict that presuppositions of those triggers can be easily targeted by such pragmatic factors. Granting that local accommodation can be used solely to change the at-issueness status of presuppositions rather than to invoke local readings relative to some operator, it should not come as a surprise that pragmatic factors, like context or focal stress, determine its application. By extension, hard triggers, which are not easily focused and more rigidly adhere to their discourse-anaphoric function, resist local accommodation as expected, while the presuppositions of soft triggers are more easily used at-issue.

Second, the proposals focused on the relation between entailment and projection behavior (e.g. Klinedinst 2016; Zehr and Schwarz 2016, 2018) are also worth discussing. According to their account, soft triggers obligatorily entail their presuppositions, so the assertion can always contribute the content as both presupposed and entailed meaning to the discourse. This property may also boost the context sensitivity of soft triggers, ease the local reading, and, unsurprisingly, improve the at-issue interpretation of their

presuppositions. For hard triggers, however, their pre-required content remains only presupposed, and such a reading is hardly available.

A natural follow-up question has to do with the issue of whether a bipartition of presupposition triggers is sufficient. Before any discussion, we need to emphasize that our experiment was not designed to detect differences between the lexical variants of the presupposition triggers of either group, as there were too few items per lexicalization to generate statistically reliable inferences. Still, we show the mean ratings per lexical realization in Figure 5 below.

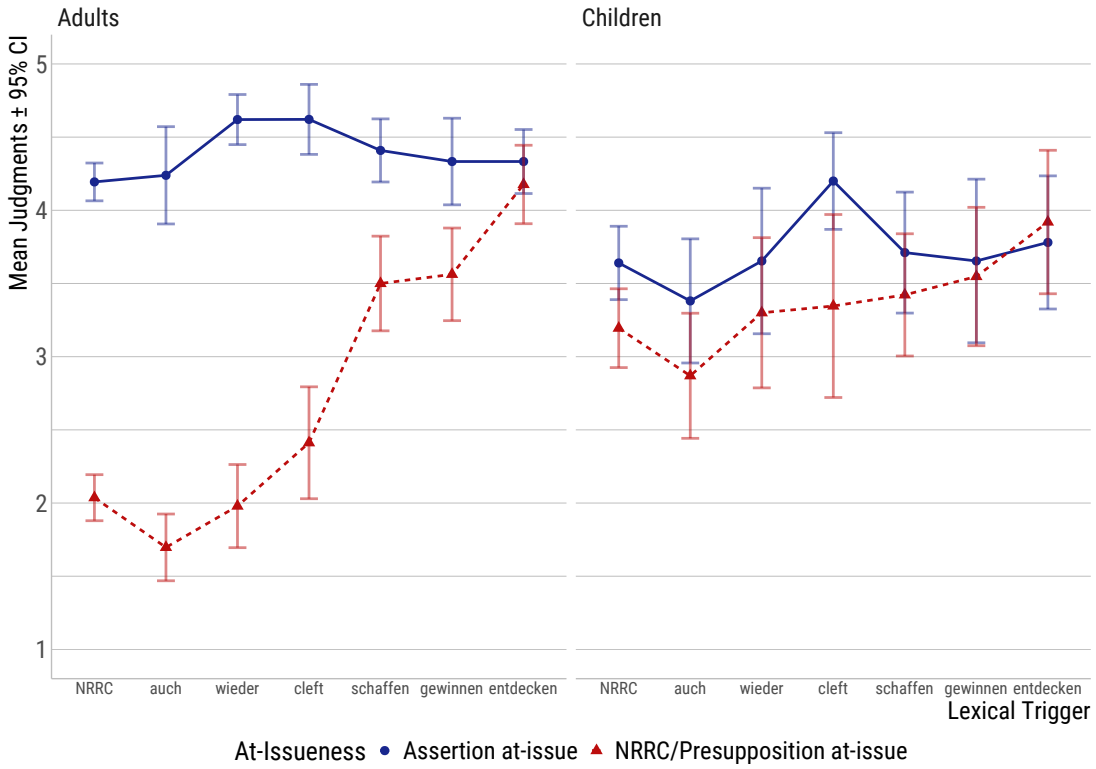


Figure 5: Mean judgments ± 95% confidence interval for the factors AGE, ISSUENESS, and lexical trigger.

The data visualized in Figure 5, especially the ratings given by adults, show some clues to answer this question. Comparing the mean rating of each trigger, the boundary between soft and hard triggers is quite clear: Although (almost) all the presuppositions used to address a QUD are ranked below all the main clause contents produced in answer to a QUD, the three soft triggers reach a position over the middle of the scale, whereas hard triggers remain at the lower end and thus constitute an inadequate response behavior indistinguishable from appositive controls. Even if we compare clefts, the trigger with the smallest Presupposition at-issue vs. Assertion at-issue ratings for the hard triggers, with *manage* ('schaffen'), the one with the largest difference among the soft triggers,

the discrepancy remains distinct. This, we think, is compatible with a bipartition of presupposition triggers together with a within-class level of variation.[34]

Note, however, that we carefully chose typical soft and hard triggers for our experiment and excluded those that did not clearly fall on either side of the dichotomy. The clear boundary between soft and hard triggers that we observe may be a consequence of our selection criteria. But given our data, it is conceivable that at least the concepts of soft and hard triggers can help to describe and investigate the heterogeneous behavior of presuppositions.

There is little research on the acquisition of at-issueness, especially regarding the questions in (108b). The significant interaction caused by AGE and ISSUENESS supports the conclusion that children and adults do indeed respond differently to at-issueness violations. Compared to adults, children had a less severe reaction to the violation, as indicated by their more centrally located rating behavior. At the same time, the *post hoc* results show a significant difference between children's ratings for the NRRC/Presupposition at-issue and Assertion at-issue conditions, indicating that preschool children between four and six are already aware of the not-at-issueness of certain contents – in line with the results in Aravind and Hackl (2017). Thus, we can conclude that children in that age bracket are already in the process of acquiring information structure.

Returning to the clustering of child ratings in the center of the scale, there are, at least, two ways to interpret this relative to the more extreme ratings by our adult participants: One possibility is that the ratings of the children reflect an incomplete acquisition of (not-)at-issueness. A second line of interpretation is to draw on previous results in the literature on implicature acquisition. Many studies show that children have a general tendency to avoid extreme ratings and that metapragmatic tasks obscure children' s pragmatic competence, possibly since children are less confident about their meta-linguistic judgments. Furthermore, children tend to penalize pragmatic oddity less strictly than adults (see e.g. Bernicot et al. 2007; Katsos and Bishop 2011). Thus, we can only speculate whether the differences between children and adults result from an incomplete acquisition or from a more general tendency to avoid extreme ratings.

Addressing the second question in (108b), our data might only offer an indirect inference. On the one hand, our results show that children are less sensitive to the

[34]It stands out that the difference between Presupposition at-issue and Assertion at-issue ratings is the smallest for *discover* ("*entdecken*"). This is not surprising, given the fact that, among others, Karttunen (1971b) and Hooper and Thompson (1973, p. 480) classify verbs like *discover* or *find out* as semifactive verbs, which can lose their factivity under certain circumstances and do not necessarily project under embedding. And, as it happens, Simons (2007, p. 1035) already observes that a complement clause embedded by semifactive *discover* can express at-issue content in English; a result, which was later confirmed empirically for German *erfahren* (Xue and Onea 2011) and *entdecken* (Antomo 2016). Following our line of reasoning, it is conceivable that semifactive verbs allow more easily for local readings than factive predicates, which are typically assumed to be hard triggers.

difference between hard and soft triggers, at least in numerical terms. In addition, we performed a further *post-hoc* analysis to see whether there is also a three-way interaction after removing the NRRCs triggers. This, again, was done using the *emmeans* package in the method described above. As with the full set data, this interaction was significant ($F(1, 1317.80) = 28.498$, $p < .0001$), indicating that once a presupposition is forced to be at-issue using an explicit QUD, adults are significantly more likely to accept sentences with soft triggers than those with hard triggers, while this discrepancy is less pronounced when we consider the results of our children participants – however, note that even for the children, we have a significant interaction between TRIGGER and ISSUENESS (see Table 5). Recall furthermore that their rating pattern for each trigger type is similar to the adults as shown in Figure 4. Thus, we can conclude that, although less sensitive than adults, children in the tested age group are sensitive towards the trigger type distinction (or its effects) and that they already figured out "which box a specific expression should go into" (Zehr and Schwarz 2018, p. 479), at least with the support of an explicit QUD.

On the other hand, recall that children are already aware of at-issueness violations in general but there is no main effect of TRIGGER (see Table 5), which differs from the adults' rating. Thus, we might be allowed to assume that preschool children struggle more with the difference between presupposition triggers than with at-issueness or backgroundedness, even if the acquisition process of both features is ongoing. If so, one possible explanation for this pattern is that softness, instantiated as contextual sensitivity according to the semantic-pragmatic split (e.g., Simons 2001; Abusch 2002, 2010), entailment (e.g., Klinedinst 2016; Zehr and Schwarz 2016, 2018) or the interaction between triggering mechanism and context (e.g., Abrusán 2011, 2016) is recognized only with additional pragmatic knowledge and working resources. In that case, the acquisition of this feature might be assumed to take place later. In this sense, our results can be considered consistent with Yatsushiro (2008a), who observes among children from six to nine that even presuppositions associated with the same word, namely *every*, the existential presupposition, which is considered as semantic, is acquired clearly earlier than anti-uniqueness and anti-duality inferences, which arise pragmatically. This result is highly interesting for theoretical accounts based on pragmatic factors like focus: If one and the same lexical item triggers separate presuppositions including both hard and soft ones, it is difficult to assume that focus is the driving factor behind hardness (or softness), since, presumably, one lexical item will not have two distinct ways of being focused. This observation may support the assumption that a stronger connection with semantic meaning, as with hard triggers, facilitates the acquisition and comprehension of language usage.

A natural follow-up question is then whether we can also detect this tendency with children. That is, whether the discrepancy between soft and hard triggers escalate with in-

creasing age. To test this, we performed a median split between the younger participants by age (median age 5;1), illustrated in Figure 6. There, we do find some incongruence between the two groups, and perhaps there is some evidence to suggest that a developmental change is underway. Note, however, that our study was in fact underpowered in this respect. What we can actually find with our data is a trigger contrast (reflected in TRIGGER×ISSUENESS) in the child data, which is weaker than within adults (reflected in the *post-hoc* three-way interaction). Thus, we do not know for sure whether the sensitivity to the trigger contrast is acquired later than the general sensitivity to at-issueness-violations. To further strengthen the reliability of this interpretation, more children on either half of the split would need to be investigated. We leave this interesting issue for future research.

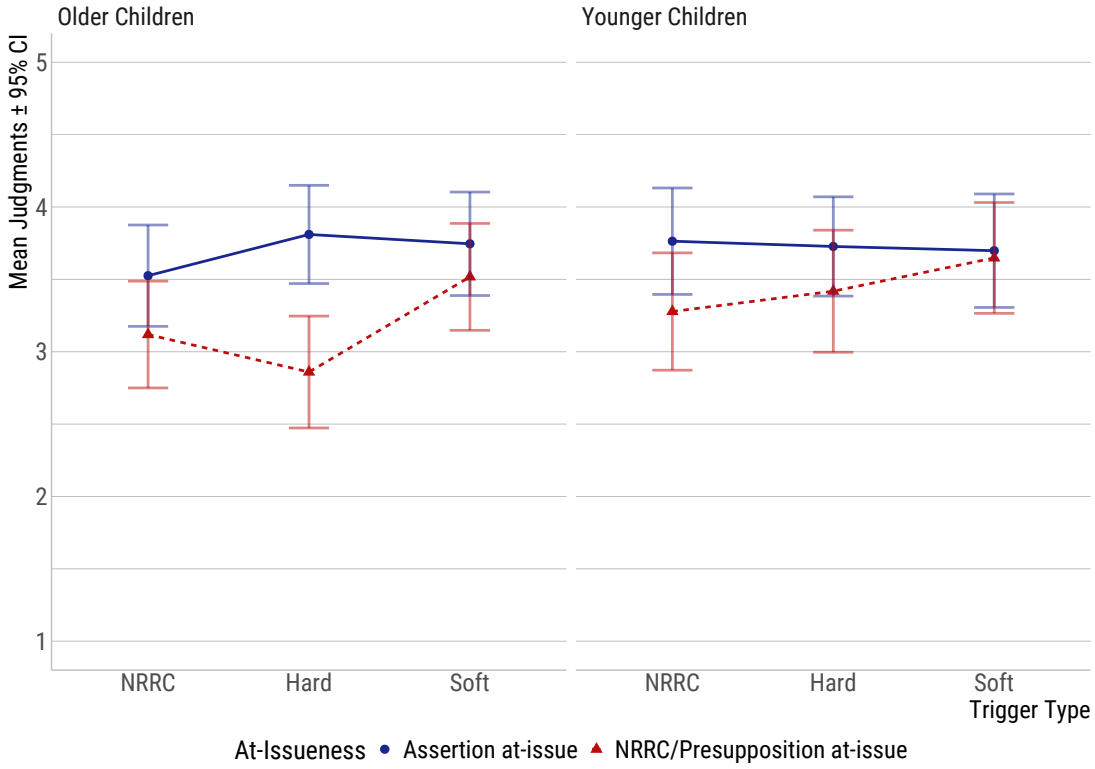


Figure 6: Median split between children in our experiment.

Taking stock, our results indicate (i) that the heterogeneity of presuppositions is not restricted to projection but also extends to their ability to answer a question directly and (ii) that children are not only able to discern felicitous discourse structures from incoherent ones, but are also aware that presuppositions are not a homogeneous class, demonstrating a sophisticated awareness of the principles that regulate discourse.

7.5 Conclusion

To conclude, using a modified acceptability judgment task with adult controls and children, we showed that preschool children, in principle, are aware of violations of the Not-

At-Issueness Constraint but are still less sensitive to it than adults. Additionally, the heterogeneous behavior between triggers, which has led to the assumption of a soft-hard split in parts of the literature, is reflected in an asymmetry regarding the ability to express at-issue content (via local accommodation). While hard triggers pattern with lexically marked not-at-issue environments (NRRCs), soft triggers are more suitable (though not perfectly acceptable) for answering the QUD. Thus, our results might be better explained in terms of a two-way-split of triggers, rather than a continuum, even if we can observe a certain level of within-class variation. So far, hard and soft triggers have been distinguished based on projection. In this paper, we have shown that the presuppositions associated with hard and soft triggers can also be distinguished on other grounds, namely their differing ability to felicitously answer a QUD.

8 Exp.II: Trigger Classification and At-issueness in L2 Acquisition[35]

8.1 Introduction

Regarding the three main research concerns and the five main questions listed in (85), the second experiment of my study mainly focuses on the first question on verifying the classification, and the fourth question on presupposition acquisition in second languages (L2), repeated below:

(110) Research question 1: verifying the classification: can the soft-hard dichotomy, a classification based on the projective strength, also be reflected in terms of the ability of different triggers to express at-issue content?

(111) Research question 4: acquisition in second languages: are nonnative speakers as sensitive as native speakers once the presuppositions of different triggers (or trigger types) are forced to be at-issue? If not, how can the difference be explained?

As the theoretical and empirical background of these two research questions has already been reviewed in detail in Chapter 3 and Section 4.2, in this section, only the most relevant findings and related aspects for the second experiment will be briefly summarized.

Regarding Research question 1, the first term that should be shortly explained again is the soft-hard split. As reviewed before in Section 3.1, generally speaking, it is observed that although all presuppositions project, their projection behavior in certain contexts varies. This heterogeneity of presuppositions can be illustrated for instance with the explicit ignorance context (Simons 2001) as shown in (47), repeated below in (112), or under the scope of some quantifiers (Charlow 2009), compare for example (49) and (50), repeated below in (113):

(112) With explicit ignorance context:

a. # I don't know whether anyone else was ill, but if Pikachu was ill *too*, his friends were worried.

[35] The experiment in this chapter was partially orally presented in *XIV. Kongress der Internationalen Vereinigung für Germanistik*, section C6 *Kontrastive Pragmatik*. A short summary in German with the title *Wenn Nicht-Muttersprachler/innen mehr Toleranz zeigen: Eine experimentelle Studie zu Präsuppositionen, At-issueness und DaF* has been published in the conference proceeding (see Y. Chen 2022).

b. I don't know whether Pikachu participated in the game, but if he *won* the game, he was happy.

(113) Under scope of quantifiers:

a. Some of these students *also* smoke[s] $Marlboros_{Focus}$.

b. More/Less than three of these students *also* smoke $Marlboros_{Focus}$.
$\rightsquigarrow$ Each of these students smokes something other than Marlboro.

c. Some of these students *stopped* smoking.

d. More/Less than three of these students *stopped* smoking.
$\not\rightsquigarrow$ Each of these students used to smoke.

(Romoli 2014, p. 176, slightly modified)

To explain this heterogeneity, one relevant and widely discussed theoretical account assumes that there are different kinds of presuppositions or triggers (for more, see also the review in Section 2.2.2), which can be traced back at least to Kripke (2009) (oral presentation in 1990) and Zeevat (1992). The terms *soft* and *hard* are firstly proposed by Abusch (2002, 2010), who claims that presuppositions of soft triggers are more suspendable or less projective as they are generated pragmatically, while presuppositions of hard triggers are rather lexical (for more details, see Section 3.1). Although her claim of a triggering difference is challenged both theoretically and empirically (see e.g. Abrusán 2011; Jayez et al. 2015; Abrusán 2016, among others), the soft-hard split based on projection strength is considered helpful in describing the heterogeneity of presuppositions, and is perhaps also reflected in some other properties, such as entailment (see e.g. Klinedinst 2016; Zehr and Schwarz 2016, 2018) and at-issueness (see e.g. Xue and Onea 2011; Tonhauser et al. 2018). Thus, the terms *soft* and *hard* are adopted in my study but used in a wider sense – that is, soft is used to refer to those triggers whose presupposition is generally assumed to be more pragmatical or conversational, more similar to scalar implicatures, more likely to be locally accommodated, or can be canceled or suspended (more) effortlessly, such as achievement verbs like *win* and *manage*, or semi-factive verbs like *discover*. Presuppositions of hard triggers, such as *too*, *again*, and *it*-clefts, on the other hand, are generally considered more stable, less pragmatic – even semantic or lexical – and highly projective.

Another relevant term in the first research question is at-issueness, a notion that is based on the concept of Question Under Discussion (QUD) following Klein and Von Stutterheim (1987) (with the concept 'Quaestio') and Roberts (1996), and which has been introduced before in Section 3.2 in detail. In brief, in this model, every discourse aims to answer an explicit or implicit question – the QUD. A QUD can be understood as a set of alternatives that relates to the current topic in discourse. The definition of at-issueness

is given by Simons et al. (2010) in (53), repeated below in (114). Taking the speaker intention also into account, they further revise at-issueness as follows:

(114) A proposition p is **at-issue** relative to a question Q iff *whether p* is relevant to Q. (Simons et al. 2010, p. 317, sightly modified)

(115) Revised definition of at-issueness:

a. A proposition p is at-issue iff the speaker intends to address the QUD via *whether p*.

b. An intention to address the QUD via *whether p* is felicitous only if:

i. *whether p* is relevant to the QUD, and

ii. the speaker can reasonably expect the addressee to recognize this intention.

(Simons et al. 2010, p. 323, sightly modified)

As we have seen before in 3.2, according to Simons et al. (2010), presuppositions belong to contents that are linguistically and conventionally marked as not-at-issue, while the entailment-canceling operators like negation or question can only target at-issue contents. Therefore, the projection of presuppositions is due to their not-at-issueness: as not-at-issue contents, presuppositions are ignored by these operators and can thus survive and project. Based on this characteristic, Aravind and Hackl (2017) propose the Not-At-Issueness Constraint of presuppositions in (55), repeated below again in (116):

(116) Not-At-Issueness Constraint:
Presuppositions cannot be used to directly target the Question Under Discussion.
(Aravind and Hackl 2017, p. 51)

Moreover, the concept of not-at-issueness also explains one important asymmetry between projection and presuppositions: although all presuppositions can project, not all contents that can project are doubtlessly presupposed. Take non-restrictive relative clauses (NRRCs) for instance: their content is projective, as shown in (56a), but should rather be new information in the context than presupposed. Therefore, similar to the first experiment, NRRCs will also be included and used as baseline for not-at-issueness in this experiment. Additionally, as mentioned before, the perspective of at-issueness may also shed light on the different projection behaviors of presuppositions. According to the experimental study by Tonhauser et al. (2018), projection behavior and not-at-issness are positively correlated: a presupposition is more projective if it is more not-at-issue (for more discussion on these two points, see Section 3.2).

Combining the definition of at-issuenss, the Not-At-Issueness Constraint, the correlation between at-issueness and projection with the soft-hard split, the hypothesis regarding Research question 1 arises:

(117) **Hypothesis 1**: Presuppositions of hard triggers, which normally project more steadily and strongly, should be less suitable for targeting the QUD directly than those of soft triggers.

As this hypothesis has already been confirmed by the German native adults in the first experiment, see Chapter 7, the experiment in this chapter further aims at unveiling whether the relation between the soft-hard split and at-issueness can also be recognized by nonnative speakers, and whether their sensitivity is influenced by the pragmatic competence and knowledge in their first and foreign languages, as is expressed in Research question 4.

As reviewed in Section 4.2, in comparison to presupposition acquisition in the first language (L1), there is far less research on presupposition acquisition in second languages (L2). According to my literature research, the only study that makes a comparison between native and nonnative speakers to investigate some properties of presuppositions is conducted by Carrell (1984). As this study has been reviewed in detail before, it is only concisely summarized in this section. Carrell's experiment has a 2×2×3 design. That is, two different types of triggers, namely factive predicates and implicative predicates as defined by P. Kiparsky and C. Kiparsky (1970) and Karttunen (1971a), with two different semantic status, namely semantically positive, such as *be thoughtful* and *remember*, and semantically negative, such as *be thoughtless* and *forget*, were tested among three group of participants: native English speakers, and nonnative English speakers at advanced level as well as at high-intermediate level. For clearer illustration, two example sets of her materials with the 4 types of inferences in test in (71) are repeated below:

(118) a. John remembered to let the dog out.
The dog is out. (Implication, true)
The dog is in. (Implication, false)
The dog is supposed to be out. (Presupposition, true)
The dog is supposed to be in. (Presupposition, false)

b. It was thoughtless of John to let the dog out.
The dog is out. (Presupposition, true)
The dog is in. (Presupposition, false)
The dog is supposed to be out. (Implication, false)
The dog is supposed to be in. (Implication, true)

(Carrell 1984, p. 9)

Using a Truth Value Judgment Task, Carrell shows that both groups of nonnative speakers are less accurate than native speakers with respect to all inferences types, indicating

that inferences related to pragmatic skills like presuppositions need to be acquired in L2, and nonnative speakers at advanced or high-intermediate levels are still in this process. Moreover, all three groups understand the implications better than the presuppositions of both implicative and factive predicates, and the semantically positive words better than the negative ones. Furthermore, the interaction of these two parameters is more significant in nonnative speakers than in native speakers. According to Carrell, this observations support the distinction between presupposed and implicated inferences (see e.g. Clark 1977; Carrell 1978) and prove the importance of native-nonnative comparison in contrastive pragmatics.

Note, however, that the mother tongues of the nonnative participants in her experiments are not the same but include several different languages like Chinese, Korean, Japanese, and more. Additionally, the triggers tested in her study are restricted to two subsets, whereas several classical triggers are not included, especially those that are typically considered hard. Moreover, recall that the review in Section 4.2 has shown that firstly, several studies on other related phenomena like implicatures show that the pragmatic skills in L2 can be influenced by L1 or the similarity between the two languages (see e.g. Taguchi 2013), demanding a strict control of nonnative participants. Secondly, the very few studies on presuppositions in L2 mainly focus on teaching strategies of writing or speaking skills (see e.g. Al-Zubeiry 2020), whereas an investigation of properties of presuppositions from the L2-perspective is generally missing, not to mention studies on trigger distinction. Thirdly, similar to children's pragmatic tolerance in L1 acquisition (Katsos and Bishop 2011), such tolerance might also be observed in nonnative speakers during the L2 acquisition process (Y. Chen 2019). This nonnative speakers' pragmatic tolerance, if it exists, can also be a relevant aspect in the investigation of the heterogeneity problem of presuppositions, especially the soft-hard split of presupposition triggers.

Considering all these points mentioned above, it is then both necessary and reasonable to investigate the heterogeneity problem of presuppositions and verify the soft-hard split by examining the nonnative speakers' reaction to presuppositions with at-issueness violations. That is, in comparison to native speakers, how sensitive are they to the (cor)relation mentioned in Hypothesis 1 in (117)? Combining the observations in Carrell (1984) and previous studies like Y. Chen (2019), the second hypothesis of the test is formulated as follows:

(119) **Hypothesis 2**: Nonnative speakers are less accurate or less sensitive towards pragmatic inferences such as presuppositions and are more tolerant towards pragmatic violations than native speakers.

Based on Hypothesis 1 and 2, this study adopts the design and the materials from the first experiment and tests both native German speakers and Chinese adults who learn

German as a foreign language (DaF). The particular research questions of this experiment can be specialized as follows:

(120) a. Can the Chinese DaF-learners also notice that presuppositions generated by soft and hard triggers differ in their ability to express an at-issue content in German, extending Hypothesis 1 in (117) to L2?

b. If the DaF-learners can notice the difference, do they react just as sensitive as the German native speakers to the at-issueness violation? If not, are nonnative speakers less sensitive but more tolerant towards the violations, as Hypothesis 2 in (119) assumes? Is the nonnative speakers' deviation from the native speakers stable for each trigger type, or are there differences between triggers?

c. If there are some differences between both groups, how can these differences be explained? Can the differences shed light on the classification issue?

8.2 Experiment

8.2.1 Design and Materials

The design and materials were adopted from the first experiment, see Chapter 7, and remain almost identical; therefore, I only summarize them shortly here.

The experiment had a $3 \times 2 \times 2$ design, i.e. with three parameters in this test: the first one was TRIGGER with three levels: NRRC vs. hard presupposition trigger vs. soft presupposition trigger, between-item and within-subject. With 10 items for each trigger type, the hard triggers used in the test were *auch* ('also', 'too', 3 items), *wieder* ('again', 3 items) and *it*-clefts (4 items). The soft triggers were *entdecken* ('discover', 3 items), *gewinnen* ('win', 3 items) and *schaffen* ('manage to', 4 items). The NRRCs were doubly marked: on the one hand, the antecedence of the relative pronoun was always a unique character in the discourse so that the interpretation had to be non-restrictive (cf. Jacobsson 1994); on the other hand, the word *übrigens* ('by the way') was added in the relative clauses in order to emphasize the not-at-issueness. As mentioned before, NRRCs are not-at-issue but not presupposed, they functioned therefore as the baseline for not-at-issueness in the test.

The second parameter was ISSUENESS with two levels: assertion at-issue vs. presupposition (PSP)/NRRC at-issue, within-item and and within-subject. Accordingly, each critical item had two variations: for one, the question is answered by the assertion, which should be natural and adequate; for the other, the explicit QUD in the item targets the presupposition of the answer or the content of NRRC in the answer, so that they are forced to be interpreted at-issue and the Not-At-issueness Constraint is violated. Using a latin-square-design, every participant judged 5 items with at-issue PSPs/NRRCs and

5 natural ones for each trigger type. Examples for each trigger type are translated into English and presented below (for the complete lists of critical items and their translations, see Appendix A.1.2 or in Y. Chen et al. 2022).

(121) **NRRC:**

The duck has a cough. She gets some cough syrup. Then she goes to bed in order to rest. The clown did not pay attention again and asks:

a. assertion at-issue:
   "Where is the duck now?"

b. NRRC at-issue:
   "What did the duck get?"

Little Peter responds, "The duck, who got some cough syrup by the way, is in bed now."

(122) **Hard trigger:** *wieder* ('again')

The panda is a member of a soccer club. Recently, he has been playing very well and scored many goals in the past weeks. Yesterday there was a football game and he scored a goal. The clown did not pay attention again and asks:

a. assertion at-issue:
   "Did the panda score a goal yesterday?"

b. PSP at-issue:
   "Did the panda score a goal for the first time yesterday?"

Little Peter responds, "The panda scored a goal again yesterday."

(123) **Soft trigger:** *gewinnen* ('win')

The panda, the duck, and the frog are having a drawing competition. All like the duck's drawing the most. The duck receives a crown as a prize. The clown did not pay attention again and asks:

a. assertion at-issue:
   "Did the duck do the best at the competition?"

b. PSP at-issue:
   "Did the duck participate in the competition?"

Little Peter responds, "The duck won the competition."

The last parameter was GROUP with two levels: the native German speakers (group DE) vs. the Chinese students who learn German as a foreign language (group DaF), within-item and between-subject. For more about the participants, see Subsection 8.2.2.

Similar as in the first experiment, the items were presented as short videos with a short story about three little animals: the duck, the frog and the panda. Two characters in picture watched the videos together with the participants: the clown, who did not pay attention and always asked a question as explicit QUD, and Peter, who paid attention and answered the question, but sometimes not with a properly formulated sentence as he was still very young (see Figure 1 for the three animals and the two characters). The participants were asked to judge how appropriate did Peters answer fit the clown's question and marked it on a 5-point scale from 1 for totally infelicitous to 5 for totally felicitous, adapted from Antomo et al. (2018). The entire experiment was programmed using the open source software OnExp (version 1.3.1).

Slightly different from the first experiment, in this one, both groups saw 10 fillers besides the 30 critical items in order to test their attitude to under-informative and over-informative answers without presuppositions. Among them, 5 fillers had partial, non-exhaustive answers that are considered as under-informative (UA-type), and the answers of the other 5 fillers were not only exhaustive but also included information that was not required. They are considered as over-informative (OA-type). One example of each type is given below:

(124) **UA-type:**

The animals are having lunch together. The panda and the frog are eating sausages. The duck is eating pizza.

The clown did not pay attention again and asks: "Which animals are eating sausages?"

Little Peter responds: "The frog is eating sausages."

(125) **OA-type:**

The animals are having breakfast together. The panda and the duck are eating muesli. The frog is eating bun.

The clown did not pay attention again and asks: "What is the panda eating?"

Little Peter responds: "The panda is eating muesli and the frog is eating bun."

(see Appendix A.1.3 or in Y. Chen et al. 2022 for the whole list of filler items)

8.2.2 Participants

The critical group of this test consists of 36 Chinese adults who learn German as a foreign language (group DaF, 30 female, mean age: 21.39 ± 0.92). All of them were bachelor students at Nanjing University and had been exchange students at a German university for at least one semester. Their German language levels were between B2 and C1 according to the Common European Framework of Reference for Languages (CEFR)

standard, meaning that they were advanced learners and able to understand complex German texts. They should possess all necessary language skills in order to understand the items in the test, at least literally, but were still not able to use German as well as native speakers like the C2 students (for more about the CEFR levels, see Council of Europe 2001).

The control group consists of 33 adult German controls (group DE, 25 female, mean age: 25.72 ± 10.3). All were monolingual German native speakers from Göttingen and the surrounding area.[36]

None of the participants had prior linguistic knowledge of presuppositions, projective contents or related issues.

8.2.3 Procedure

The procedure also followed the routine used in the first experiment. The participants were invited to the test room/office room at Universität Göttingen or Nanjing University, with laptops running the test program. After the introduction of the three animals and the two characters, clown and Peter, three warm-up items were presented, so that they could understand the task (translation of the task question used in the test: 'To what extent does Peter's answer match up with the clown's question?') and the rating scale. If there was no problem, the critical part started, in which the 30 critical items and 10 fillers were mixed and randomized. Note that the factor Issueness is within-item, therefore, two lists of critical items were created according to latin-square. Every participant only saw one list with all 30 items, but each item only in one Issueness variation. During the test, the occurrence of both lists were controlled as balanced as possible. Among the DaF group, it was also announced that in the critical part, only questions about vocabulary could be answered, in other cases, the participants had to make a decision based on their intuition.

8.2.4 Predictions

Regarding the question in (120a), if the DaF-learners can also recognize the trigger distinction with regard to at-issueness in German, their judgments on at-issue presuppositions of soft triggers should be statistically different from those of hard triggers. A significant interaction of Trigger and Issueness can also be expected as the control group can definitely perceive the soft-hard split, which has already been observed in the first experiment (see Chapter 7).

Regarding the research questions in (120b), the predictions about the experiment results are as follows: If the pragmatic skills or intuition in one's mother tongue can be

[36]They also functioned as the control group in the first experiment.

automatically and completely adopted and used in a foreign language, the Chinese DaF-learners should react as sensitive as the native German speakers. If so, neither a significant effect of GROUP nor a significant interaction of GROUP and ISSUENESS should be observed. If their sensitivity is in line with Hypothesis 2, that is, the nonnative speakers react less sensitive than the native speakers, then their judgments on PSP/NRRC at-issue items should be higher than those of the control group. Thus, at least a significant interaction of ISSUENESS and GROUP should be observable. Additionally, if the nonnative speakers' (in)sensitivity varies between trigger types, then a three-way interaction should also be observed.

As the questions in (120c) mainly focus on theories, I will discuss them later in section 8.3.

8.2.5 Results

Starting the analysis with the judgments on the critical items by excluding all judgments on the warm-up items and fillers, the data then encompassed 2066 judgments taken from 69 individuals out of 2 groups for 30 critical item-pairs. The mean judgments with 95% confidence interval are presented in Figure 7.

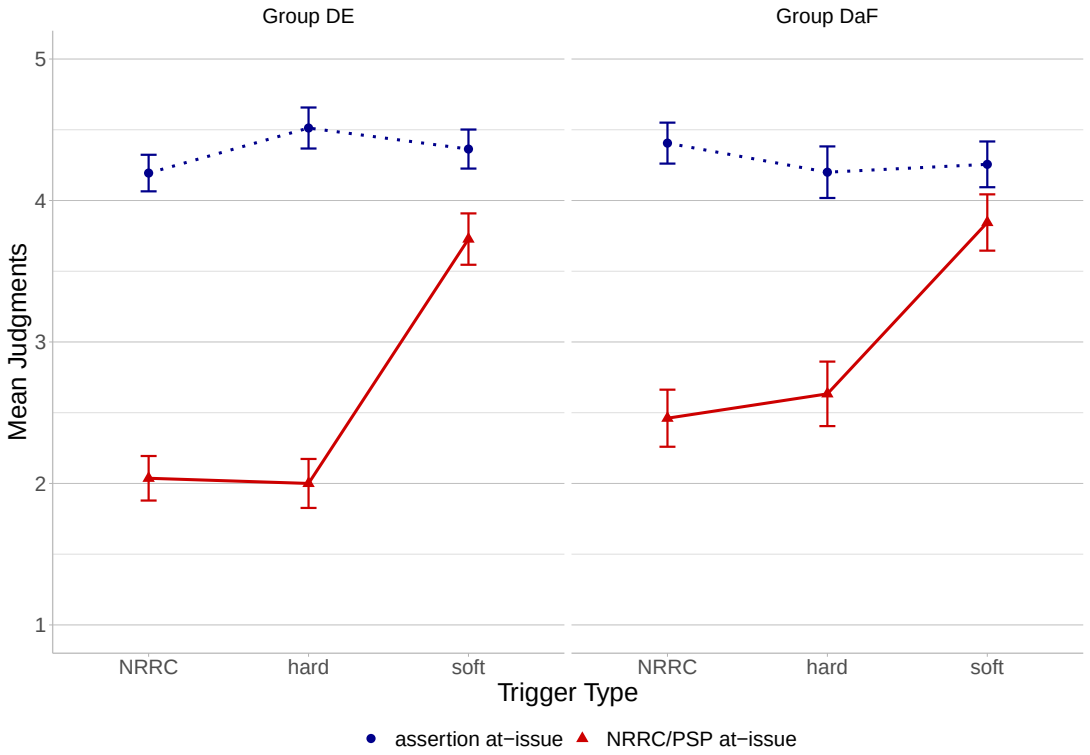


Figure 7: Mean judgments ± 95% confidence interval for the factors GROUP, TRIGGER and ISSUENESS.

To estimate the effects of TRIGGER, ISSUENESS, GROUP and their interactions, the data was analyzed using R (version 4.0.3, R Core Team 2020). Specifically, a linear mixed

model (LMM) was fitted using the *lme4* package from Bates, Mächler, et al. (2015) by maximum likelihood. In order to achieve the maximal random effects structure (cf. Barr et al. 2013), the random effects structure should consist of by-participant random intercepts and random slopes for the factor ISSUENESS and TRIGGER with their interaction, and by-item random intercepts with random slopes for the factor ISSUENESS and GROUP also with their interaction. However, this model obtained then a singular fit. Following Bates, Kliegl, et al. (2015), the model was too complex and the random structure needed to be reduced until the random-effect Principal Components Analysis (rePCA) does not report overidentification. Therefore, the final maximal random effects structure that neither caused converge failure nor led to a singular fit included only the random intercepts and the random slope ISSUENESS by subject[37]. With 121.53 observations per estimated term, the model was then not excessively complex. In the model, the reference levels of the fixed effects were set as follows: *NRRCs* for TRIGGER as they are the baseline among trigger types, *assertion at-issue* for ISSUENESS as it is the natural, standard answer type and *group DE* for GROUP as they are native speakers and constituted the control group in the test.

As the research questions in this experiment mainly focus on the difference between groups, the statistic analysis started with the full-null comparison (Forstmeier and Schielzeth 2011), in which the null model lacked only the fixed effect of GROUP and its interaction with other effects, and otherwise was the same as the full model.[38] The aim is an overall test of the effect GROUP and its interactions in order to avoid cryptic multiple testing. The full-null comparison shows that GROUP and its interactions had a clear, significant effect on the judgments ($\chi^2 = 33.85, df = 6, p < 0.001$).

Then, in order to generate p-values from the linear mixed model, the fully specified model and restrictive models which left out one parameter at a time were compared by using the `mixed` function with Type 3 tests from the *afex* package (Singmann et al. 2020). An overview of the results can be found in Table 6.

According to results of the analysis, both TRIGGER and ISSUENESS can clearly influence the appropriateness of the answer, indicating that the three types of triggers have a significantly different potential to express at-issue contents, and the violation of the Not-At-Issueness Constraint significantly declines the pragmatic felicity in general. The fact that GROUP achieved no main effects shows that the mean values of all ratings from the two groups were very similar overall.

With respect to the research questions, more important are the interactions. The significant interaction of GROUP and ISSUENESS indicates that the two groups indeed

[37] Y ~ TRIGGER*ISSUENESS*GROUP + (1+ISSUENESS| Participant) + (1 | Item).

[38] Y ~ TRIGGER*ISSUENESS + (1+ ISSUENESS | Participant) + (1 | Item).

Table 6: LMM overview with p-values based on likelihood ratio tests.

Model Parameter	df	χ^2	p-value
TRIGGER	2	36.32	<0.001
ISSUENESS	1	160.05	< 0.001
GROUP	1	2.40	0.121
TRIGGER×ISSUENESS	2	214.90	<0.001
TRIGGER×GROUP	2	7.52	0.023
ISSUENESS×GROUP	1	13.12	<0.001
TRIGGER×ISSUENESS×GROUP	2	13.14	0.001

react differently to at-issueness violations. In comparison to the German native speakers, the Chinese DaF-learners were, at least numerically, less strict or more tolerant towards the violation of the Not-At-Issueness Constraint.

Nevertheless, results also prove that the nonnative speakers were still able to recognize pragmatic infelicity. According to the paired comparisons given by *emmeans* package (Lenth 2021), the DaF-learners rated answers with at-issue PSPs/NRRCs significantly more infelicitous than those with at-issue assertions (with hard triggers: $\Delta = 1.53$, $t(335) = 12.12, p < 0.001$, with soft triggers: $\Delta = 0.41$, $t(324) = 3.30, p = 0.006$, with NRRCs: $\Delta = 1.94$, $t(324) = 15.60, p < 0.001$[39]). In other words, the nonnative speakers were able to recognize the violation of the Not-At-Issueness Constraint caused by at-issue presuppositions and NRRCs, but they were less sensitive than the native speakers.

The interaction of ISSUENESS and TRIGGER was also significant, indicating that the presupposed contents of the three trigger types are infelicitous to a varying degree once they are at-issue. Using the *emmeans* package (Lenth 2021) again, the paired comparison shows that, once the presuppositions or NRRCs are at-issue, there are significant differences between soft and hard triggers as well as between soft triggers and NRRCs but not between hard triggers and NRRCs. More importantly, this split can not only be observed among the native speakers, but also among the DaF-learners. The results are presented in Table 7.

This difference between trigger types also influenced the interaction of GROUP und ISSUENESS and caused the significant 3-way interaction. In other words, the difference between the reaction of the DaF-learners and the native speakers to Not-at-issueness Constraint violation varies between trigger types. Using the `joint_tests` function, again from the package *emmeans*, the results in Table 8 show that the interaction of ISSUENESS and GROUP is neither significant with respect to the soft triggers nor to NRRCs, but mainly significant with respect to the hard triggers.

[39]Degrees-of-freedom method: kenward-roger, p-value adjustment: tukey method for comparing a family of 6 estimates.

Table 7: Contrasts of judgments on at-issue presuppositions/NRRCs separated by group.

Group	Trigger contrasts	estimate	df	t	p-value
group DE	hard − soft	-1.69	132	-11.50	<0.001
	NRRC − soft	-1.69	133	-11.60	<0.001
	NRRC − hard	-0.01	133	-0.04	1
group DaF	hard − soft	-1.17	118	-8.19	<0.001
	NRRC − soft	-1.38	119	-9.76	<0.001
	NRRC − hard	-0.21	118	-1.50	0.67

Table 8: Interaction of Issueness and Group for each trigger type.

Interaction	Trigger type	dfs	F	p-value
Issueness × Group	hard	1, 326.56	27.04	<0.001
	soft	1, 324.36	1.52	0.219
	NRRC	1, 325.08	1.41	0.236

Lastly, recall that there were two types of filler items used in experiment: the under-informative type with partial, non-exhaustive answers (UA-type) and the over-informative type, in which the answers were not only exhaustive but also included information that was not required (OA-type). The data consisted of 690 judgments of both answer types from both groups. The mean values of both groups with 95% confidence interval are presented in Figure 8.

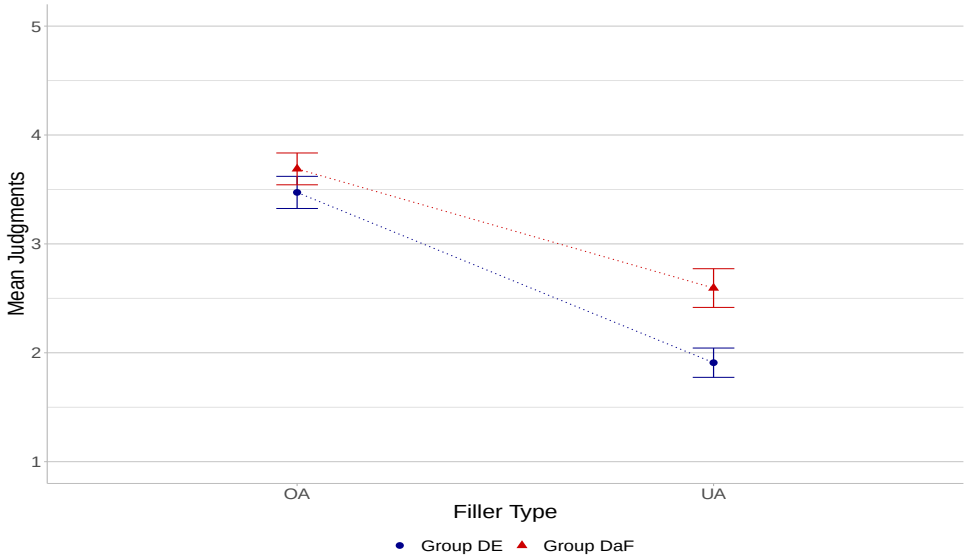


Figure 8: Mean judgments ± 95% confidence interval for filler items.

Similarly, the data was also analyzed by fitting a LMM using the *lme4* package from Bates, Mächler, et al. (2015) by maximum likelihood. In order to achieve the maximal random effects structure (cf. Barr et al. 2013), the random effects structure should consist of by-participant random intercepts and random slopes for Answer Type, and by-item

random intercepts with random slopes for the factor GROUP. This full model obtained a singular fit. However, the simplified model with reduced random effects following Bates, Kliegl, et al. (2015) does not significantly differ from the full model. Therefore, the full model was used for analyses[40], with 62.78 observations per estimated term. Analyzing the full model with the *emmeans* package (Lenth 2021), the paired contrasts show that the group DaF's deviation from the group DE (i.e. DE − DaF) was significant with respect to the under-informative, partial answers ($\Delta = -0.69$, $t(64.9) = -3.22, p = 0.002$), but not to the over-informative answers ($\Delta = -0.22$, $t(63.2) = -1.09, p = 0.280$).

8.3 Discussion

Based on the data and the statistical analysis, the research questions in (120), repeated below in (126), can be discussed.

(126) a. Can the Chinese DaF-learners also notice that presuppositions generated by soft and hard triggers differ in their ability to express an at-issue content in German, extending Hypothesis 1 in (117) to L2?

b. If the DaF-learners can notice the difference, do they react just as sensitive as the German native speakers to the at-issueness violation? If not, are nonnative speakers less sensitive but more tolerant towards the violations, as Hypothesis 2 in (119) assumes? Is the nonnative speakers' deviation from the native speakers stable for each trigger type, or are there differences between triggers?

c. If there are some differences between both groups, how can these differences be explained? Can the differences shed light on issues like trigger classification?

Let us start with the question in (126a): the paired contrasts prove that the DaF-learners are also able to perceive the soft-hard distinction and at-issueness violations in German. Just like the native speakers, their judgments also confirm that in comparison to hard triggers and NRRCs, the presuppositions of soft triggers can be easily locally accommodated and therefore more properly or naturally used to target the QUD – although still worse than assertions. In short, the Hypothesis 1 in (117) is confirmed by the nonnative speakers' data, too.

Addressing the questions in (126b), the significant interaction of GROUP and ISSUENESS shows that once the projective contents are at-issue, native and nonnative speakers react differently to this pragmatic violation. In comparison to the native speakers, the DaF-Learners, although competent to recognize the infelicity in their L2, are still less sensitive but more tolerant towards at-issueness violations, which is in line with Hypothesis

[40] Y ∼ ANSWER TYPE*GROUP + (1+ANSWER TYPE | Participant) + (1+GROUP | Item).

2 in (119). Moreover, the three-way interaction provides an answer to the last question in (126b), that is, regarding the pragmatic infelicity caused by at-issue presuppositions and NRRCs, the nonnative speakers' deviation from the native speakers in terms of sensitivity varies between trigger types. According to the results in Table 8, a significant interaction of Group and Issueness can be mainly observed with respect to the presuppositions of hard triggers. In other words, it is the hard trigger which primarily illustrates the nonnative speakers' insensitivity or tolerance towards at-issueness violations.

Another important observation in this comparison is that the Chinese DaF-learners are also able to recognize the infelicity caused by at-issue NRRCs, just like the native German speakers. This similarity is highly interesting because in the last 20 years, most theoretical and empirical research on relative clauses in Chinese either claim that there are only restrictive relative clauses (see e.g. N. Zhang 2001; Del Gobbo 2005) or, although there are relative clauses which can be interpreted as non-restrictive, syntactically they still differ from the NRRCs in European languages like English or Italian (see e.g. Del Gobbo 2010; Lin and Tsai 2015; Del Gobbo 2017). Del Gobbo (2017) for instance classifies non-restrictive clauses in Chinese as fully integrated as both the relative pronoun and an intonational break are absent, whereas they have to be present in non-integrate NRRCs – such as the NRRC items used in this experiment. No matter whether there are no NRRCs or only integrated NRRCs in Chinese, what is relevant is that the similarity observed in this experiment clearly shows that the Chinese DaF-learners have already acquired and understood the grammar of the non-integrate NRRCs in German, together with their not-at-issueness.

With this similarity between both groups in their judgments of NRRCs in mind, the difference in their judgments on hard triggers can hardly be explained by the absence of adequate counterparts in Chinese. That is, if we assume that the difference is due to the reason that the counterparts of hard triggers in Chinese have some potentially different syntactical or semantic properties which can interfere with the nonnative speakers' judgments, then for the NRRCs, which clearly differ from or even have no counterpart in Chinese, the difference between both groups has to be more significant. However, this is not the case, indicating that the corresponding hard triggers in Chinese can hardly cause the difference.

Now, turning to the questions in (126c): how can it be explained that nonnative speakers are clearly less sensitive or more tolerant towards the violation of Not-At-Issueness Constraint? Furthermore, why is this tolerance dependent on triggers, especially, why is it significantly more observable with respect to hard triggers? Can this sensitivity or the observations expose some clues to the properties of presuppositions in general?

The first possible explanation might be that the difference between both groups is due to the participants' native languages, based on the assumption that the speakers of certain languages are in general more tolerant towards pragmatic violation, or less sensitive to the boundaries between foregrounded and backgrounded information. If so, the violation of backgroundness in these languages can be considered less serious or pragmatically less infelicitous, and projective contents like presuppositions might also be understood as less not-at-issue – in other words, the Not-At-Issueness Constraint in (116) is not universal but rather language-dependently restrictive. If so, then native speakers of these languages might adapt their insensitive intuition of L1 to L2, and they are therefore more tolerant towards at-issueness violations.

However, this assumption can hardly be very persuasive. On the one hand, although there is few empirical cross-linguistic research on at-issueness, some experiments investigate properties that are related. Amaral and Cummins (2015) for instance observe that the behaviors of presuppositions in Spanish are parallel with that of those in English, which has been examined in their previous study (Cummins et al. 2013). In both experiments, participants were asked to judge four different types of response to polar questions carrying presuppositions: responses with 'yes' or 'no' and with or without a denial of the presupposition. In their experiments, the native speakers of English and Spanish clearly preferred the responses that address the assertion without denying or targeting the presupposition (see the review in Section 5.3 for more details). This preference confirms that in both English and Spanish, presuppositions are less suitable or less possible to be addressed in conversation, in other words, less at-issue than the assertions.[41] Moreover, Tonhauser (2012) applies six diagnostics for at-issueness, which are mainly based on data from European languages, to Paraguayan Guarani, a South American language. Results show a cross-linguistic validity of the diagnostics for projective contents, including presuppositions and NRRCs, and cross-linguistically support the correlation between not-at-issueness and projection, proposed by Simons et al. (2010) based on data in English. Therefore, we are allowed to assume the not-at-issueness of presuppositions to be cross-linguistic.

On the other hand, although the DaF group judged less strictly with regard to all three kinds of triggers once they were at-issue, the interaction of Group and Issueness is mainly significant with respect to hard triggers. If the Chinese native speakers are in general less sensitive to the difference between foregrounded and backgrounded information, this interaction would at least be significant with respect to the NRRCs, too. Additionally, as shown in the next chapter, the third experiment proves that the Chi-

[41] Note however, as mentioned before, that Snider (2017a,b) claims that a potential weakness of these studies is that the probability or possibility for a direct denial does not illustrate the backgoundedness but rather the anaphoric availability.

nese and German native speakers share a similar intuition about projection behavior of presupposition triggers in their L1. Considering that not-at-issueness and projection are correlated (Tonhauser et al. 2018), assuming that the difference between both groups is purely due to the mother tongue of the DaF group seems rather questionable.

The second possible explanation is that the pragmatic skills or intuitions cannot be completely adopted and then automatically used in a foreign language, but need to be acquired again, which is in line with Carrell (1984). Recall that in this experiment, all DaF participants have a German language level between B2 and C1, which can be understood as advanced and competent enough to understand some complex texts but not as good as the native speakers (see Council of Europe 2001 for more details about the levels for L2 acquisition). Interestingly, this level can be used to describe their pragmatic ability as well: their reaction to pragmatic violation shows that they are able to recognize the infelicity caused by at-issue presuppositions, but not as sensitively as the control group. Additionally, this assumption can also explain why there is no significant deviation with respect to the NRRCs: although there are no German-like NRRCs in Chinese, the DaF-learners have acquired this structure particularly in their foreign language and have understood their pragmatic properties, too.

Still, one problem is that all presupposition triggers used in the test are very basic vocabulary words and most of them are taught already in the primary, A-level courses. According to the word lists given by Goethe institute (see e.g. Perlmann-Balme 2004; Glaboniat et al. 2013; Perlmann-Balme 2015), triggers like *'auch' (too)*, *'gewinnen' (win)* are acquired very early at the A1-level, followed by *'wieder' (again)* and *'schaffen' (manage)* at A2. Only the word *'entdecken' (discover)* is considered as more advanced and belongs to the B1-vocabulary.[42] A question that needs to be answered then is what is missing so that the B2/C1-learners do not acquire the pragmatic usage of the triggers that they have already learned before? In other words, why is the pragmatic understanding or properties of a word not acquired together with its semantic meaning and syntactic structure?

In order to answer this question, a visualization of the judgments on each single trigger can be very helpful. Note that due to the reason that there are too few items per trigger word, the contrasts in Figure 9 are not able to generate an inferential statistic conclusion.

Surprisingly, as shown in Figure 9, it is precisely the A1-level words for which the DaF-learners' deviation from the native speakers seems, at least numerically, extremely obvious: *'auch' (too)* among the hard triggers, and *'gewinnen' (win)* among the soft triggers. Besides the clear insensitivity to at-issueness violations, the DaF-learners' judgments on natural

[42] I did not find out to which level clefts belong, but according to my experience, the cleft is taught much later than *'wieder'* in DaF courses.

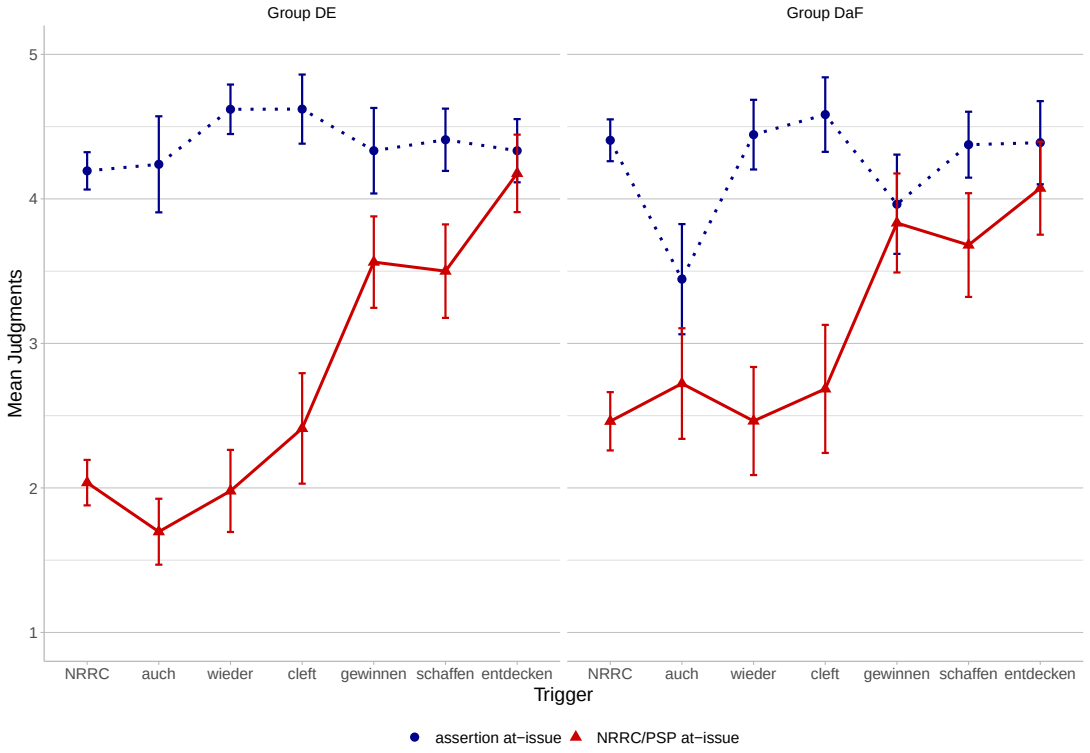


Figure 9: Mean judgments ± 95% confidence interval for the factors GROUP, ISSUENESS, and each single trigger.

answers (i.e. where the assertion targets the question directly) also differ sharply from those of the native speakers with respect to these two triggers. After having excluded the influence of their counterparts in Chinese, as discussed before, I have to admit that I have no explanation for such observation and can only leave this issue for future research.

Another observation that can hardly be explained by the acquisition assumption – at least not solely – is the ratings for the filler items. Recall that on the one hand, in the over-informative filler items, Peter's assertion answers the question exhaustively, although with additional information that is not required. With respect to this answer type, there is no significant contrast between both groups, showing that the DaF-learners and the native speakers share a similar intuition. On the other hand, with respect to the under-informative fillers, in which Peter's assertion does not offer an exhaustive answer, the nonnative speakers show their tolerance again. It does not seem very reasonable to assume that in L2, over-informativeness is acquired before under-informativeness. On the contrary, this contrast rather supports the idea that more perspectives are needed and more effects should be considered.

The third possible explanation assumes nonnative speakers' tolerance towards pragmatic issues in general, that is, all speakers are less sensitive to pragmatic violations in foreign languages, regardless of which language is their mother tongue, which language they learn as L2, and which pragmatic inference they judge. More specified, the less

sensitive reaction of nonnative speakers is not primarily due to the reason that they lack certain pragmatic skills or knowledge, or need to learn them in the foreign language again, but mainly due to the reason that nonnative speakers are generally more tolerant towards pragmatic violations, similar to children's pragmatic tolerance in first language acquisition, proposed by Katsos and Bishop (2011). This assumption can thus combine the observations in this study on presuppositions with those on scalar implicatures in my previous study (Y. Chen 2019) with a general principle – at least possibly.

This pragmatic tolerance of nonnative speakers may be in line with the observation in the field of psychology that we are more logical or less confident when we use or think with a foreign language. For instance, Costa et al. (2014) show in their experiment that facing moral dilemmas, participants are more likely to choose the more logical or utilitarian option if the question is given in a foreign language. Moreover, Geipel et al. (2015) propose that the use of a foreign language can reduce severe judgments and increase uncertainty. Note that under the condition NRRCs/PSP at-issue, the answer to the question still offers the information that is required by the question, although with an inappropriate information status. Therefore, it is possible to explain the difference with the preference of logic interpretation or the reduction of severe judgments in foreign languages. Furthermore, these two trends might be related to limited working resources, like Chierchia et al. (2001, 2004) assumed to explain children's insensitivity to pragmatic inappropriateness: similar to children in L1 acquisition, the adults also have to assign more working memory or cognitive resources to interpret pragmatic inferences in L2, which may reduce their pragmatic sensitivity.

This assumption seems to fail to explain why pragmatic tolerance is trigger-dependent itself: the critical point here is that it was the hard triggers that received the most obvious group difference instead of the soft ones, which are classically described as more pragmatic or conversational. However, the non-native speakers' pragmatic tolerance should still be considered a relevant part of the whole mechanism that influences speaker intuition in L2, as it offers us more possibilities to investigate pragmatic enrichment and is sometimes critical for explaining the observation made. For instance, by assuming the non-native speakers' tolerance to play a role , we may conclude that the observations in this experiment might offer some relevant clues to certain properties of presuppositions.

Firstly, this result can be understood as empirical evidence against the assumption that some presuppositions are conversational or scalar implicatures. Recall that in my previous experiments (Y. Chen 2019), a similar tolerance among nonnative speakers has been observed for typical scalar terms like *<all, some>* and *<must, it can be>*, but disappeared with respect to numerals, which are considered rather semantic (see also Y. T. Huang et al. 2013). As the pragmatic violation and the experiment design are different

from Y. Chen (2019), I do not think that the results of this experiment can be interpreted as evidence for presuppositions of hard triggers being pragmatic and those of soft triggers being semantic. However, the absent nonnative speakers' tolerance with respect to soft triggers should be a convincing evidence against the idea that the mechanism behind (some) presuppositions and scalar implicatures is the same. This observation is in line with Bill et al. (2016, 2018), and seriously challenges the uniform hypothesis for these two inferences (e.g. Chemla 2008; Romoli 2014).

Secondly, the results also contribute to the discussion on the similarity and difference between NRRCs and presuppositions. Although presuppositions in certain cases can be considered as new information (see e.g. Abbott 2000; Tonhauser 2015 for more discussion about *informative presuppositions*), the observation in this experiment offers additional evidence supporting the distinction between NRRCs and presuppositions: while NRRCs and hard triggers share a similar not-at-issueness and projection behavior, the nonnative speakers' reaction shows that these pragmatic properties could be acquired and/or processed differently. Thus, NRRCs should be classified as another subset of projective/not-at-issue contents, as most theories predict, rather than a special kind of hard presupposition triggers.

In sum, the results of the experiment show that 1) nonnative speakers are able to recognize at-issueness violations and the soft-hard split of presupposition triggers in terms of their different potential to express at-issue contents. 2) Still, in comparison to the native speakers, the nonnative speakers are less sensitive to at-issueness violations, and their insensitivity is especially obvious with respect to hard triggers. 3) The insensitivity of nonnative speakers might be explained by the acquisition of pragmatic skills in L2 and nonnative speakers' pragmatic tolerance in general, which can shed light on the boundary between presuppositions and NRRCs on the one hand, and between presuppositions and scalar implicatures on the other.

8.4 Chapter Conclusion

Adopting the experiment design from the first experiment, in this chapter, the soft-hard split of presupposition triggers has been investigated by comparing Chinese learners of German with German native speakers, with NRRCs as a baseline. Results show that both groups perceived that the presupposition of soft triggers are clearly more suitable to express at-issue content than NRRCs and hard triggers, confirming Hypothesis 1 in (117). Moreover, regarding the violation of the Not-At-Issueness Constraint in (116), the nonnative speakers are clearly less sensitive or more tolerant than the native speakers, confirming Hypothesis 2 in (119). Possible explanations for these findings might be the acquisition of pragmatic competence in L2 and/or nonnative speakers' pragmatic toler-

ance. Last but not least, such tolerance of the nonnative speakers or their deviation from the native speakers is mainly significant with regard to the hard triggers, but not to the soft triggers or NRRCs. This observation might give us some important clues to issues like the difference between soft triggers and scalar implicatures on the one hand, and NRRCs and hard triggers on the other.

9 Exp.III: Comparison of Triggers in German and Chinese[43]

9.1 Introduction

Regarding the three main research concerns and the five main research questions listed in (85), the third experiment of my study mainly focuses on the second question of applying the classification by using certain typical triggers for measuring and sorting, and the fifth question concerning the cross-linguistic stability of presuppositions, repeated below in (127):

(127) a. Research question 2: applying the classification: choosing certain typical presuppositions or triggers as markers for certain properties, can they function like a thermometer that measures and sorts other triggers into certain subclasses?

b. Research question 5: cross-linguistic comparison of presuppositions: can a presupposition be triggered steadily across languages, as long as an adequate and equivalent translation of its trigger can be found? Are their properties like projection behavior or information status independent from linguistic and cultural influence?

Regarding these two research questions, this experiment aims to measure the projection strength of factive verbs with typical hard and soft triggers as anchors, verifying the distinction between cognitive and emotive factive predicates. Moreover, this test will be performed in two very different languages, German and Chinese, in order to test the cross-linguistic stability of the projection behavior and the variety of projection strength. As the most relevant theoretical and empirical research background for these two questions has already been extensively reviewed before in Chapter 3 and 5, in this introduction, they will be only shortly and briefly summarized. We will start with the heterogeneity within the class of factive predicates, especially the split between cognitive and emotive factives with a focus on their projection strength, and then come to the classical soft-hard split of presupposition triggers and the projection of presuppositions across languages.

[43]The experiment in this chapter is joint work with Mailin Antomo. A paper partly based on it has been accepted for publication in the Journal *Intercultural Pragmatics*, see Y. Chen and Antomo to appear.

Let us start with factive predicates, which were first listed and investigated by P. Kiparsky and C. Kiparsky (1970) and have exceedingly extended the presupposition trigger list and the discussion on presuppositions (for more background, see e.g. Karttunen 2016). According to Hooper and Thompson (1973), predicates with *that*-complements can be classified into five groups, as shown in Table 9. Among them, only group D and E are relevant for this study; therefore the first three groups of verbs, which cannot presuppose their complement clauses as fact, will not be further discussed here.

Table 9: Hooper and Thompson's (1973) classification of *that*-S verb.

non-factive	A	say, report, explain,...
	B	suppose, believe, think,...
	C	be (un)likely, be (im)possible, doubt,...
factive	D	resent, regret, be sorry, be surprised, bother, be strange, be odd, be interesting
	E	realize, learn, find out, discover, know, see, recognize

As shown in Table 9, the factive predicates are split into two groups: emotive factive, such as *regret*, *be sorry* and *bother*, and cognitive factive, such as *discover* and *realize*. As summarized in Djärv et al. (2018), while predicates from the former sub-class express the subjects' emotional attitude to the facts expressed in the complements, verbs from the later one can illustrate the subjects' cognitive manner towards or knowledge status of the complements. In terms of presuppositions, predicates from both groups D and E can presuppose their *that*-complement as a fact. However, their complements have different projection strength, which can be illustrated once the sentences are embedded under question or conditionals, as observed by Karttunen (1971b), or even by all four operators from the family of sentences (Chierchia and McConnell-Ginet 2000) as claimed by Karttunen (2016).

According to Karttunen (1971b), the complement clauses of cognitive factive verbs like *discover* might lose their presupposed status once the main clause is questioned or conditional, while the complement of verbs like *regret* remains projective, as shown below in examples (128) and (129). When the sentences are embedded under question, Karttunen (1971b) claims that although both sentences (128a) and (128b) can be interpreted with the presupposition that the addressee did not tell the truth, with (128b), another interpretation is also possible – that is, the speaker asks the question as s/he is not certain about the truth of the complement clause and sincerely requires information. Such an interpretation is hardly possible for (128a). Similarly, by expressing (129a), the speaker admits that s/he has not told the truth, i.e., the complement is still presupposed and projects the conditional (for more discussion of these two sentences, see also e.g Abrusán 2011). In (129b) on the contrary, the complement is not certainly presupposed anymore,

as a local reading without the presupposition becomes quite natural. In other words, comparing (129b) with (129a), and (128b) with (128a), it can be seen that although both *discover* and *regret* are classified as factive verbs, the projection strength of their complements is very different.

(128) a. Did you regret that you had not told the truth?
b. Did you discover that you had not told the truth?
(Karttunen 1971b, p. 63)

(129) a. If I regret later that I have not told the truth, I will confess it to everyone.
b. If I discover later that I have not told the truth, I will confess it to everyone.
(Karttunen 1971b, p. 64)

Based on their different projection strength, traditionally, only the emotive predicates are classified as factive, while the cognitive predicates are classified rather as semi-factive (see e.g. Karttunen 1971b; Hooper and Thompson 1973; Stalnaker 1977).

However, about a half century later, Karttunen (2016) offers a different classification of factive verbs. Regarding the heterogeneity of factive predicates, they are sorted into five groups instead of a emotive-cognitive contrast:

(130) classification of factive verbs in Karttunen (2016, p. 711)
a. Certain predicates with that-clause subjects:
that S be odd/tragic (as opposed to likely)
that S count/matter/suffice (as opposed to happen)
b. Certain emotive adjectives with complements:
NP be happy/glad/furious that S (as opposed to hopeful)
NP be sad/delighted/disappointed to VP (as opposed to willing)
c. Certain propositional attitude verbs:
NP know/regret/forget/remember that S (as opposed to believe)
d. Verbs of discovery:
NP discover/find out/notice/observe that S (as opposed to suspect)
NP be discovered/found out/noticed/observed to VP (as opposed to suspected)
e. Certain verbs of communication:
acknowledge, admit, confess (as opposed to say)

In contrast to the classification based on Hooper and Thompson (1973), Karttunen (2016) claims that *regret* should pattern with propositional attitude verbs like *know* instead of emotive predicates like *be sad/happy*, and assumes a separate class for verbs of discovery, including *discover*. Moreover, Karttunen (2016) proposes that discovery verbs are

not presupposition triggers at all as "negative sentences, questions, and conditionals with coming-to-know verbs are in principle noncommittal" (p.713), and therefore clearly distinguishes them from verbs like *regret*.

In sum, the classification of factive verbs is still controversial in the theoretical works while the distinction between *regret* and *discover* in non-affirmative utterances is generally accepted. Both accounts, that is, the emotive-cognitive contrast and the five-group classification, predict that the presupposition of *regret* should be more stable and projective than the presuppositions of *discover*, if it can trigger one. In this study, we will use the concepts of emotive and cognitive factive verbs as it is widely accepted and the difference between them has also been experimentally investigated (see e.g. Egré 2008; Djärv et al. 2018; Schwarz et al. 2020).

The second key word of the study, the soft-hard split of presupposition triggers, has been extensively explained in Chapter 3; therefore, it will not be repeated here anymore. Interestingly, the difference between emotive and cognitive factive verbs is considered part of the varying projection strength between presuppositions of hard and soft triggers: traditionally, emotive factive verbs are considered as typical hard triggers, while congitive factive are assumed to be soft (see e.g. Simons 2001; Abbott 2006; Simons 2007; Abusch 2010, among others; for more discussion, see the review before in Chapter 3). Therefore, we are allowed to expect that the projection behavior of emotive factive verbs patterns with that of other hard triggers like *too* or clefts, while that of cognitive factive verbs patterns with that of soft triggers like *win*.

Experimentally, the projection strength can be measured by the certainty judgment task proposed and used in Tonhauser et al. (2018). In this test, presuppositions are considered as commitments of the speakers and the participants are asked to judge how certain a speaker is about a certain content of the utterance. One example given in Tonhauser et al. (2018, p. 502) is repeated below[44]. This design with question as operator and certainty as indicator is adopted in our experiment.

(131) Patrick asks: Was Martha's new car, a BMW, expensive?
'certain that' question (projectivity): Is Patrick certain that Martha's new car is a BMW?

Turning to the third key word of the study, the cross-linguistic comparison of presuppositions: as reviewed before in Chapter 5, there are only a few cross-linguistic studies on presuppositions. This might be due to the reason that for a long time, presuppositions are generally considered to be a universal phenomenon that occurs in all languages and less dependent on culture or language specific elements (see the review on theoretical

[44]Note that in their experiments, not only the projection strength of presuppositions but also that of other projective contents, like appositives in this example, is also tested.

predictions in Section 5.1). However, while Amaral and Cummins's (2015) experiments show the parallel behaviors of English and Spanish native speakers when rating presuppositions, both Schwarz et al. (2020) and Reins et al. (2021) observe that there may be differences between the ratings given by English and Italian, and English and Russian speakers, respectively (for more details, see the review in Section 5.3). Moreover, regarding other pragmatic features, the East-West divide, or the difference between Western and non-Western societies, has also been empirically observed (see the review in Section 5.2). Therefore, the universality of presuppositions needs to be experimentally investigated. Additionally, as the previous studies have mainly focused on European languages, in this experiment, a non-Western language with a very different cultural background is compared with German: Mandarin Chinese.

The first sub-experiment in this chapter is aimed at contributing to the universality issue of presuppositions with further empirical observations in German and Chinese. Regarding the difference between these two languages, the triggers used in the test were carefully chosen in order to avoid as many potential problems as possible. As the selection of triggers is extremely relevant and complex, it is explained separately in Section 9.2, which ends with the specialized research questions for the experiments. The empirical study is then presented in Section 9.3. As there is an alternative form of clefts in Chinese, the two options were shortly compared in the second sub-experiment, which is presented in Section 9.4. The observations and results are discussed in their sections and the summarized in then conclusion at the end.

9.2 Trigger Selection

Addressing the Research question 2 on applying the classification in (127a), the experiment has to test 4 triggers: a typical emotive and a typical cognitive factive predicate, together with a typical soft and a typical hard presupposition trigger that are able to function as anchors in the rating task. All 4 triggers need to be used commonly and frequently in both languages, and more importantly, their German and Chinese counterparts have to be both semantically and syntactically as parallel as possible.

As for the emotive factive predicate, the trigger *regret* was chosen. In German, this word can be translated as '*bereuen*' (see for instance Büring and Hartmann 2001; Antomo 2012; Ito 2017). In Chinese, the counterpart of *regret* is assumed to be '后悔 *(houhui)*' (see for instance Ji 2009; Xu 2014). Besides their correspondence in semantic meaning, syntactically, '*bereuen*' in German and '后悔 *(houhui)*' in Chinese can both take clausal complements and presuppose their content, as shown below:

(132) a. Ich bereue, dass ich den Kuchen gegessen habe.
I regret, that I the cake eaten have

'I regret that I ate the cake' (Ito 2017, p. 114)

⇝ I ate the cake.

b. 很长一段时间里，我都后悔自己那次选择站岗。

⇝ 我那次选择了站岗。(Xu 2014, p. 67)

hen chang yi duan shijian li, wo dou houhui ziji naci
very long one measure word time in, I always regret myself that time
xuanze zhangang.
choose stand guard.

'For a long time, I always regretted that I chose to stand guard that time.'

⇝ wo naci xuanze le zhangang.
⇝ I that time choose PST-particle stand guard.
⇝ 'I chose to stand guard that time.'

The verb *discover* was chosen to represent the cognitive factive verbs in the test. The first advantage is that *discover* is one of the most frequently discussed presupposition triggers (see e.g. Simons 2001, 2007; Abusch 2010; Abrusán 2011, 2016; Djärv et al. 2018, among others). Concerning the cognitive-emotive split, it is typical cognitive and concerning the soft-hard split, it is typical soft. If semi-factive/cognitive factive verbs are similar to other soft triggers, *discover* has to be the best candidate to illustrate this. The second advantage is that the counterparts of *discover* in German ('*entdecken*') and Chinese ('发现 *faxian*') are syntactically parallel – at least with regard to presuppositions. Both '*entdecken*' and '发现 *(faxian)*' can trigger presuppositions when taking clausal complements but are ambiguous when taking simple objects. Using negation as a test method, this property can be seen in examples below:

(133) a. Die Experten haben nicht entdeckt, dass ein Panda hinter
The experts have-PRS-3PL NEG discover-PPRT, that a panda behind
den Babmuspflanzen gesessen hat.
the bamboos sit-PPRT have-PRS-3SG.

'The experts did not discover that a panda sat behind the bamboos.'

b. 专家没有发现竹子后面坐着一只熊猫。

zhuanjia meiyou faxian zhuzi houmian zuozhe yi zhi
expert NEG discover bamboo behind sit a measure word
xiongmao.
panda.

'The experts did not discover that a panda sat behind the bamboos.'

⇝ A panda sat behind the bamboos.

(134) a. In dem Gebiet haben die Experten die Spuren von Pandas nicht
In the area have-PRS-3PL the experts the traces of pandas NEG
entdeckt.
discover-PPRT

'In the area, experts have not discovered traces of pandas.'

b. 专家在这片区域没有发现熊猫的痕迹。

zhuanjia zai zhe pian quyu meiyou faxian xiongmao de henji.
expert at this measure word area NEG discover panda of trace.
'In the area, experts have not discovered traces of pandas.'

?/↛ There are traces of pandas.

For the anchors in the test, the verb *win* was selected for soft triggers, as it is considered to be typical soft in several previous theoretical and empirical studies (see e.g. Abusch 2002, 2010; Bill et al. 2016, and the review in Section 3.1). Additionally, its meaning is clear and the syntax of the counterparts in German ('*gewinnen*') and Chinese ('赢 *ying*') can be parallel:

(135) a. Er hat das Spiel gewonnen.
He have-PRS-3SG the game win-PPRT.
'He won the game.'

b. 他赢了比赛。

ta ying le bisai.
he win PST-particle game.
'He won the game.'

⇝ He participated in the game.

As for hard triggers, the cleft structure was chosen due to three reasons: first, as all other triggers are lexical items, adding a structural trigger can be interesting for the cross-linguistic comparison. Second, other classical hard triggers like *again* or *too/also* were excluded due to their potentially non-parallel behaviors in German and Chinese. Take *again* for example: according to Xu (2014), the counterpart of *again* in Chinese is '又 *(you)*', a word with 6 meanings in the Contemporary Chinese Dictionary (现代汉语词典 Xiandai hanyu cidian, 6th Edition, The Commercial Press 2012), and not all of them can trigger presuppositions. Even when it expresses the meaning of repetition or iteration, the usage of '又 *(you)*' is still different from that of '*wieder*', the counterpart of *again* in German: as shown in Xu (2014), '又 *(you)*' can be used in all four contexts listed below, while '*wieder*' can only be used with the same subject and the same predicate, similar to *again* in English:

(136) a. Same subject, same predicate:
1849 年，英国的凯利最先制成一架滑翔机 [...]。1853 年，他**又**制成一架滑翔机。

1849 nian, yingguode kaili zuixian zhicheng yi jia
1849 year English Cayley first build a measure word
huaxiangji[...]. 1853 nian, ta **you** zhicheng yi jia huaxiangji.
glider[...]. 1853 year, he **you** build a measure word glider.

'In 1849, Cayley from England was the first to build a glider[...]. In 1853, he built a glider **again**.'

⇝ Cayley has built a glider before.

b. Different subjects, same predicate:
第一年姐姐考上了大学，[...]，第二年弟弟**又**考上了大学。

diyi nian jiejie kaoshang le daxue, [...], di'er nian
first year elder sister pass the exam PST-particle university, [...], second year
didi **you** kaoshang le daxue。
younger brother **you** pass the exam PST-particle university.

'The first year, the elder sister got into university, [...], the second year, the younger brother got into university, **too**.'

⇝ Someone has gotten into university before.

c. Same subject, different predicates:
他刚才表演了个魔术，现在**又**唱了一首歌。

ta gangcai biaoyan le ge moshu, xianzai **you**
he just perform PST-particle measure word magic trick, now **you**
chang le yi shou ge.
sing PST-particle a measure word song.

'He just performed a magic trick, and now he sings a song.'

⇝ He has done something silimar before.

d. Different subjects, different predicates:
早在三四十年代，美国就推广了杂交玉米，50 年代，墨西哥**又**出现了 [...] 杂交小麦。

zao zai sansishi niandai, meiguo jiu tuiguang le
early in three-four-ten period, USA already promote PST-particle
zajiao yumi, 50 niandai, moxige **you** chuxian le [...] zajiao
hybrid corn, 50 period, Mexico **you** emerge PST-particle [...] hybrid
xiaomai.
wheat.

'As early as the 1930s and 1940s, hybrid corn was already promoted in USA, and in the 1950s, hybrid wheat emerged in Mexico.'

⇝ Other hybrid crops have appeared elsewhere.

(All examples above are adopted from Xu 2014, pp. 97–99, sightly modified and then annotated.)

In short, while '*wieder*' in German mainly triggers the presupposition that the same subject has done the same thing before just like *again* in English, '**又** *(you)*' can trigger several different kinds of presuppositions, some of them even overlap with *too* and *also* in English. Therefore, *again* cannot be used as an example of a hard trigger anchor in the test.

There are similar problems with *too* and *also*. Besides the potential overlap with the trigger *again* in Chinese, i.e. '又 *(you)*', as shown above in (136b), there is also an ambiguity problem with the German counterpart. In German, both *too* and *also* are mainly translated as '*auch*'. Depending on whether '*auch*' refers to the subject or the object, a sentence can also become ambiguous with different presuppositions, especially in question form:

(137) Hat Elisa auch Kekse gegessen?
have-PRS-3SG Elisa too/also cookies eat-PPRT.
'Did Elisa eat cookies, too?'/'Did Elisa also eat cookies?'
'*auch*' refers to the object: Elisa also ate something else than cookies.
'*auch*' refers to the subject: Someone else than Elisa also ate cookies.

As both readings are natural and possible, and disambiguity via context seems difficult within our test design, *too* and *also* are also excluded.

Last but not least, it is generally assumed that the counterpart of the *it*-cleft is the '是… 的 *(shi...de)*'-structure in Chinese. This assumption can be traced back at least to Teng (1979) and C.-T. J. Huang (1982, 1988) and is still widely accepted nowadays (see e.g. Zhan and C. Sun 2013; Liu and Kempson 2018; Zhou 2020). Some examples can be seen in (138a) - (138d), in which the subject, time, place, and the action are focused respectively. However, '是 *(shi)*' can hardly be used to mark post-verbal phrases, such as the object in (138e) (for more discussion, see e.g C.-T. J. Huang 1982; Liu and Kempson 2018), while this is quite natural in German, as shown in (139b).

(138) a. **是**皮卡丘昨天在市中心看见了喵喵**的**。
Shi pikaqiu zuotian zai shizhongxin kanjian le miaomiao **de**.
shi Pikachu yesterday in city-center see PST-particle Meowth **de**.
'It was Pikachu who saw Meowth in the city center yesterday.'

b. 皮卡丘**是**昨天在市中心看见了喵喵**的**。
Pikaqiu **shi** zuotian zai shizhongxin kanjian le miaomiao **de**.
Pikachu **shi** yesterday in city-center see PST-particle Meowth **de**.
'It was yesterday when Pikachu saw Meowth in the city center.'

c. 皮卡丘昨天**是**在市中心看见了喵喵**的**。
Pikaqiu zuotian **shi** zai shizhongxin kanjian le miaomiao **de**.
Pikachu yesterday **shi** in city-center see PST-particle Meowth **de**.
'It was the city center where Pikachu saw Meowth yesterday.'

d. 皮卡丘昨天在市中心**是**看见了喵喵**的**。
Pikaqiu zuotian zai shizhongxin **shi** kanjian le miaomiao **de**.
Pikachu yesterday in city-center **shi** see PST-particle Meowth **de**.
'It was the case that Pikachu saw Meowth in the city center yesterday.'

e. # 皮卡丘昨天在市中心看见了**是**喵喵**的**。

Pikaqiu zuotian zai shizhongxin kanjian le **shi** miaomiao **de**.
Pikachu yesterday in city-center see PST-particle **shi** Meowth **de**.
'It was Meowth whom Pikachu saw in the city center yesterday.'

(139) a. Es war Peter, der Marie gestern im Stadtzentrum gesehen
It be-PST-3SG Peter, REL-NOM Marie yesterday in city-center see-PPRT
hat.
have-PRS-3SG
'It was Peter who saw Marie in the city center yesterday.'

b. Es war Peter, den Marie gestern im Stadtzentrum gesehen
It be-PST-3SG Peter, REL-ACC Marie yesterday in city-center see-PPRT
hat.
have-PRS-3SG
'It was Peter whom Marie saw in the city center yesterday.'

In contrast, the reversed structure '... 的是 *(...de shi)*' can easily be used to mark both subject and object but it is considered more similar to the pseudo-cleft (*wh*-cleft) or the headless relative clause in European languages (C.-T. J. Huang 1988; Cheng 2008):

(140) a. 昨天在市中心看见了喵喵**的是**皮卡丘。

Zuotian zai shizhongxin kanjian le miaomiao **de shi** pikaqiu.
Yesterday in city-center see PST-particle Meowth **de shi** Pikachu.
'Who saw Meowth in the city center yersterday was Pikachu.'/'The person who saw Meowth in the city center yesterday was Pikachu.'

b. 皮卡丘昨天在市中心看见了**的是**喵喵。

Pikaqiu zuotian zai shizhongxin kanjian le **de shi** miaomiao.
Pikachu Yesterday in city-center see PST-particle **de shi** Meowth.
'Whom Pikachu saw in the city center yesterday was Meowth.'/'The person Pikachu saw in the city center yesterday was Meowth.'

The problem now is: should the *it*-cleft in German (or '*es*'-cleft) then rather be compared with the '是... 的 *(shi...de)*'-structure, or with the '... 的是 *(...de shi)*'-structure?

According to C.-T. J. Huang (1988), the '是 *(shi)*' in *it*-cleft functions like an adverb, but in pseudo-cleft, '是 *(shi)*' is a copula. According to Zhan and C. Sun (2013), '是 *(shi)*' is a copula in both structures, but "it [pseudo-cleft] has the form [NOM COP NP], which differs from the cleft [NP COP NOM]" (p.773), in which NOM means nominalizated S/VP/NP with '的 *(de)*' at the end. According to Cheng (2008, p. 249), "the difference between them is simply a difference between canonical predication and inverse predication. Under this analysis, the inverse predicate cases are identical to specificational pseudoclefts in both structure and meaning". From this point of view, the '是... 的 *(shi...de)*'-structure is clearly more compatible with the *it*-cleft than the '... 的是 *(...de shi)*'-structure.

However, as is also reviewed in Cheng (2008), Merchant (1998) argues that there is no pseudo-cleft in Japanese: in English, wh-phrases like *who* cannot replace the focus or occur as the pivot in the matrix clause and then, by doing so, transform the pseudo-cleft into a question, but this is possible in seemingly pseudo-clefts in Japanese. Similarly, it is also suitable in the '*…* 的是 *(…de shi)*'-structure in Chinese, which might support that the '*…* 的是 *(…de shi)*'-structure is not a pseudo-cleft either:

(141) a. # Who/which director was who Ben met?

b. Jon-ga kubinisita-no-ga dare desu ka?
Jon-NOM fired-NW-NOM who is Q
'Who is it that Jon fired?'

c. Lisi zui xihuan de shi shei?
Lisi most like DE be who?
'Who is it that Lisi likes most?'

(Cheng 2008, pp. 247–248)

Additionally, most discussions so far are mainly concerned with '是*…* 的 *(shi…de)*' and '*…* 的是 *(…de shi)*' in declarative, affirmative sentences. Recall that in this experiment, the projection of presuppositions is tested with the question operator. Embedded under question, another issue needs to be considered: the '的 *(de)*'-omit.

According to Zhan and C. Sun (2013), the '的 *(de)*' in the '是*…* 的 *(shi…de)*'-structure is frequently omitted if the sentence ends with particles like the current related state particle '了 *(le)*', intonation particles '呢 *(ne)*', '吧 *(ba)*', '呀 *(ya)*' or question particles like '吗 *(ma)*'. In other words, a question like (142a) sounds more natural and is used more often than (142b):

(142) a. **是**皮卡丘昨天在市中心看见了喵喵吗？

Shi pikaqiu zuotian zai shizhongxin kanjian le miaomiao
Shi Pikachu yesterday in city-center see PST-particle Meowth
ma?
QST-particle.
'Was it Pikachu who saw Meowth in the city center yesterday?'

b. ?/# **是**皮卡丘昨天在市中心看见了喵喵**的**吗？

Shi Pikachu zuotian zai shizhongxin kanjian le miaomiao **de**
Shi Pikachu yesterday in city-center see PST-particle Meowth **de**
ma?
QST-particle.
'Was it Pikachu who saw Meowth in the city center yesterday?'

Then, a new problem occurs, namely that a sentence like (142a) cannot be distinguished from bare '是 *(shi)*' sentences. According to Paul and Whitman (2008, p. 420), "shi ... de patterns with cleft constructions, while bare '是 *(shi)*' behaves like association

with focus". Nevertheless, they also propose that sentences with initial bare '是 *(shi)*' with subject focus can be considered clefts.

Considering all these points mentioned above, there are two possibilities for the counterpart of *it*-cleft in Chinese: the first is using the '是… 的 *(shi…de)*'-structure but with initial '是 *(shi)*', omitted '的 *(de)*' and a focus on subject, so that regardless of whether these sentences are interpreted as '是… 的 *(shi…de)*' with implicit '的 *(de)*' or initial bare '是 *(shi)*' with subject focus, they are understood as clefts. The second is using the '… 的是 *(…de shi)*'-structure, so that the '的 *(de)*' occurs in the midfield and then the question particle '吗 *ma*' can occur at the end of the sentence. The two options are listed below:

(143) a. **'是… 的 *(shi…de)*'-strucuture: 是**伴娘给新娘送了一条珍珠项链吗？

Shi banniang gei Xinniang song le yi tiao
Shi maid of honor give bride present PST-particle a measure word
zhenzhu xianglian ma?
pearl necklace QST-particle.

'Was it the maid of honor who gave the bride a pearl necklace?'

b. **'… 的是 *(…de shi)*'-structure:** 给新娘送了一条珍珠项链**的是**伴娘吗？

Gei Xinniang song le yi tiao zhenzhu xianglian
give bride present PST-particle a measure word pearl necklace
de shi banniang ma?
de shi maid of honor QST-particle.

'Was it the maid of honor who gave the bride a pearl necklace?'

Regarding the advantages and disadvantages of both options, in the experiment, both structures were tested. That is, there were two groups of Chinese participants, one group who saw the '是… 的 *(shi…de)*' option and another group who was confronted with the '… 的是 *(…de shi)*' option, while all other items remained the same. As the '是… 的 *(shi…de)*'-structure is generally accepted and preferred as the counterpart of *it*-cleft, for the language comparison, the German group (group DE) was compared with the '是… 的 *(shi…de)*'-group (group CNS). After that, there was a separate comparison between the two Chinese groups (group CNS and CND) in order to check whether there are differences between the projection behavior of presuppositions triggered by these two structures.

In sum, the four selected triggers used in the experiment are listed below in Table 10 with their traditional classification and translations in both Chinese and German.

With these four triggers, the research concerns are specialized in the following particular research questions:

(144) a. Addressing Research question 2: combining the soft-hard split and emotive-cognitive split, can we observe that the emotive factive verb *regret* patterns

Table 10: The list of triggers used in Exp.III.

Trigger	Classification	Chinese	German
clefts	typical hard	是... 的 (shi...de)/ ... 的是 (...de shi)	es war
win	typical soft	赢 (ying)	gewinnen
regret	emotive factive, assumed to be hard	后悔 (houhui)	bereuen
discover	cognitive factive, assumed to be soft	发现 faxian	entdecken

with clefts and the cognitive factive verb *discover* patterns with *win*? Additionally, is there a significant difference between *regret* and *discover* as both the emotive-cognitive split and the five-group classification would predict?

b. Addressing Research question 5: when the translation is controlled as strict as possible, is the varying projection behavior of the presuppositions stable across German and Chinese, two very different languages?

c. Addressing the two cleft structures in Chinese: apart from their syntactical differences, are they both able to trigger presuppositions? If so, do their presuppositions receive similar projection strength?

9.3 Sub-Experiment a: Comparison of Triggers and Languages

9.3.1 Design

To address the research questions (144a) and (144b), we adopted the certainty judgment task from Tonhauser et al. (2018), measuring the projection strength by asking the participants whether the speaker is certain about some contents in her/his question. According to Tonhauser et al. (2018), speaker's certainty can estimate the speaker commitment more directly than other indicators[45]. Instead of a yes/no-rating, a 7-point certainty scale from 1 (absolutely uncertain) to 7 (absolutely certain) was used in our test. The paper-and-pencil questionnaires in both German and Chinese were built with OnExp (version 1.3.1) and collected online due to Covid-19.

[45]As summarized in Tonhauser et al. (2018), some other relevant indicators are the following: in Xue and Onea (2011) for instance, the assertion is embedded under *if* (the operator of conditionalization) and the projection strength is measured by asking whether it is possible that the presupposition does not hold. In Smith and Hall (2011), the assertions are embedded under negation or question, and the projection strength is evaluated by asking how surprising it is when the presupposition holds. We chose the certainty judgment task because of its stable performance and wide acceptance, for instance also in Tonhauser (2016) and Stevens et al. (2017)

With two parameters, the experiment had a 4×2 design: the first one was TRIGGER with four levels: *it*-clefts as typical hard, *win* as typical soft, *regret* and *discover* as typical emotive and cognitive factive, respectively. This parameter was within subjects but between items. The second one was LANGUAGE with two levels, German and Chinese, within items but between subjects.

To improve the understanding of the items and avoid confusion caused by item contents, we also designed a background story to introduce the context and all speakers that occur in the items, followed by an explanation of the rating task. This introduction into the questionnaire is translated into English and presented below:

(145) She, an athlete, and he, an elementary school teacher, are celebrating their wedding with numerous guests. After many games and performances in the afternoon and a big dinner, the bridal couple now wants to open the dance floor. Suddenly, police officers storm into the hall: a guest, Mr. Müller, has been found dead in the men's room. The guests are startled and whisper to each other.

You are an detective who is also there by accident. At your table, there are six other guests: Anna, Ben, Charlotte, David, Emma and Felix. You can clearly hear the quiet conversations between these six people. Anna, Ben, Charlotte, David, Emma and Felix don't know that you are a detective and are listening carefully, therefore they are all honest.

Based on the conversations, please try to conclude what level of knowledge the speaker has and rate it on a scale from 7 (= absolutely certain) to 1 (= absolutely uncertain). Please try to use the whole range of the scale for rating.

Last but not least, in order to check whether the participants paid attention and understood the task correctly, we also created filler and control items in the test. The ratings of a participant were excluded if s/he did not pass the attention check. We will explain more about the fillers/control items and the attention check in 9.3.2 and 9.3.5.

9.3.2 Materials

The materials used in our test consisted of three types of items: critical, control, and fillers. All items consisted of one question uttered by a speaker and a target question regarding the speaker's certainty about some contents, following Tonhauser et al. (2018).

Regarding the critical items, all of them had one of the four selected triggers in the uttered question, and the 'certain that' question addressed the presupposition of the utterance. For each trigger, 6 critical items were constructed, resulting in 24 critical items in total. Table 11 below shows one critical item per trigger in both German and Chinese with their English translation.

Table 11: Examples for critical items in Exp.III

trigger	Chinese	German	English translation
clefts	丁问道：是发型师女士在晚饭前和王先生在酒吧见了面吗？丁确定有人在晚饭前和王先生在酒吧见面吗？	David fragt: War es die Friseurin, die Herrn Müller kurz vor dem Abendessen an der Bar getroffen hat? Ist sich David sicher, dass jemand Herrn Müller kurz vor dem Abendessen an der Bar getroffen hat?	David asks: Was it the hairdresser who met Mr. Müller at the bar shortly before dinner? Is David certain that someone met Mr. Müller at the bar shortly before dinner?
win	丙问道：酒保先生下午赢了歌唱比赛吗？丙确定酒保先生下午参加了歌唱比赛吗？	Charlotte fragt: Hat der Barkeeper nachmittags den Gesangswettbewerb gewonnen? Ist sich Charlotte sicher, dass der Barkeeper nachmittags an dem Gesangswettbewerb teilgenommen hat?	Charlotte asks: Did the bartender win the singing competition in the afternoon? Is Charlotte certain that the bartender participated in the singing competition in the afternoon?
regret	甲问道：伴郎后悔他在茶歇期间吃了四块胡萝卜蛋糕了吗？甲确定伴郎在茶歇期间吃了四块胡萝卜蛋糕吗？	Anna fragt: Bereut der Trauzeuge, dass er während der Kaffeepause vier Stück Karottenkuchen gegessen hat? Ist sich Anna sicher, dass der Trauzeuge während der Kaffeepause vier Stück Karottenkuchen gegessen hat?	Anna asks: Does the best man regret that he ate four pieces of carrot cake during the coffee break? Is Anna certain that the best man ate four pieces of carrot cake during the coffee break?

Table 11 – *Continued from previous page*

trigger	Chinese	German	English translation
discover	已问道：录影师先生发现了新娘晚餐前换了发饰吗？ 已确定新娘晚餐前换了发饰吗？	Felix fragt: Hat der Kameramann entdeckt, dass die Braut vor dem Abendessen ihren Haarschmuck gewechselt hat? Ist sich Felix sicher, dass die Braut vor dem Abendessen ihren Haarschmuck gewechselt hat?	Felix asks: Did the cameraman discover that the bride changed her hair ornament before dinner? Is Felix certain that the bride changed her hair ornament before dinner?

Note that due to the inherent differences between the triggers, in the German version, the main clauses of *win* and *discover* were in perfect tense while the main clauses of *regret* were in present tense. The clefts were in simple past as the perfect form of the copula is usually replaced by simple past in German. As verbs do not have grammatical conjugation in Chinese, the items were either chronologically obvious or have a time phrase so that potential temporal ambiguity could also be avoided in the Chinese version.

Another difference between the two languages is that most job titles are gendered in German, for instance, a female hairdresser is a 'Friseurin', a male is a 'Friseur'. In contrast, the job titles in Chinese are mainly semantically gender neutral, but with potential gender preferences in a default imagination. Therefore, the gender information in German was also translated into the Chinese version by adding '女士 (nüshi)' (Ms.) or '先生 (xiangsheng)'(Mr.) after the job titles.

In the fillers and the control items, the target question always addressed the assertion of the uttered question. The difference between them was that the four selected presupposition triggers also occurred in the control items, but not in the fillers. So if the participants do not pay attention and simply rate all sentences with the four triggers as certain, or have a very different understanding of certainty, they can be detected and discarded. There were three control items for each trigger, thus 12 controls in the test. Additionally, there were also 12 fillers without triggers, resulting in 24 non-critical items in total. As the certainty rating for the fillers and control items should be rather on the uncertain half of the scale, the whole scale should be used and with a balance between its two halves in the experiment. Examples for control and filler items are presented below.

Table 12: Examples for control items and fillers in Exp.III.

item type	Chinese	German	English translation
control	丁问道：是发型师女士点了一份无乳糖饮食吗？	David fragt: War es die Friseurin, die ein laktosefreies Menü bestellt hat?	David asks: Was it the hairdresser who ordered a lactose-free menu?
	丁确定发型师女士点了一份无乳糖饮食吗	Ist sich David sicher, dass die Friseurin ein laktosefreies Menü bestellt hat?	Is David sure it was the hairdresser who ordered a lactose-free menu?
filler	丙问道：新娘的父亲典礼期间看了好多次手机吗？	Charlotte fragt: Hat der Vater der Braut während der Trauung mehrmals auf sein Handy geschaut?	Charlotte asks: Did the father of the bride look at his cell phone several times during the wedding ceremony?
	丙确定新娘的父亲典礼期间看了好多次手机吗？	Ist sich Charlotte sicher, dass der Vater der Braut während der Trauung mehrmals auf sein Handy geschaut hat?	Is Charlotte sure that the father of the bride looked at his cell phone several times during the wedding ceremony?

If a participant understands the task correctly and pays attention during the rating, ideally, her/his mean rating for each item type should look more or less similar to the diagram in Figure 10.

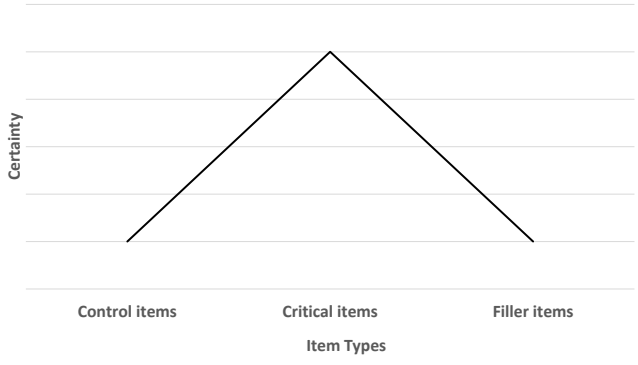


Figure 10: Attention Check: an expected diagram of the mean ratings per subject for each item type.

9.3.3 Procedure

The test links were sent online and after the participants had opened the questionnaire, they were firstly greeted and asked to fill in some personal data like their age, mother tongue and foreign languages. Then they read the introduction with the background story for all items and the explanation of the task, as shown in (145). In order to clarify what the participants would have to rate and how the certainty scale should be understood properly, two examples were presented with rating suggestions. As the critical items were concerned with presuppositions, the judgment on or rating for presuppositions was not allowed to be explained or implied in the introduction. Thus, we provided an appositive expression in the first example:

(146) a. Example:
Anna asks: Was the dead man, Mr. Müller, found in the men's room?
Is Anna certain that the dead man is Mr. Müller?
Rating suggestion:
In this example, it is appropriate to mark 7: The expression of this question is only meaningful if Anna is very certain that the dead man is Mr. Müller.

b. Example:
Anna asks: Is the dead man in the men's room Mr. Müller?
Is Anna certain that the dead man in the men's room is Mr. Müller?
Rating suggestion:
In this example, it is appropriate to mark 1: Anna asks the question exactly because she does not know whether the dead man in the men's room is Mr. Müller.

Before the beginning of the test session, it was emphasized that the participants did not have to value whether the utterance is grammatically appropriate or how likely the speaker can receive a *yes* or *no* as an answer. Instead, they were instructed to always focus on the certainty level of the speakers.

In the test session, the 48 items occurred randomly and the whole task took 15-20 minutes on average.

9.3.4 Predictions

Considering the predictions and observations of the previous theoretical and empirical studies (reviewed above and in Chapter 3 and 5), and the careful selection of triggers with a strict translation, we made the following predictions with a graphical representation in Figure 11.

(147) a. Addressing the research questions in (144a): the emotive factive verb *regret* patterns with clefts and they both are judged to express a high level of certainty, whereas the cognitive factive verb *discover* patterns with *win* and the speaker's certainty is rated comparatively lower. Additionally, significant differences between clefts and *win* on the one hand, and between *regret* and *discover* on the other are also expected.

b. Addressing the research question in (144b): the projection behavior of the presuppositions and the differences between them are stable across German and Chinese.

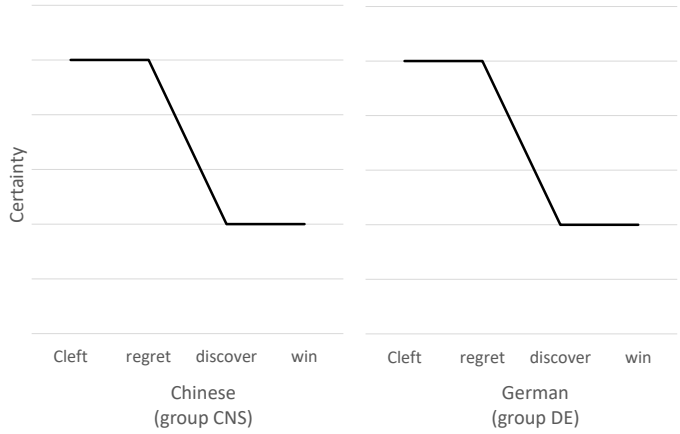


Figure 11: Predictions about observations.

9.3.5 Participants

As explained before in the introduction, the two groups compared in this experiment were German native speakers and Chinese native speakers who saw the '是... 的 *(shi...de)*' option of the cleft.

The Chinese group consisted of 51 participants; 6 of them did not pass the attention and understanding check (plot per subject see Figure 24 in Appendix A.3.4), leaving 45 participants (group CNS, mean age 22.91 ± 4.54) for the statistical analysis.

The German group consisted of 34 participants; 2 of them did not pass the attention and understanding check (plot per subject see Figure 25 in Appendix A.3.4). Therefore, 32 participants (group DE, mean age 27.45 ± 8.95) were included in the statistical analysis.

9.3.6 Results

For the analysis of the judgments on the critical items, all ratings for the control items and fillers were excluded. The data then encompassed 1848 judgments taken from 77

individuals out of 2 groups for 24 critical items. The mean judgments with 95% confidence interval are presented in Figure 12.

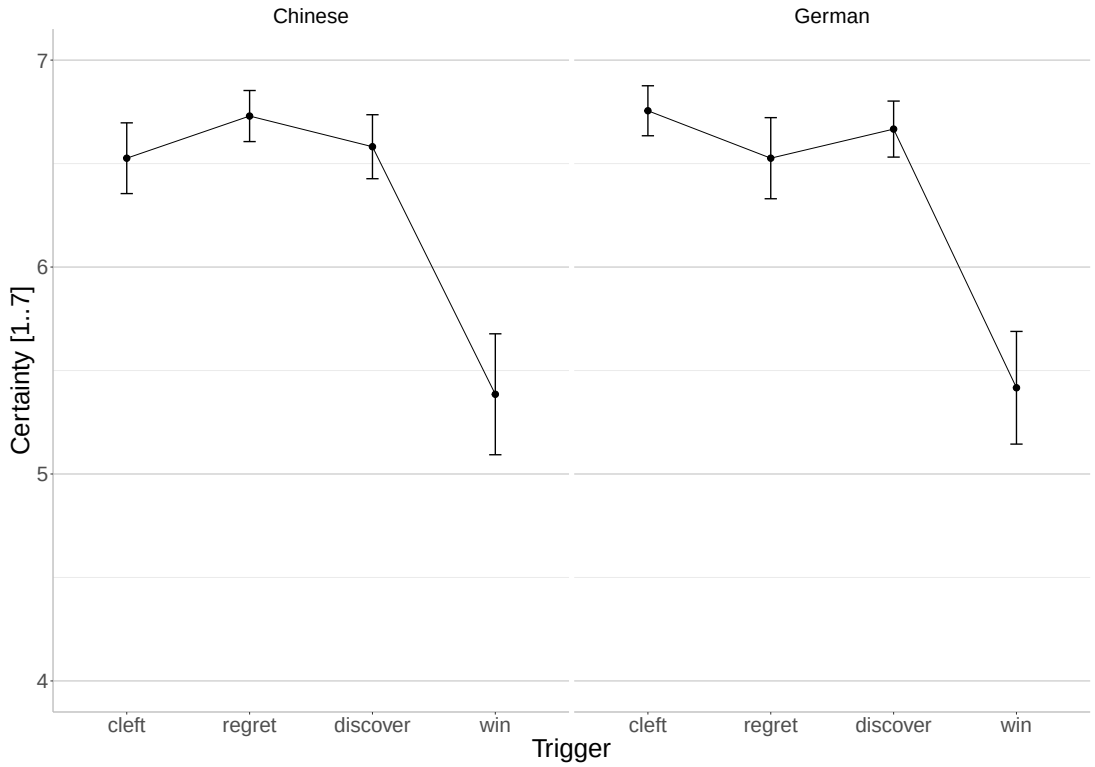


Figure 12: Mean judgments ± 95% confidence interval for the factors LANGUAGE and TRIGGER.

To estimate the effects of TRIGGER, LANGUAGE and their interactions, the data was analyzed using R (version 4.0.3, R Core Team 2020). Specifically, a linear mixed model (LMM) was fitted using the *lme4* package from Bates, Mächler, et al. (2015) by maximum likelihood. The factor LANGUAGE was centered with the *sum*-contrast, coded with (-0.5, 0.5), while the factor TRIGGER was dummy coded. In order to achieve the maximal random effects structure (cf. Barr et al. 2013), the random effects structure should consist of by-participant random intercepts and random slopes for TRIGGER, and by-item random intercepts with random slopes for the factor LANGUAGE. However, this model obtained a singular fit and failed to converge. Following Bates, Kliegl, et al. (2015), the model was too complex and the random structure needed to be reduced until the random-effect Principal Components Analysis (rePCA) did not report overidentification. Therefore, the final maximal random effects structure that neither caused converge failure nor led to a singular fit included only the by-participant random intercept.[46] The model had then 184.80 observations per estimated term.

In order to generate *p*-values from the linear mixed model, the fully specified model and restrictive models which left out one parameter at a time were compared by using

[46] Y ~ TRIGGER*LANGUAGE + (1|Participant).

the `mixed` function with Type 3 tests from the *afex* package (Singmann et al. 2020). An overview of the results can be found in Table 13.

Table 13: LMM overview with *p*-values based on likelihood ratio tests.

Model Parameter	df	χ^2	*p*-value
TRIGGER	3	234.75	<0.001
LANGUAGE	1	0.07	0.797
TRIGGER×LANGUAGE	3	5.37	0.147

As shown in Table 13, only the factor TRIGGER had a significant effect, indicating that the four different triggers clearly influence the certainty rating; in other words, their presuppositions differ in projection strength. Additionally, both the factor LANGUAGE and the interaction of TRIGGER and LANGUAGE had no significant effect. Considering the similarity between patterns, the difference between triggers can be seen as similar or parallel in both languages.

However, as the 7-points-scale we used in test can also be understood as a kind of ordinal scale on the one hand (for more discussion about ordinal data and metric models, see Liddell and Kruschke 2018), and the residual of this LMM fitting does not seem to be as ideally symmetric as required (see Figure 27 in Appendix A.3.4), the data was then double checked with the Cumulative Link Mixed Model (CLMM) using the `clmm` function from the *ordinal* package (Christensen 2019). Similar to the LMM model, the CLMM full model also included TRIGGER, LANGUAGE and their interaction as fixed effects and the maximal random effects structure[47] and used logit as the link function. The factor LANGUAGE was centered with the *sum*-contrast, coded with (-0.5, 0.5), while the factor TRIGGER was dummy coded. The results of the full CLMM model are listed in Table 14. As the estimates of effects and thresholds are on the log odds scale, the probabilities of ratings predicted by the model were also calculated and then visualized in Figure 13.

Regarding the main research question on cross-linguistic stability, the full model was firstly compared with the restrictive model which comprised only the fixed effect TRIGGER and the random effects[48], that is, the effect of LANGUAGE and the interaction were removed. The comparison between these two CLMM models was not significant (LR.stat=6.46, df=4, p=0.17), indicating that LANGUAGE and the interaction have on significant influence of the certainty rating, which is in line with the results given by the LMM comparisons above.

According to both LMM and CLMM analyses, TRIGGER is the only main effect that significantly influences the certainty rating value. Applying the CLMM full model to

[47] Y ∼ TRIGGER*LANGUAGE + (1+TRIGGER|Participant) + (1+LANGUAGE|Item).

[48] Y ∼ TRIGGER + (1+TRIGGER|Participant) + (1+LANGUAGE|Item)

Table 14: Overview of the CLMM full model, fitted with the Laplace approximation, with estimates, standard errors, z- and p-values of fixed effects and threshold coefficients.

Effect	Estimate	SE	z-value	p-value
TRIregret	0.48	0.38	1.28	0.20
TRIdiscover	0.78	0.47	1.67	0.09
TRIwin	-2.03	0.29	-6.92	<0.001
LAN.con1	-0.22	0.42	-0.52	0.60
TRIregret×LAN.con1	-0.93	0.53	-1.74	0.08
TRIdiscover×LAN.con1	-0.28	0.65	-0.43	0.67
TRIwin×LAN.con1	-0.48	0.50	-0.97	0.33

Threshold	Estimate	z-value
1|2	-4.44	-15.44
2|3	-4.14	-14.69
3|4	-3.89	-14.01
4|5	-3.52	-12.92
5|6	-3.17	-11.83
6|7	-2.23	-8.73

the `emmeans` function from the *emmeans* package (Lenth 2021), the four triggers were compared pairwise with each other within LANGUAGE using z test[49]. Results that are relevant to our research questions are listed in Table 15 below.

Table 15: Comparison of triggers based on the CLMM full model with z test, p-value adjusted with tukey method.

LANGUAGE	comparison of triggers	estimate	z-ratio	p-value
Chinese	cleft − win	1.79	4.58	<0.001
	regret − discover	0.02	0.03	1
	cleft − regret	-0.94	-1.88	0.23
	discover − win	2.72	4.47	<0.001
German	cleft − win	2.28	5.99	<0.001
	regret − discover	-0.63	-1.03	0.73
	cleft − regret	-0.02	-0.05	1
	discover − win	2.92	4.61	<0.001

According to the comparisons, the similarity and the difference between the individual triggers can also be considered stable across languages. In both Chinese and German, the difference between the typical hard trigger clefts and the typical soft trigger *win* is highly significant. Additionally, the similarity between the typical hard trigger clefts and

[49]The degrees of freedom were given as infinite (Inf) as label for the z test. According to the vignette of the *emmeans* package (Lenth 2021): "obtaining quantiles or probabilities from the t distribution with infinite degrees of freedom is the same as obtaining the corresponding values from the standard normal."

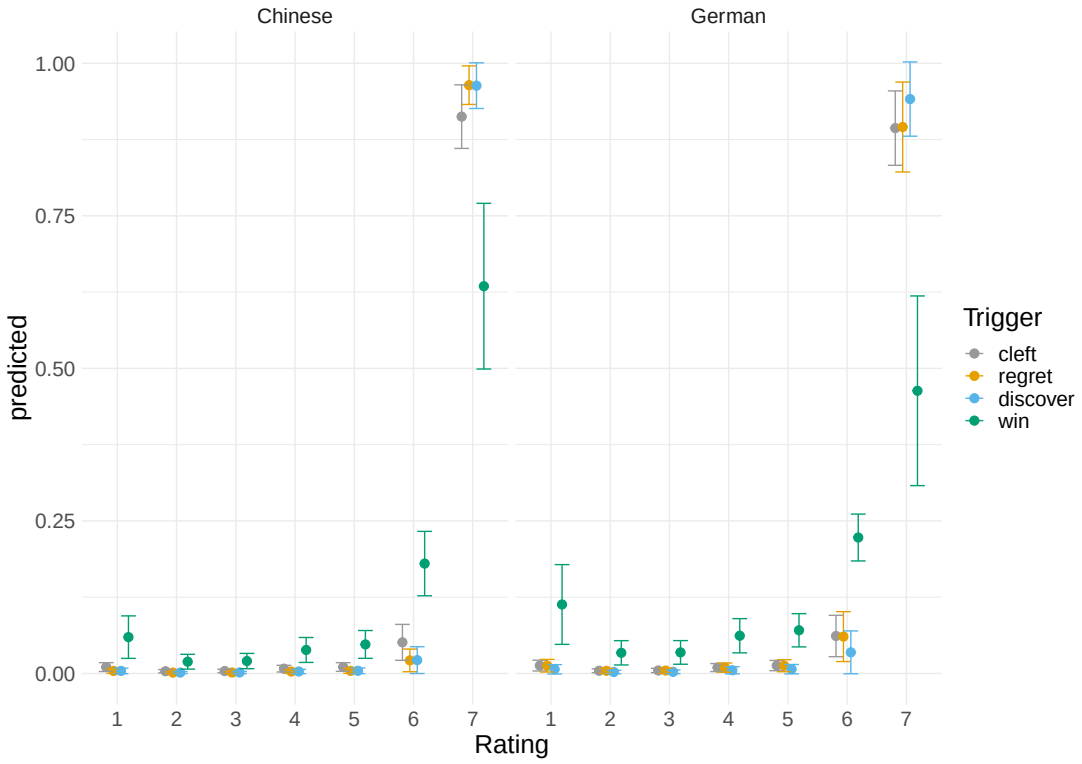


Figure 13: Probabilities of ratings predicted by the CLMM full model with 95% confidence interval.

the typical emotive trigger *regret* can also be clearly observed, as expected. However, in contrast to our predictions in (147a), the typical cognitive trigger *discover* does not pattern with the soft trigger *win* as most theories on the soft-hard classification propose. Quite on the contrary, its rating probabilities are similar to those of the typical hard trigger clefts and the emotive trigger *regret*, although a difference between them is expected.

In the *post-hoc* analysis, we investigated whether the difference between the two soft triggers is related to the understanding of the task. More precisely, we wanted to know whether the difference between the ratings for *discover* and for *win* ($\bar{x}_{discover} - \bar{x}_{win}$) is correlated to the difference between the critical and the non-critical items ($\bar{x}_{critical} - \bar{x}_{non-critical}$). As there is no significant difference between LANGUAGE, the 77 participants were considered as one group and their data are plotted in Figure 14. The `cor` function from R (version 4.0.3, R Core Team 2020) shows a negative correlation between these two differences with a r-value of -0.30. Using the `lm` function from R (version 4.0.3, R Core Team 2020), the effect of the task understanding on the difference between these two soft triggers was analyzed with the simple linear regression. According to the results, the difference between *discover* and *win* decreases by about a half point if the participants judge the critical items one point more certain than the non-critical items (estimate ± SE: -0.46 ± 0.17, t_{75}=-2.73, p=0.008). The linear regression with 95% confidence interval is presented in Figure 14, too.

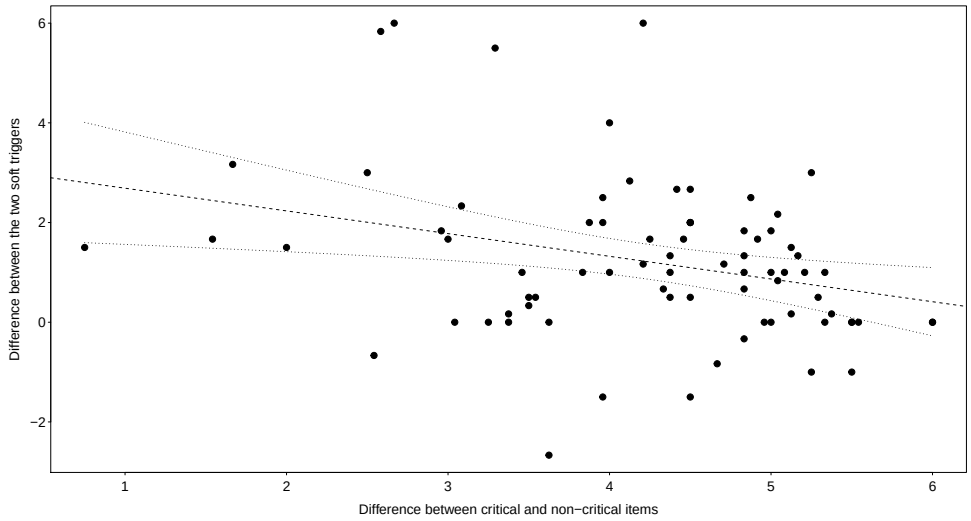


Figure 14: Results of the *post hoc* analysis with linear regression and 95% confidence interval.

9.3.7 Discussion

The Sub-Experiment a is designed to address the research questions (144a) and (144b), repeated below as (148a) and (148b):

(148) a. Addressing Research question 2: combining the soft-hard split and emotive-cognitive split, can we observe that the emotive factive verb *regret* patterns with clefts and the cognitive factive verb *discover* patterns with *win*? Additionally, is there a significant difference between *regret* and *discover* as both the emotive-cognitive split and the five-group classification would predict?

b. Addressing Research question 5: when the translation is controlled as strict as possible, is the varying projection behavior of the presuppositions stable across German and Chinese, two very different languages?

Regarding the first question, our observations are only partially in line with the predicitions in (147a): with regard to the projection strength of presuppositions, which was measured by the certainty of the speakers in this design, a significant difference between the typical hard trigger clefts and the typical soft trigger *win* was observed. Additionally, the emotive factive verb *regret* patterned with clefts and they were both judged as expressing a very high level of certainty, which distinguishes them from the soft trigger *win* and is in line with most assumptions made in the context of the traditional classifications (e.g. Simons 2001; Abbott 2006; Simons 2007; Abusch 2010).

However, in contrast to our predictions, the cognitive factive verb *discover* did not pattern with the soft trigger *win* but received rather high ratings of certainty. Additionally, the difference between *regret* and *discover*, which is expected by researchers like Karttunen (1971b), Hooper and Thompson (1973), and Karttunen (2016), was also absent in this experiment. According to the judgments given by the participants, the speakers were rated

being similarly certain about the information given in the complement clauses of *regret* and *discover*, indicating a similar projection strength of emotive and cognitive factive verbs. This finding is different from the observations in previous empirical studies, such as Djärv et al. (2018). Considering that the presuppositions of *discover* can lose their projectivity in some contexts, the following question arises: which elements in our design have increased the projection strength of *discover*?

For illustration, take the item given in Table 11, repeated below in (149). The question then can be specialized as follows: why is the reading as a sincere requirement of information about the bride's hair ornament, similar to the example (128b) suggested by Karttunen (1971b), not available, or how does the question become committal, thus different from what Karttunen (2016) proposes?

(149) Felix asks: Did the cameraman discover that the bride changed her hair ornament before dinner?

One might argue that the complement clause of the question (149) is about something, namely a change in the bride's hair ornament, which can be observed or noticed by some or even most of the guests at the party. Therefore, the truth of the embedded clause can be considered to be part of the common ground or committed via the context, and interpreting the whole sentence as a sincere question is hence not felicitous. However, there are two arguments against this assumption: first, there were also critical *discover*-items with complements about things that, if they can be seen as committed at all, should only be known by a very limited group of people or took place outside of the party, such as a missing make-up brush in item 1 or the classmate relation between the groom's father and Mr. Müller in item 6 (for the whole lists of items used in the experiment, see Appendix A.3.1). However, the ratings of these two items did not decline or differ from the other items, as shown in Figure 15. Second, there were also items with the other soft trigger *win* that described events that happened during the party and should also be widely known among the guests. Events like the singing competition in item 3, the quiz in item 5 and the tug-of-war in item 6 all took place in the afternoon at the wedding party. However, their ratings were not homogeneously higher than those of the other three items, which described competitions that took place in the last year: in fact, they were rated quite similarly, only the mean rating given by group DE for item 6 was a little higher, but it was still lower than most ratings given for items containing the other three triggers. So even if the absence or presence of the events in the hearer's mind could influence the reading, it should only be a minor effect.

Another possibility might have to do with the exciting background story of a wedding followed by the death of one guest. According to Simons (2007), a *discover*-sentence can lose its presupposition because it is the complement that offers the answer to the Question

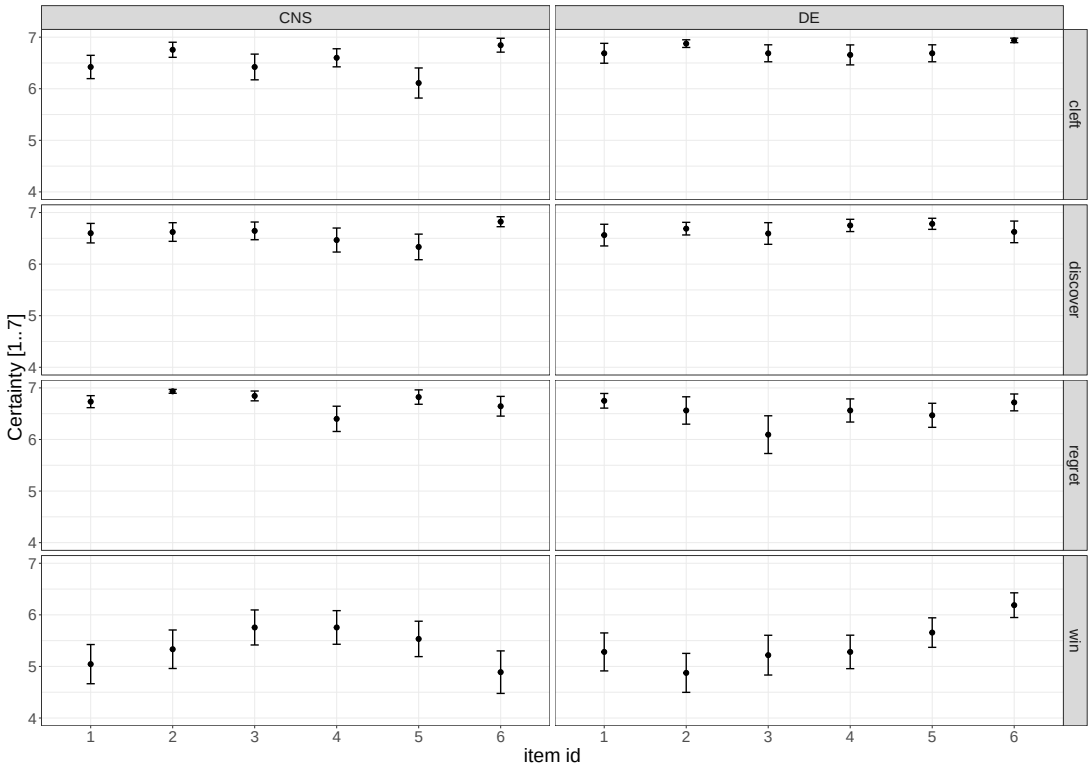


Figure 15: Mean ratings for each critical item used in the experiment, split by TRIGGER and GROUP.

Under Discussion (QUD), not the main clause. Therefore, the complement is at-issue and and cannot project, while the word *discover* only offers the source of information, called the evidential function of the main clause (see Simons 2007; Abusch 2010; Simons et al. 2010 for more discussion):

(150) A: Where was Harriet yesterday?
B: Henry discovered that she had a job interview at Princeton.
(Simons 2007, p. 1035)

In our test, on the contrary, with such a thrilling background story, expressing a question explicitly with the word *discover* can effectively boost the relevance of the cognitive status itself and thus that of the matrix clause. As for the question in (149), this means that the discovery status of the cameraman is marked as highly relevant to the QUD and the main clause becomes clearly at-issue, while other contents can be more or less pushed into the background.

Additionally, our background story may also boost the preference of a more informative interpretation of the whole sentence. Recall that the matrix sentence of (149) is about *whether discover* while the embedded clause is about *whether change*. While the two different readings might be similarly informative in declarative expressions, this is not the case with questions. As shown below, the two interpretation possibilities of the

question have different information quantities, and the one with the presupposition is more informative than the other, and possibly therefore preferred in this situation.

(151) Did the cameraman discover that the bride changed her hair ornament before dinner?
p: The cameraman discovered q.
q: The bride changed her hair ornament.
Interpretation with presupposition: certain q, uncertain p
Interpretation as a sincere question (Karttunen 1971b) with the evidential function of the main clause (Simons et al. 2010): uncertain q, uncertain p

A question that arises, then, is: why is the typical soft trigger *win* impervious to such a background story? As reviewed before in Chapter 3, in most previous theoretical analyses, the softness of *win* and *discover* is always explained with the same mechanism. According to the account based on the two-way split (see the review in Section 3.1), Abusch (2002, 2010) for instance claims that presuppositions of both semi-factive verbs and achievement verbs are triggered pragmatically from their alternatives.[50] Similarly, Romoli (2012, 2014) discusses the similarity between scalar implicatures and presuppositions of soft triggers based on the assumption that all soft triggers like *win, stop* and *discover* are similarly soft. In a similar manner, in the account focusing on the entailment relation between assertions and presuppositions (Klinedinst 2016; Zehr and Schwarz 2016, 2018, reviewed in Section 3.3), soft triggers are discussed as a homogeneous class without distinction between triggers as well. In sum, most existent research on presupposition classification cannot straightforwardly explain the difference between the two classical soft triggers that we observed in our test. Moreover, note that the question could be even more complicated, as the negative correlation observed in the *post-hoc* analysis should also be considered, or at least not be contradicted – that is, the difference between the certainty ratings for *discover* and for *win* would decrease, if participants distinguished presuppositions more clearly from assertions.

One explanation is that the difference between the two soft triggers is due to the events: in comparison, the competition results are perhaps less interesting or less relevant to Mr. Müller's death than the discovery status of some guests. Therefore, the assertion of *win* is not as at-issue as that of *discover*, and hence, its complement clause does not have to be interpreted as completely presupposed or as not-at-issue as that of *discover*. However, there are several events mentioned in other items that are not relevant to the

[50]Based on my research, among all literature on presupposition triggering and classification, *win* is one of the triggers that are generally accepted to be soft but are less explicitly analyzed. The most precise analysis is given by Abusch (2002) with the pragmatic triggering from alternatives, see (46), which was also applied to factive verbs like *know*. Whether this strategy can be improved in order to describe the difference within the sub-class of soft triggers could be a interesting theme, but it will not be further discussed here.

death either, for instance the best man's feeling about the four pieces of carrot cake in the *regret*-item, which can hardly be assumed to be directly related to Mr. Müller's death either. Nevertheless, these ratings were still significantly different from the *win*-items. In short, the possible lack of relations between the competitions and Mr. Müller's death should not be the main reason for the difference between *win* and *discover*.

A more plausible explanation is, then, that the soft triggers are not similarly soft, that is, the softness of *win* and *discover* is due to different reasons. While the latter is affected by our design, the former is not. Summarizing the predictions about *discover* given by Karttunen (1971b), Simons (2007), and Karttunen (2016), the softness of *discover* is due to the fact that the verb can receive more than one reading: the one with the presupposition that is clearly not-at-issue and hence projects, and some others which can be described as sincere questions, or non-committal expressions, or with a evidential function, so that they do not presuppose the complement. With our design, the latter option is less available due to the reasons discussed above, while the former option is preferred.

With regard to the verb *win* on the other hand, we are allowed to assume that its softness is due to some different reasons. One possibiliy among them might be that, in contrast to *discover*, the presupposition of the achievement verb *win* is part of the process it describes, so that the presupposition *participate* and the assertion *win* are temporally and logically extremely closely related and stored with strong connection in our mental lexicon. During the sentence comprehension, once *win* becomes relevant in the discourse and at-issue, our world knowledge might also activate *participate*. Therefore, assuming a continuum of at-issueness, *participate* as the presupposition of *win* is not-at-issue but not as strongly not-at-issue as the presuppositions of hard triggers like clefts or factive verbs. Thus, there is a significant difference between the ratings for *win* and for the other three triggers on the one hand, and on the other hand, the presuppositions of *win* should still be seen as projective, as the majority of the ratings were given above the middle point of the scale.

This possibility can also explain the negative correlation observed in the *post-hoc* test. Assuming that the clear difference in certainty ratings for critical and non-critical items indicates that the participant has a sensible pragmatic intuition and can distinguish sharply between presuppositions and assertions, we are also allowed to assume that s/he is more likely to achieve or access the non-presuppositional reading of the *discover*-questions, too, in spite of the background story.

There is an alternative explanation that needs to be discussed shortly here. One might argue that in our design, the complement of the target question and of the uttered question are always identical in the critical items with the two factive verbs. As a consequence,

participants were uncertain when confronted with the critical items containing *win* because of the difference between the two questions. However, we think this is rather unlikely. First, the target question in the critical items with clefts is also different from the uttered question, but the results show no uncertainty like in those items with *win*. Second, the complements of the target questions are more identical with the uttered questions in the filler and control items, but most participants still judged the speaker as being uncertain. However, we admit that the design can be improved, and the repetitions of complements could sometimes be annoying for our participants.

The second research question (148b) addresses the universality issue and focuses on the triggering and projection behavior of presuppositions in Chinese and German. According to our results, the projection behavior of presuppositions under the question operator can be considered stable between the two languages if the trigger is carefully selected, that is, if there are adequate counterparts of the trigger in both languages, and the translation of items is strictly controlled. This finding is in line with the theoretical predictions, in which a cross-linguistic stability of presupposition projection is implied or assumed (see the review in Section 5.1). Moreover, it is also consistent with the main findings in previous studies like Amaral and Cummins (2015) and Schwarz et al. (2020), who observe a similarity among some European languages. Additionally, the ratings illustrate that not only the triggering of presuppositions by certain expressions can be cross-linguistic, but also their projection strength and their interaction with contextual effects can be stable across languages as well.

Recall that Simons (2001) notices the nondetachability of presuppositions within one language, i.e. the observation that expressions with the same semantic content give rise to the same presupposition, as shown in (39b), repeated below again in (152):

(152) Jane didn't stop laughing/didn't quit laughing/didn't cease laughing.

$\rightsquigarrow$ Jane laughed before. (Simons 2001, p. 435, slightly modified)

Our results suggest that this observation can be extended to a cross-linguistic perspective: if a trigger has a semantically and syntactically equivalent expression (or construction as in the case of clefts) in another language, it leads to the same presupposition with the same projective behavior. For Simons (2001), this nondetachability is proof for a pragmatic source of the presupposition. However, in our opinion, this does not follow straightforwardly. Especially if we consider the observed crosslinguistically stable projective behavior of clefts, which have to be considered a hard trigger, it is not convincing that the presuppositions of these constructions in German and Chinese necessarily arise via pragmatic reasoning. Nevertheless, what we can say is that our results suggest that the triggering process, semantic or pragmatic, seems to be the same in both languages we investigated. It is more than likely that the presupposition triggered by an expres-

sion like *win*, *regret* or *discover* is closely related to the semantic content of its trigger. Thus, an expression in another language with the same semantic content should typically lead to the same presupposition. However, this line of reasoning cannot be applied to cleft-structures: without semantic content, the arising presupposition cannot be a result of this content. Instead, giving rise to a presupposition seems to be the raison d'Être of cleft-constructions, their specific function that distinguishes them from a mere main clause. As a consequence, we predict that if a language offers the possibility to use a cleft, this construction should give rise to a presupposition. We will leave the question on how presuppositions are triggered for future research – contrastive investigations, however, seem to be a promising way to shed light onto the triggering process.

However, note that a cross-linguistic or cross-cultural difference cannot be absolutely excluded. Firstly, the triggers used in this experiment are strictly selected. Thus, some potential differences or the absence of similarity or correspondence are already excluded before the test. This is in fact a paradox between controlling and comparing in the test, and such paradox can only be solved or reduced if we learn more about similarities and differences of presuppositions between languages.

Secondly, in this experiment, no typical cultural or social specific features are involved, such as politeness or lying. Interestingly, recall that in the experiments by Reins et al. (2021), presuppositions are used for indirect lies. According to their results, the misleading effect of presuppositions is judged, at least numerically, differently by English and Russian speakers (for more details, see the review in Section 5.3). Combining our results with their observations, we may assume that this cross-linguistic difference, if any, might have more likely to do with a different understanding of *misleading* in different cultures or societies, rather than because of a difference in the commitment effect or the information status of presuppositions in different languages. The interaction between presuppositions and those cultural or social features could be very interesting for contrastive pragmatic research in the future.

Last but not least, although German and Chinese are two very different languages, they can of course not represent all languages. Nevertheless, the stability observed in our experiment is a relevant piece of the puzzle and can be a useful baseline for all studies that compare presuppositions or related phenomena in different languages and cultures. Additionally, this similarity is also important for studies on presuppositions with contrastive pragmatics involving a L2-perspective, especially when differences are observed.

In sum, our experiment shows that the group or sub-class of soft triggers has to be more heterogeneous than theoretical research has expected so far. In other words, if hard triggers are alike, it is possible that each soft trigger might be soft in its own way.

Hence, the hard-soft distinction perhaps rather illustrates a distinction between typical and less typical presuppositions or triggers, and the softness or less typicality might result from different reasons or mechanisms. With them, we can illustrate or better explain the shared properties of presuppositions. Moreover, our results also show that the projection behavior and the heterogeneity of presuppositions are cross-linguistically stable, as long as suitable equivalences can be found.

9.4 Sub-Experiment b: Comparison of '是...的 *(shi...de)*' and '...的是 *(...de shi)*'

9.4.1 Design and Materials

In order to address the questions in (144c), the two options of clefts in Chinese were compared. Therefore, this experiment had only one parameter, GROUP, with two levels, CNS and CND, within items but between subject. That is, the participants from group CNS read and rated the cleft with the '是... 的 *(shi...de)*'-structure, in which '的 *(de)*' was omitted as the sentence-final position is blocked by the question particle '吗 (ma)'. In group CND, participants were confronted with the other option, namely the '... 的是 *(...de shi)*'-structure, which is traditionally considered more compatible with pseudo-clefts or headless relative clauses, as reviewed in 9.2. For a better understanding, take the example with the cleft used in Sub-Experiment a, presented in Table 11: in the test, group CNS saw the clefts like in (153a) whereas group CND judged variations like (153b), while the target questions remained the same.

(153) **Uttered question:**

a. **'是... 的 *(shi...de)*'-structure:** 丁问道：**是**发型师女士在晚饭前和王先生在酒吧见了面吗?

ding wendao: **Shi** faxingshi nüshi zai wanfan qian he wangiansheng
Ding ask: **Shi** hairdresser lady at dinner before with Mr.Wang
zai jiuba jianlemian ma?
at bar meet-PST QST-particle?

'Ding (corresponding to David in German version) asks: Was it the hairdresser who met Mr. Müller at the bar shortly before dinner?'

b. **'... 的是 *(...de shi)*'-structure:** 丁问道：晚饭前和王先生在酒吧见面**的是**发型师女士吗?

ding wendao: wanfan qian he wangxiansheng zai jiuba jianmian **de shi**
Ding ask: dinner before with Mr.Wang at bar meet **de shi**
faxingshi nüshi ma?
hairdresser lady QST-particle?

'Ding (corresponding to David in German version) asks: Was it the hairdresser who met Mr. Müller at the bar shortly before dinner?'

Target question:

丁确定有人在晚饭前和王先生在酒吧见面吗？

'Is Ding certain that someone met Mr. Müller at the bar shortly before dinner?'

Despite the critical and control items with clefts, all other items in the test were exactly the same as in Sub-Experiment a. As the procedure of the experiment also remained the same, they will not be repeated here.

9.4.2 Predictions

Recall that on the one hand, most discussions about the difference between the two cleft options only mention the syntactical difference between them, and on the other hand, both *it*-clefts and pseudo-clefts (*wh*-clefts) are considered presupposition triggers (see e.g. the trigger list given by Levinson 1983c in Table 1). Thus, with respect to projection strength, we expect no difference between the two cleft structures. That is, we expect that both '是... 的 *(shi...de)*' and '... 的是 *(...de shi)*' can trigger presuppositions with a high certainty of the speakers, and the ratings are distributed similarly in both groups.

9.4.3 Participants

As mentioned before, the group CNS, which rated the '是... 的 *(shi...de)*'-structure and was compared with the group DE in Sub-experiment a, consisted of 51 participants. 6 of them did not pass the attention and understanding check (plot per subject see Figure 24 in Appendix A.3.4), leaving 45 participants (group CNS, mean age 22.91 ± 4.54) for the statistical analysis.

The group CND, which rated the '... 的是 *(...de shi)*'-structure, consisted of 33 participants; one of them did not pass the attention and understanding check (plot per subject see Figure 26 in Appendix A.3.4). Therefore, 32 participants (group CND, mean age 27.06 ± 4.64) were included in the statistical analysis.

9.4.4 Results

Recall that the only parameter in this sub-experiment was GROUP with two levels, CNS for 'shi...(de)' and CND for '...de shi', and that only items with clefts were varied between groups. Therefore, for the analysis of this test, all judgments on the control items, fillers and the items with the other three triggers were excluded. The data then encompassed 462 judgments taken from 77 individuals out of 2 groups for 6 critical items. Again

considering the rating scale as ordinal[51], the proportion of judgments is given in Figure 16.

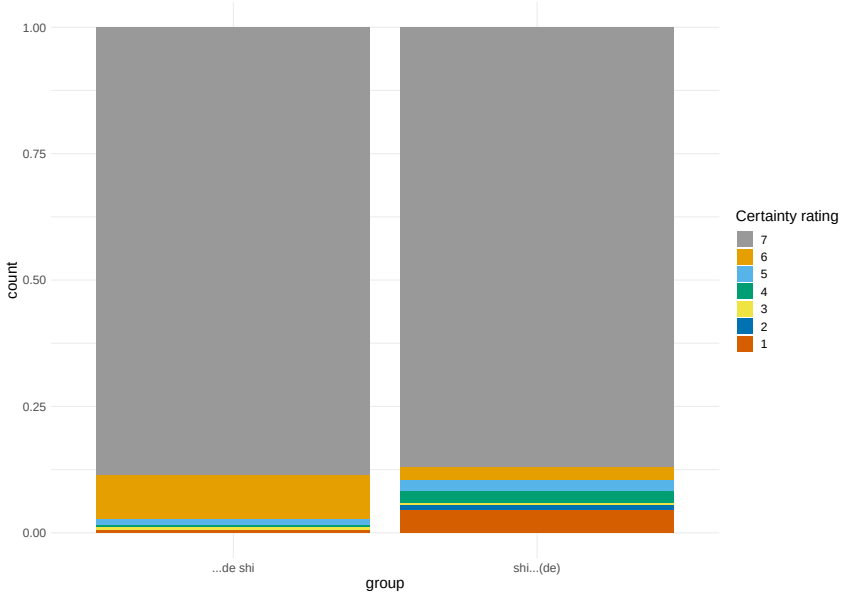


Figure 16: Proportion of ratings given to the cleft option '...de shi' (group CND) and 'shi...(de)' (group CNS).

To estimate the effects of GROUP, the data was analyzed using R (version 4.0.3, R Core Team 2020) with the Cumulative Link Mixed Model (CLMM) using the `clmm` function from the *ordinal* package (Christensen 2019). Similarly, the CLMM full model included GROUP as fixed effects and the maximal random effects structure[52], logit was used as the link function. The parameter GROUP was centered with *sum*-contrast, coded with (-0.5, 0.5). The results of the full CLMM model are listed in Table 16. As the estimates of effects and thresholds are on the log odds scale, the probabilities of ratings predicted by the model were also calculated and are visualized in Figure 17.

Table 16: Overview of the CLMM full model, fitted with the Laplace approximation, with estimates, standard errors, z- and p-values of fixed effects and threshold coefficients.

	Estimate	SE	z-value	p-value
GROUP1	0.61	0.66	0.94	0.35

Threshold	Estimate	z-value
1|2	-5.26	-8.29
2|3	-5.01	-8.11
3|4	-4.88	-7.99
4|5	-4.48	-7.60
5|6	-4.12	-7.21
6|7	-3.30	-6.22

[51]For plots with mean ratings of each group, in which the rating scale is treated as interval, see Appendix A.3.4.

[52]Y ~ GROUP + (1|Participant) + (1+GROUP|Item).

According to the analyses, there is no statistic difference between the two groups. Considering the similarity between patterns, the projection strength of the presuppositions triggered by the two variations of clefts can be considered similar, as we expected.

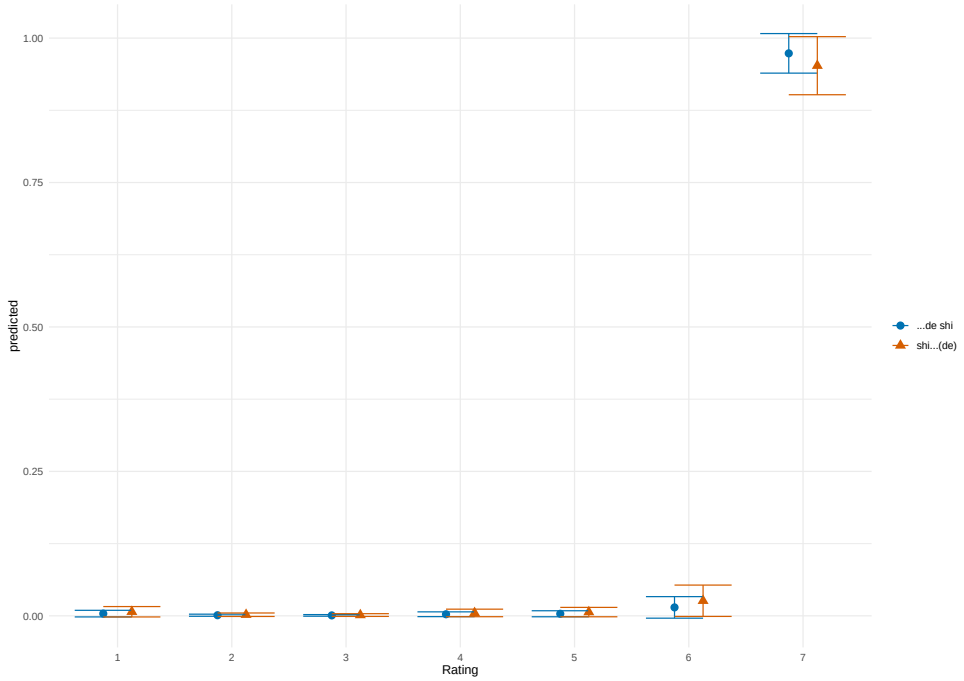


Figure 17: Probabilities of ratings predicted by the CLMM full model with 95% confidence interval.

9.4.5 Discussion

Addressing the research questions in (144c), repeated below in (154), our results clearly show that despite the syntactical differences, the two cleft structures in Chinese, '是… 的 *(shi…de)*' and '… 的是 *(…de shi)*', can both trigger presuppositions with high projectivity. In other words, regardless of whether they should be seen as two variations of *it*-clefts or rather counterparts of *it*-clefts and pseudo-clefts respectively, from the perspective of presuppositions, these two structures can been taken as equivalent triggers.

(154) Addressing the two cleft structures in Chinese: apart from their syntactical differences, are they both able to trigger presuppositions? If so, do their presuppositions receive similar projection strength?

This observation can be useful especially for further experimental cross-linguistic studies. As mentioned before, although the '是… 的 *(shi…de)*'-structure is a generally accepted counterpart for the *it*-cleft in Chinese (see e.g. Teng 1979; C.-T. J. Huang 1982, 1988; Zhan and C. Sun 2013; Liu and Kempson 2018; Zhou 2020), there are certain limitations that can affect or restrict a comparison between languages: first, the '是… 的 *(shi…de)*'-structure is highly infelicitous once post-verbal phrases like the object have

to be focused, while such a post-verbal focus is usual and natural for *it*-clefts in English or German. Second, once sentence-final position is blocked with elements like question particles, the '的 *(de)*' has to be dropped so that the structure cannot be differentiated from the bare '是... *(shi...)*'-structure – a structure that does not always pattern with clefts (Paul and Whitman 2008). Now, at least for experiments on presuppositions, these problems can be solved by using the '... 的是 *(...de shi)*'-structure, as our results support empirically that it can be an equivalent alternative trigger of the '是... 的 *(shi...de)*'-structure. In comparison to the '是... 的 *(shi...de)*'-structure, the '... 的是 *(...de shi)*'-structure can not only mark post-verbal elements like objects but also has less troubles with the sentence-final position. These properties might be relevant, for instance, to avoid unnecessary infelicities during the material drafting. Note however that this similarity is neither an argument for nor against the assumption that the '... 的是 *(...de shi)*'-structure is rather compatible with pseudo-clefts (*wh*-clefts), as both cleft structures can trigger presuppositions (see e.g. Levinson's (1983) trigger list, adopted in Table 1 before, based on Halvorsen 1978; Prince 1978; Atlas and Levinson 1981).

9.5 Chapter Conclusion

Modifying the experiment design from Tonhauser et al. (2018), the first sub-experiment in this chapter investigated the soft-hard split in Chinese and German by comparing the emotive factive verb *regret* and the cognitive factive verb *discover* with the typical hard trigger clefts and the typical soft trigger *win*. Results show that firstly, the projection behavior and the various projection strength of presuppositions are stable across these two languages if an equivalent counterpart of the trigger can be found and the translation is strictly controlled. Secondly, in line with most theoretical predictions (e.g. Simons 2001; Abbott 2006; Simons 2007; Abusch 2010), the emotive factive verb *regret* clearly patterns with the hard trigger clefts and differs from the soft trigger *win* in both languages. Thirdly, the ratings of *discover* did not pattern with *win* but rather with *regret* and clefts in our test. This observation is not only different from most theoretical predictions on the classification of *discover* (see e.g. Karttunen 1971b; Hooper and Thompson 1973; Abusch 2002, 2010; Karttunen 2016), but also from results of the previous empirical studies (e.g. Xue and Onea 2011; Djärv et al. 2018; Tonhauser et al. 2018). This rating possibly illustrates that the class of soft triggers, if any, is more heterogeneous than has been so far assumed, as their softness might be due to different reasons or mechanisms: while hard triggers are similarly projective, each soft trigger might be soft in its own way. The second sub-experiment proves that the two different cleft structures in Chinese, namely '是... 的 *(shi...de)*' and '... 的是 *(...de shi)*', can trigger presuppositions with similarly high projectivity, despite their syntactical differences. This observation can be helpful

for further empirical studies on cross-linguistic comparisons of presuppositions involving Chinese, especially for their item design.

10 Summary of Part II

Regarding the three main research concerns and focusing on the five main research questions listed in (85), three experimental studies were preformed and reported in this part. These main research concerns and questions are repeated below in Figure 18, together with the relation between them and the three experiments. This chapter is a brief summary of the experiments with a focus on their background, hypotheses, design and main findings.

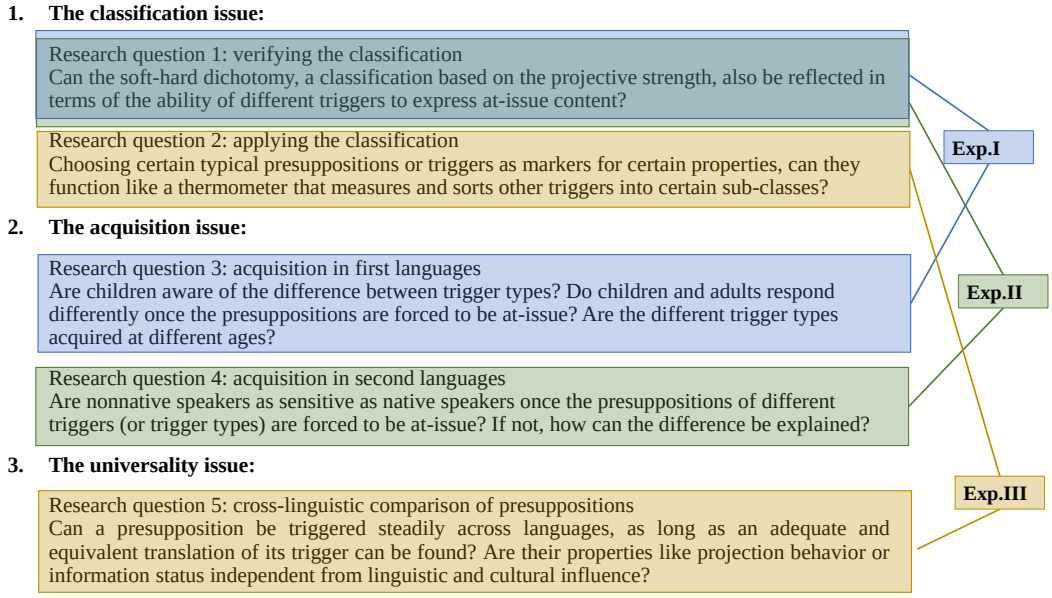


Figure 18: Relation between the five research questions and the three experimental studies.

The first experiment (Exp.I) mainly aims at Research question 1 and 3. In the context of the controversy over the classification of presuppositions (see the review in Section 2.2.2 and 3.1) and the correlation between projection and (not-)at-issueness (see e.g. Xue and Onea 2011; Tonhauser et al. 2018 and the review in Section 3.2, among others), the soft-hard split of triggers was verified from the perspective of at-issueness. Note that the terms *soft* and *hard* are used in a wider sense in my study than originally proposed by Abusch (2002, 2010). It was hypothesized that in comparison to hard triggers, soft triggers are more able to express at-issue contents and therefore their presuppositions are more suitable for targeting the Question under Discussion (QUD) directly than the presuppositions of hard triggers.

Another relevant perspective in order to investigate the trigger classification is presupposition acquisition: that is, when and in how far are children aware of presuppositions and their trigger distinction. According to previous studies on presuppositions in first language acquisition, reviewed in Section 4.1, the age range between 4 and 6 years old is considered to be critical. Hence, 23 monolingual German preschool children between 4 and 6 years of age formed the critical group in Exp.I and they were compared with 33 adult controls.

During the test, the at-issueness, or the ability to express at-issue contents, was measured by means of the appropriateness between the question and its answer: that is, how felicitous or acceptable it is when the relevant information for the explicit QUD is offered by the presupposition of the answers, violating the Not-At-Issueness Constraint in (55). The experiment also included non-restrictive relative clauses (NRRCs) as baseline, as they are not-at-issue but not presupposed either. Moreover, no operators from the family of sentences (Chierchia and McConnell-Ginet 2000) were involved in the test. For ease of understanding, video stimuli were provided with a 5-point smiley scale for rating. The results mainly support the following findings:

- First, the soft-hard split can be reflected in terms of at-issuenss. In particular, the presuppositions of soft triggers are indeed more able to felicitously answer a QUD than those of hard triggers. In other words, soft triggers are more able to express an at-issue content than hard triggers.
- Second, preschool children between 4 and 6 years of age are not only generally aware of violations triggered by at-issue presuppositions, but can also recognize the difference between hard and soft triggers in their abilities to express at-issue contents.
- Third, in comparison to adults, preschool children are still less sensitive to the trigger distinction and at-issueness violations. Possible explanations might be the acquisition process, limited working resources and/or children's pragmatic tolerance (Katsos and Bishop 2011).
- Fourth, no difference in at-issueness between presuppositions of hard triggers and non-restrictive relative clauses was observed.
- Fifth, numerically, the presuppositions triggered by *discover* were rated most suitable for being at-issue of all other triggers by both adults and children.

The second experiment (Exp.II) focuses on Research question 1 and 4, investigating the relation between the soft-hard split and (not-)at-issueness in second language acquisition. In comparison to first language acquisition, the acquisition of presuppositions in second languages has hardly been experimentally explored (see the review in Section 4.2). However, studies on related phenomena, such as implicatures, show that gaining

pragmatic skills and sensitivity in foreign languages may not happen as naturally or self-evidently as expected. Moreover, the difference between native and nonnative speakers can shed light on the mechanism behind certain phenomena. Therefore, 36 Chinese learners of German were compared with German native speakers (the same control group as in Exp.I) in order to explore whether and in how far at-issueness violations and the distinction between triggers are recognized by the nonnative speakers.

The second experiment used exactly the same design and material in the first experiment. Results in particular display the following findings:

- First, the relation between projection strength and at-issueness is proved again, as both native and nonnative speakers perceived that the presuppositions associated with soft triggers are more suitable for answering a QUD than those of hard triggers.
- Second, while the ratings are similar between both groups with respect to trigger distinction, regarding the at-issueness violation, nonnative speakers are significantly less sensitive or more tolerant than native speakers. Possible explanations might be the uncompleted acquisition of pragmatic competence in the foreign language and nonnative speakers' pragmatic tolerance.
- Third, the nonnative speakers' deviation from the native speakers was only significant for hard triggers, illustrating the boundary between scalar implicatures and soft triggers on the one hand, and between non-restrictive relative clauses and hard triggers on the other.
- Fourth, numerically, the presuppositions triggered by *discover* were rated most suitable for being at-issue of all other triggers by both native and nonnative speakers.

The third experiment (Exp.III) mainly addresses Research question 2 and 5. Therefore, slightly different from the former two experiments, this experiment serves rather as a baseline experiment in my whole research. Its aim is not only to investigate whether there could be a 'thermometer' for measuring and sorting presupposition triggers according to their projection strength, but also to verify whether there is a baseline for comparisons between speakers with different first languages, in the expectation of cross-linguistic stability of presuppositions.

In order to measure the projection strength of presuppositions, the certainty judgment task proposed by Tonhauser et al. (2018) was adopted in Exp.III, as it can evaluate speaker commitment more directly than other diagnostics and has been widely used with a stable function (see the review in Chapter 9 and also Tonhauser 2016; Stevens et al. 2017, among others). In this test model, assertions with a presupposition trigger were embedded under the question operator and uttered by a speaker. The projection strength was then measured by how certain the speaker is about the presupposition. Assuming that there is a soft-hard split of triggers, regardless of due to which mechanism or criterion, it

is expected that using certain typical triggers as markers or anchors, other triggers can be classified by measuring their projection strength and comparing them with the markers. In Exp.III, clefts and *win* were selected as anchors for hard and soft triggers respectively, while two factive verbs, *regret* from the emotive sub-class and *discover* from the cognitive or semi-factive sub-clase, were chosen and measured.

As for the cross-linguistic comparison, recall that as we have seen in Chapter 5, the pragmatic principles and features are not as universal or language-neutral as different theories assumed. Considering especially the East-West divide or the contrast between Western and non-Western societies, observed in the pragmatic field, the cross-linguistic stability of presuppositions cannot be simply assumed but needs to be empirically verified. Thus, for the comparison between languages, one Western and one Eastern language were chosen, namely German and Chinese.

In Exp.III, the speaker's certainty was rated with a 7-point-scale from absolutely uncertain to absolutely certain. Judgments given by 32 German native speakers and 45 Chinese native speakers were analyzed in the first sub-experiment. Moreover, as the cleft in Chinese might have two different syntactic structures, another group of 32 Chinese participants were included and compared with the first group in the second sub-experiment, in order to inquire whether these two structures also differ with regard to presuppositions. Based on the experiment results, the main findings of Exp.III are summarized as follows:

- First, there is a clear split between the two anchor triggers, confirming again the heterogeneity within the category of presuppositions and the need for classification.
- Second, the cognitive factive or semi-factive trigger *discover* received an unexpected high rating of certainty in this test and did not pattern with the soft anchor but rather with the hard one, against most theoretical assumptions. This might indicate that the class of soft triggers is more heterogeneous, and the mechanism behind this softness might be more complex and diverse than it has been assumed so far.
- Third, both projection behavior and the difference in projection strength can be cross-linguistically stable, as long as appropriate counterparts are available.
- Fourth, no difference between the presuppositions triggered by the two cleft structures in Chinese was observed, implying that they can be exchanged with each other if needed, at least for studies on presuppositions.

In sum, regarding the five main research questions, the three main research concerns were investigated with three different experiments. Although all observations and findings in Part II were discussed in the context of each experiment in detail, with the three main research concerns in mind, they should be all combined and discussed in a broader context. This general discussion can be found in the next Part, Chapter 11, with a focus on the classification, acquisition, universality and methodological issues.

Part III

General Discussion

11 Discussion

As the empirical observations have already been extensively discussed in the context of each experiment in separate chapters above, this chapter aims at a general discussion on issues that need to be considered inter-experimental, that is, by means of combining or comparing results and/or assumptions from more than one of the studies presented in Part II. After a discussion of the three main research concerns the first three sections, some relevant or interesting methodological aspects will be reviewed in Section 11.4.

11.1 Discussion of the Classification Issue

The first main research concern of my study is about the classification issue of presuppositions or their triggers, and it consists of two main research questions, namely Research question 1 and 2 in (85a), repeated below in (155):

(155) The classification issue:

a. Research question 1: verifying the classification: can the soft-hard dichotomy, a classification based on the projective strength, also be reflected in terms of the ability of different triggers to express at-issue content?

b. Research question 2: applying the classification: choosing certain typical presuppositions or triggers as markers for certain properties, can they function like a thermometer that measures and sorts other triggers into certain subclasses?

Addressing these two main research questions, note firstly that the terms *soft* and *hard trigger* are used in a wider sense than they are originally proposed by Abusch (2002, 2010): instead of claiming different triggering processes, *soft* in this work refers to triggers whose presuppositions are considered, depending on different theories, to be weaker or less often projective, more defeasible or suspendible, more sensible to certain contextual influence, or more easily to lose their projectivity and receive a local reading. *Hard triggers*, on the other hand, are expressions whose presuppositions are in general stable, highly projective and reluctant to these effects.

Research question 1 is mainly targeted by the first and second experiments (Exp.I and Exp.II), in which the soft-hard split was tested with regard to at-issueness, that is, whether and how properly their presuppositions can target the Question Under Discussion (QUD,

see e.g. von Stutterheim and Klein 1989; Roberts 1996) directly. The observations in both experiments illustrate that the presuppositions of typical soft triggers like *win, manage* and *discover* are significantly more suitable for answering the QUD or less reluctant to an at-issue reading than those of typical hard triggers like *too, again* and *it*-clefts. In other words, both experiments support the idea that the classical soft-hard split of triggers – based mainly on their varying projection strength – is actually reflected in the at-issueness ability of their presuppositions, even if they are not embedded under entailment-canceling operators, such as the family of sentences suggested by Chierchia and McConnell-Ginet (2000), see e.g. (4). Moreover, results show that this reflection can also be perceived by native preschool children and nonnative adults, although both groups were less sensitive than the native adult speakers.

These observations powerfully confirm that the projection strength of a content is correlated with its potential to express at-issue content, i.e. its relevance to the QUD, which is also consistent with the prediction given by the at-issueness account (see e.g. Simons et al. 2010; Tonhauser et al. 2018 and the review in Section 3.2). As has already been discussed extensively in Section 7.4 (and in Y. Chen et al. 2022), a key concept of the explanation for this correlated difference can be local accommodation, a strategy that can modify the at-issueness state of a content and therefore also its projectivity. Moreover, local accommodation does not only take place in clauses embedded under entailment-canceling operators, but also in matrix sentences if needed and available – needed if the presupposed content offers an answer to the QUD in the discourse, and available if the trigger is not impervious to such local accommodation.

With this in mind, recall that the theories from the at-issueness account (see e.g. Simons et al. 2010; Tonhauser et al. 2018 and the review in Section 3.2) in fact expand and relocate the question of the projection heterogeneity of presuppositions to the question of the at-issueness heterogeneity of all projective contents. The observations in Exp.I and Exp.II, especially ratings given by the native adult speakers, reveal clearly that even when triggers are not embedded under entailment-canceling operators and even when their presupposed contents are all relevant to the explicit QUD, the at-issueness of presuppositions still varies: with some within-class variation, the ratings clearly demonstrate a split of at-issuneness, which reflects the soft-hard distinction of triggers. This connection or correlation sheds light on the question of the at-issueness heterogeneity by indicating that the soft-hard split – at least the concepts of trigger classification – can be helpful or even necessary for explaining it, and supports the idea that the assumptions about a trigger split and the at-issueness perspective can be combined. Take again the example (109) given in Section 7.4 (see also Y. Chen et al. 2022, p. 33), repeated below in (156), and compare it with the hard trigger in (157): the different level of at-issueness of presuppositions, or

the different ability of triggers to express at-issue contents can probably be explained by the trigger split as below:

(156) "Did the duck participate in the competition?"
Little Peter responds: "The duck won the competition."

a. The presupposed content triggered by *win* is relevant for the explicit QUD, local accommodation is needed.

b. *win* is a soft trigger and not resistant to local accommodation, local accommodation is available.

⇝ Local accommodation takes place, the answer is understood as:
Little Peter responds: "The duck participated in the competition and won (it)."

(157) "Did the panda take swimming lessons for the first time yesterday?"
Little Peter responds: "The panda took swimming lessons again yesterday."

a. The presupposed content triggered by *again* is relevant to the explicit QUD, local accommodation is needed.

b. *again* is a hard trigger and resistant to local accommodation, local accommodation is not available.

⇝ Local accommodation does not take place, little Peter's answer is rated as odd.

In short, the at-issuenss of a presupposition, which can predict its projection strength, is not only determined by whether the content includes an answer to the QUD, but might also depend on how hard the expression resists local accommodation: if the content is relevant to the QUD and the trigger is soft, that is, it is not resistant to local accommodation, then restrictions like the Not-At-Issueness constraint in (55), repeated below again, can also be broken, as the content of the presuppositions is shifted to the foreground by means of local accommodation.

(158) Not-At-Issueness Constraint
Presuppositions cannot be used to directly target the Question Under Discussion. (Aravind and Hackl 2017, p. 51)

A question that arises, then, is: what mechanism or property prevents hard triggers, or their presuppositions, from local accommodation? The experiments in this study hardly provide any direct answer, but the results in Exp.II and the third experiment (Exp.III) potentially exclude or at least challenge some classical assumptions, such as the idea that presuppositions of hard triggers are just lexically stored or semantically triggered. Note that in Exp.III, it can be clearly observed that the projectivity of presuppositions is judged cross-linguistically stable, at least between German and Chinese. But in Exp.II, the results show that the Chinese learners of German are significantly more tolerant

towards or less sensitive to at-issueness violations involving hard triggers than the native German speakers – this insensitivity can thus hardly be explained without a pragmatic perspective on hard triggers (for more discussion on nonnative speakers' tolerance, see discussions in Section 8.3 above and in Section 11.2 on Research question 4 below).

Research question 2 is mainly investigated in Exp.III, in which the projection strength of presuppositions triggered by clefts, *win*, *regret* and *discover* was compared between German and Chinese. According to the results, the typical hard trigger clefts and the typical soft trigger *win* used in the experiment can indeed function as markers for measuring the projection strength, not only because their presuppositions differ in their projection strength, but also because their readings are less ambiguous and their difference in projection strength is stable across languages.

However, while the markers and measurement worked well with the hard triggers in that ratings of the emotive factive verb *regret* patterned with the hard marker clefts and differed from the soft marker *win* as expected, this was not the case for the soft triggers. Judgments of the cognitive factive trigger *discover*, a classical, typical soft trigger, surprisingly patterned with the hard marker and differed from the soft one in Exp.III. Combining the high projection strength of presuppositions of *discover* found in Exp.III with their high level of at-issueness and ease with local accommodation observed in Exp.I and Exp.II, and comparing this variability with the stable performance of *win*, it is reasonable to argue that the softness of *discover* may not be provided by the same mechanism as that of *win*. In other words, while the hard triggers are similarly hard, each soft trigger might be soft in its own way. This raises the following question: how can this heterogeneity within the class of soft triggers be explained?

As has been discussed in detail in Subsection 9.3.7, the softness of *win* is perhaps due to the reason that the presupposition (x has won y $\rightsquigarrow$ x participated in y) is both temporally and logically more closely connected with its trigger. Therefore, it is more at-issue than presuppositions of hard triggers. In comparison to *win*, the projection behavior of presuppositions triggered by *discover* is in fact not soft but either hard or not projective at all. That is, a *discover*-sentence has normally two readings: If the QUD focuses on the cognition status, the matrix sentence is at-issue and the complement can be considered to be presupposed and projective, as shown in Exp.III with a murder story as background. However, if the matrix clause has only an evidential function (Simons 2007) and it is the content of the complement which is relevant to the QUD, then the complement becomes naturally at-issue and loses its projectivity, as shown by the high ratings in Exp.I and Exp.II, in which the explicit QUDs target the complement. In short, the at-issueness or projection strength of a complement of *discover* depends mainly on which reading of *discover* is preferred in the given context. This softness is thus caused by a either-or choice

and differs from the softness of *win*, but is rather in line with Schlenker's (2010) concept of part-time triggers.

The different kinds of softness indicate that soft triggers, as a sub-class of expressions with presuppositions, are actually a more heterogeneous group than most theoretical and empirical studies predict. In fact, with reference to the classification possibilities reviewed in Section 3.1 and 3.3, these triggers are classified as a separate group mainly because they differ more or less from some typical triggers in certain criteria: for instance, based on their suspendability or contextual defeasibility (see e.g. Simons 2001; Abusch 2002, 2010, but also Abrusán 2011, 2016[53]), these triggers are classified as not-hard because they are less stable or less resistant to contextual changes than other triggers. From the perspective of the resolution-lexical distinction (see e.g. Zeevat 1992; Kripke 2009), these triggers are considered less anaphora-alike than clefts or *too* but their presuppositions are rather logically required. Therefore, they are classified into the not-resolution class. In terms of possible readings, triggers like *announce* or *discover* are separated from the other triggers because their complements may but not always have to be presupposed (see e.g. Simons 2007; Schlenker 2010). Additionally, in Exp.I and Exp.II, observations in the realm of both first and foreign language acquisition disprove the hypothesis that presuppositions of soft triggers share the same or similar processes with scalar or conversational implicatures.

Regarding the question of whether the classification of presuppositions should be considered a binary or graded split, with the observations of my three studies, we still do not have enough data to draft a conclusive answer. However, considering all of those points mentioned above, we may hypothesize that instead of a pragmatic-semantic or a conversational-conventional distinction, the soft-hard split might rather illustrate a distinction of less typical and typical triggers, or triggers with less typical and typical presuppositions, and this assumption might offer us a new perspective on the question. That is, the assumption of typicality is able to combine the two relevant hypotheses from the two main pragmatic accounts, indicating that the heterogeneity issue needs to be explained by taking several dimensions into account, such as the lexical meaning of presupposition triggers, their sensitivity toward contextual changes, their at-issueness or information structure, the relation between the presupposition and the trigger, and so on. If so, then presuppositions cannot be measured on only one scale, as there are more criteria in which some soft triggers may diverge from hard ones, and the measurement should also be multidimensional. Take a watermelon for example. Depending on how we cut it, the distribution of seeds in it can be observed as a circle, an oval, or in the form of some lines. Similarly, depending on with which criterion or properties we measure the

[53]Note that although she argues against a classification of triggers based on the distinction between semantic and pragmatic triggering processes, Abrusán (2011, 2016) also claims that presuppositions of some triggers are more sensitive to contextual changes than those of other triggers.

presuppositions, or from which perspective we "cut" the complex into slides, the classification can be both two-way split and gradient. Moreover, children's awareness of and difficulties with presuppositions (or not-at-issue contents in general) can also be broken down into questions of their acquisition of certain properties or abilities.

However, note that currently, there is not enough evidence for such a typicality hypothesis either, and there are still many unanswered questions. Nevertheless, my experimental studies have shown that recent theories on the heterogeneity and classification of presuppositions might not be the be-all and end-all solution, but rather some in-progress parts of it.[54]

In sum, based on the observations of all three experiments, further discussions and assumptions regarding the first main research concern on the classification of presuppositions can be summarized mainly as follows:

- First, the soft-hard split is reflected in the level of at-issueness, and this reflection can be explained with the key concept of local accommodation. While at-issueness may predict the projection behavior of presuppositions, the property itself is influenced not only by how relevant the content is to the QUD, but also possibly by how hard the trigger is – that is, how resistant the trigger is to local accommodation.
- Second, to explain the reluctance of hard triggers to local accommodation and at-issue readings, it is probable that assumptions like lexical or semantic triggering are neither sufficient nor appropriate, as pragmatic perspectives are needed.
- Third, the sub-class of soft triggers is more heterogeneous than most theories on classification predict. Among them, there is at least a distinction between *win*-like softness and *discover*-like part-timeness, as their accessibility for local accommodation might result from different sources. This variation might indicate that the soft-hard split rather illustrates a distinction between less typical and typical triggers or presuppositions, and the less-typicality can result from several distinct properties or factors.

[54] Recently, Schlenker (2021b) proposes a triggering algorithm that "take[s] as an input a contextual meaning, and [...] turn[s] some contextual entailments into presuppositions". Briefly summarized, he assumes that presuppositions are not lexically stored or somehow embedded in the lexical meaning of the triggers. Instead, this kind of information is considered "an epistemic precondition of the global meaning", and it becomes presupposed or projective as the speaker for instance requires its linguistic triviality or givenness in certain contexts. According to Schlenker, this algorithm should replace the notion of presupposition triggers. This assumption provides some novel perspectives for typical questions surrounding presuppositions, especially the heterogeneity problem and the universality issue. With regard to the experimental observations in my dissertation, especially the homogeneity among soft triggers, this algorithm can doubtlessly offer some inspiring explanations and a new way to slice or break down the complex. However, I think a more crucial question arises: with such a triggering process that is based mainly on contextual and epistemic effects, how can we explain some other observations like the nonnative speakers' tolerance, which can be observed, above all, regarding hard triggers? Whether it is possible and necessary to combine the typicality hypothesis with such a triggering algorithm is a highly interesting problem and I would like to investigate it in detail in the future.

- Fourth, sorting presuppositions or triggers by means of certain typical examples as measuring markers seems possible, but assuming that the softness is due to different properties, the complete 'thermometer' (if any) should be rather multidimensional.

11.2 Discussion of the Acquisition Issue

The second main research concern of my study focuses on the acquisition of presuppositions in both first and second languages (L1 and L2) with two main research questions, namely Research question 3 and 4 in (85b), repeated below in (159):

(159) The acquisition issue:

a. Research question 3: acquisition in first languages: are children aware of the difference between trigger types? Do children and adults respond differently once the presuppositions are forced to be at-issue? Are the different trigger types acquired at different ages?

b. Research question 4: acquisition in second languages: are nonnative speakers as sensitive as native speakers once the presuppositions of different triggers (or trigger types) are forced to be at-issue? If not, how can the difference be explained?

Research question 3 is only targeted by Exp.I. As the observations of Exp.I have already been discussed in-depth before and observations of the other two experiments hardly contribute to this question, the discussion in Section 7.4 (and in Y. Chen et al. 2022) will be only shortly summarized below.

Regarding the first two questions in (159a), Exp.I shows that German preschool children are already aware of the at-issueness difference between foregrounded and backgrounded contents in conversation and they are sensitive to at-issueness violations, which is in line with the observations in Aravind and Hackl (2017) for English. Moreover, the experiment further illustrates that, although significantly less sensitive than the adult speakers, children between 4 and 6 years of age are also competent to recognize the distinction between soft and hard triggers by noticing the different degrees of possibility or felicity for an at-issue reading of their presuppositions, at least with the support of an explicit QUD. This observation further supports the idea that the development stage before 7 years of age is not only crucial for the acquisition of presuppositions (see e.g. Bill et al. 2016) but also for perceiving their heterogeneity.

Regarding the last question in (159a), recall first that in previous studies, preschool children at 3 years of age are able to understand the presuppositions of hard triggers like *also* and the definite article *the* (Syrett et al. 2010; Berger and Höhle 2012) but still struggle with *know* under negation (Dudley et al. 2015). Second, children at early

school age have a high level of accuracy when rating the existence presupposition of *every* but not with the implied anti-uniqueness presupposition (Yatsushiro 2008a). Third, with the children's ratings given in Exp.I, especially the median split by age (see Figure 6), some divergences of the acquisition processes might be seen between trigger types. Together with these observations, these divergences might give us some indirect clues to this question. However, as the data in Exp.I is in fact insufficient to answer this question, this issue has to be left for further research.

Although Research question 4 is mainly targeted by Exp.II, the results need to be discussed in combination with observations of Exp.I and Exp.III, especially with regard to the last question in (159b).

Addressing the first question in (159b), results show that the Chinese learners of German with an advanced foreign language level are aware of the soft-hard split of presupposition triggers and their different potential to express at-issue contents. However, in comparison to the native speakers, they are still less sensitive to or more tolerant towards at-issueness violations, and their ratings on at-issue presuppositions are, at least numerically, less strict than those of the German controls. Additionally, the statistically significant difference was surprisingly mainly observed with the hard triggers but absent for the baseline items with non-restrictive relative clauses (NRRCs). As discussed in Section 8.3 above, this contrast confirms the distinction between presupposed and non-presupposed projective contents and supports the idea that although presuppositions of hard triggers and NRRCs are both highly not-at-issue and projective, they cannot be sorted into one class, which is in line with the claims in A. Holler (2005) and Roberts et al. (2009). Moreover, the difference between L1 and L2 groups regarding soft triggers was not significant in Exp.II, while it was significant for the ratings of scalar implicatures in my previous study (Y. Chen 2019). With this contrast, the assumption of a uniform approach for presuppositions of soft triggers and scalar implicatures (see e.g. Chemla 2008; Romoli 2014) is clearly and seriously challenged, which is in line with the evidence from the perspective of L1 acquisition (Bill et al. 2016).

Regarding the last questions in (159b), one possible explanation for the insensitivity of nonnative speakers might be the on-going acquisition of some pragmatic skills or constraints at the discourse level in L2, as has been discussed before in Section 8.3. Interestingly, with the discussion on the resistance of hard triggers to local accommodation in Exp.I in mind, some criteria used for other classification possibilities may perhaps add some details to this explanation (for more about these classification possibilities, see the review in Section 3.3).

The first interesting criterion could be the similarity between some presuppositions (mainly of hard triggers) and anaphora, proposed firstly by Zeevat (1992). Briefly speak-

ing, according to Zeevat (1992), typical hard triggers like *again, too* and clefts are assumed to be more anaphora-alike than other triggers like *stop*.[55] This discourse-anaphoric function might not only strongly demand the existence or the possibility of an assumption on the existence of presupposed contents as an antecedent, but may also limit the possibility for local accommodation and increase the difficulty of an at-issue reading of the presupposed content (for more discussion about the difference between anaphoric availability and not-at-issueness, see Snider 2017b). It is possible that nonnative speakers, in contrast to native children, are aware of the possibility of a local reading for soft triggers. However, they are less sensitive to these limits or resistance with regard to hard triggers, so that their tolerance or deviation from native speakers is more obvious with respect to the at-issue presuppositions of the hard triggers than those of the soft ones.

Another criterion that might be interesting for this explanation is proposed by Klinedinst (2016), based on observations in Sudo (2012): the hard triggers do not obligatorily entail their presupposition in a non-monotonic context. However, note that the items used in the experiment did not contain any typical non-monotonic quantifier. Thus, if the absence of entailment for hard triggers in a non-monotonic context can affect the nonnative speakers' comprehension in the experiment, it may influence them indirectly and possibly via some related properties in the test.

In short, whether these features or limits can affect the comprehension and how they – or some properties related to them – can cause the insensitivity among the nonnative speakers, has to be investigated by future research. Additionally, note that this explanation might not be sufficient as it has to face some challenges observed in Exp.II. For instance, when rating the filler items, the nonnative speakers only show such insensitivity with regard to under-informative answers but not to over-informative answers. However, it is hard to argue that over-informativeness is acquired earlier in L2 so that the nonnative speakers and native speakers share the same intuition, while the acquisition of under-informativeness takes place later so that they do not.

Another possible explanation is nonnative speakers' pragmatic tolerance, comparable to children's pragmatic tolerance (Katsos and Bishop 2011): in comparison to native speakers, nonnative speakers are generally more tolerant or less strict towards pragmatic violations. Note first that there is no conflict between these two explanations but they can be smoothly combined. Second, this assumption is also supported by the results in

[55] Also focusing on the requirement on a linguistic antecedent, Goebel (2020) distinguishes triggers like *too, also* and clefts from triggers like *win, regret* by claiming that the former ones are focus-sensitive, require an antecedent in the context and are therefore sensitive to the QUD-structure and more difficult to globally accommodate than the later ones. Slightly different from most classifications reviewed so far, Goebel (2020) also claims that *again* lacks the focus-sensitivity and therefore cannot be sorted into the group with *too, also* and clefts. As this sorting cannot be reflected in or supported by the results of the experiments with the perspective of QUD/at-issueness in this work, this classification will not be further discussed.

Exp.III: when participants were asked to rate presuppositions in their mother tongue, no significant effect of LANGUAGE was observed. That is, German and Chinese native speakers comprehend the projection of presuppositions similarly in their mother tongue. Recall that the projection strength is related to or even dependent on the at-issueness (see e.g. Simons et al. 2010; Tonhauser et al. 2018); the results then indicate a cross-linguistic stability of projection behavior of presuppositions and thus their at-issueness, at least between German and Chinese. Now, considering the stability and similarity of ratings in German and Chinese observed in Exp.III, and the significantly more tolerant ratings given by Chinese learners of German in Exp.II, it is reasonable to assume that nonnative speakers have pragmatic tolerance in general. Third, this assumption is also supported by the observations in Y. Chen (2019), in which the logically correct but pragmatically inappropriate scalar terms were rated significantly more acceptable among nonnative speakers than among native speakers. Moreover, nonnative speakers' pragmatic tolerance is also in line with the increased logical choice preference or decreased certainty in L2 observed in psychological studies (see e.g. Costa et al. 2014; Geipel et al. 2015). For more discussion on this assumption, see Section 8.3.

Last but not least, the experiments also confirm the relevance and usefulness of the L2-perspective for research on contrastive pragmatics. More discussion will take place in Section 11.4.

In sum, based on the empirical observations of all three experiments, further discussions and assumptions regarding the second main research concern on the acquisition of presuppositions can be summarized mainly as follows:

- First, although significantly less sensitive than the adult speakers, German preschool children between 4 and 6 years of age are competent to recognize at-issueness violations and the distinction between soft and hard triggers. In comparison to the previous studies, this observation further confirms that the developmental stage before 7 years of age is not only relevant for the acquisition of presuppositions, but also for the perception of the trigger distinction.
- Second, nonnative speakers are aware of the trigger distinction and at-issueness violations, but still less sensitive than native speakers. This insensitivity is significant mainly with respect to hard triggers but not to NRRCs and soft triggers. This contrast confirms the distinction of NRRCs from presuppositions on the one hand, and challenges the uniform account of presuppositions and scalar implicatures (see e.g. Chemla 2008; Romoli 2014) on the other.
- Third, nonnative speakers' insensitivity can be possibly explained by the acquisition process of certain pragmatic skills or constraints at discourse level in L2. Some criteria used for other classification probabilities can be interesting, such as the

strong discourse-anaphoric function of hard triggers or their absent entailment in certain contexts. Another possible explanation is nonnative speakers' pragmatic tolerance, analogical to children's pragmatic tolerance observed in L1-acquisition (Katsos and Bishop 2011). These assumptions also highlight the relevance and usefulness of contrastive pragmatics with a L2-perspective.

11.3 Discussion of the Universality Issue

The third main research concern of this study is about the universality of presuppositions, and it consists of only one main research question, namely Research question 5 in (85c), repeated below as (160):

(160) The universal issue:
Research question 5: cross-linguistic comparison of presuppositions: can a presupposition be triggered steadily across languages, as long as an adequate and equivalent translation of its trigger can be found? Are their properties like projection behavior or information status independent from linguistic and cultural influence?

The universality issue arises due to the reason that although most classical theories assume or imply that pragmatic phenomena or enrichments are universal or general as they are based on human rationality or cognition, differences between languages and societies have been experimentally observed (see the review in Chapter 5). Therefore, the universality or cross-linguistic stability of presuppositions cannot be just assumed, but needs to be empirically verified. Moreover, in order to explain the more tolerant ratings given by the nonnative speakers in Exp.II, the influence of their native language has to be investigated.

This research concern is mainly addressed by Exp.III, and again, as the results have already been extensively discussed in Chapter 9, they will only be summarized here. In short, Exp.III shows that the projection strength of presuppositions can indeed be cross-linguistically stable, if 1) the triggers are carefully selected, that is, the triggers should have syntactically parallel and semantically equivalent counterparts in the languages tested in the experiment, with a minimum possibility of ambiguity and, if possible, a high frequency in daily communication; 2) the translations of the sentences used in the test are strictly controlled and 3) the background story or context avoids typical culture-specific events.

Therefore, one has to keep in mind that the results of Exp.III cannot totally exclude the possibility of a culture- or language-specific understanding or processing of presuppositions. First of all, the trigger selection has already excluded certain controversial issues that may vary between languages and then influence the rating – a paradox between accurate control and comprehensive investigation. This problem can only be solved if we learn

more about whether, and if yes, which cultural elements play a role in presupposition triggering, especially by means of studies on language comparison or pragmatics in second languages. Second, the certainty judgment task adopted from Tonhauser et al. (2018), which aims to evaluate speaker commitment more directly, in the experiments only measure the projection strength without any interaction with politeness, lying or speech acts which can vary between languages. Once those effects are added and presuppositions are tested as a strategy in this context, it could still be possible to observe differences between languages and societies, for instance the (at least numerically) different misleading effect of presuppositions observed in Reins et al. (2021). Third, only two languages, German and Chinese, were compared in the test, which of course cannot represent all languages and their usage in all social groups.

Nevertheless, the stability observed in Exp.III is in line with the main findings in Amaral and Cummins (2015) and Schwarz et al. (2020), and it offers relevant evidence for the universality of presuppositions by means of an East-West comparison. It is persuasive enough to support the idea that there are some cross-linguistically stable properties of presuppositions, such as their projection behavior and the heterogeneity of their projection strength. If such projection variation leads to a classification of presupposition triggers, the split, regardless of it being soft-hard or less typical-typical, should also be considered cross-linguistic as well, perhaps with the triggers like clefts in one group and those like *win* in another. Moreover, the varying information status of the cognitive factive verb *discover* and the mechanism behind this variability, discussed before in Section 11.1, should also be stable between languages. This cross-linguistically stable variability possibly illustrates furthermore how at-issueness (Simons et al. 2010) and relevance (Sperber and Wilson 1996) cross-linguistically affect our interpretation of cognitive factive predicates, at least in both German and Chinese.

In sum, based mainly on the observations of Exp.III, the discussions and assumptions regarding the third main research concern on the universality of presuppositions can be summarized as follows:

- First, extending the nondetachability of presuppositions within one language proposed by Simons (2001), we may assume that presuppositions, or at least their triggering and projection, can be considered stable across languages, as long as an adequate and equivalent counterpart of their triggers can be found in another language.
- Second, if adequate counterparts can be found and no culturally or linguistically varying pragmatic features are involved, the heterogeneity of projection strength of presupposition and the split of presupposition triggers can be expected to be stable between languages, too.

11.4 Discussion of Methodological Issues

In this study, the three main research concerns and the five main research questions on presuppositions have all been investigated empirically by means of contrasts: a comparison of native speakers at different developmental stages of one language in Exp.I, a comparison between native and nonnative adult speakers of one language in Exp.II, and a comparison of adult native speakers of two languages in Exp.III. Therefore, some methodological aspects and remarks regarding the empirical research are worthy of discussion.

Firstly, regarding the test comparing preschool children with adult controls, difficulties and challenges mainly arise on the side of children. Besides the need to control the choice of words and events in that they are understandable for children during the test, another requirement which I want to discuss here already occurs during the test design: a balance between a representative set of items and children's limitations of cognitive resources has to be maintained. On the one hand, there should be an adequate number of items for each parameter combination so that the items can be considered representative for the latent linguistic population. According to the analysis by Mahowald et al. (2016), this number should be at least 5 – that is, at least 5 items for each parameter combination need to be rated in the test, so that the tendency, if any, can be assumed to be representative, at least to some extent. On the other hand, the amount of items should not exceed children's limitations of cognitive resources, so that they are able to follow the task during the whole test.

With these two requirements in mind, it is worth considering the previous studies on presupposition acquisition among children reviewed in Section 4.1 again. Considering their number of critical items, it seems that only some of them fulfill both requirements. One example is the study of Berger and Höhle (2012), who test the preschool children in their experiment with 5 items with *'auch' (also)* and 5 items without this trigger, totaling 10 items for each participant. On the contrary, several studies did not provide enough items: Dudley et al. (2015) presented only three trials for each combination of their two parameters, that is, verb (*think*/*know*) × negation (no/matrix/embedded), totaling 18 critical items per subject. Similarly, Aravind and Hackl (2017) presented four items for each condition type, generated by verb (*forget*/*remember*)× polarity (affirmative/negative), so that participants saw 16 critical items in total. In Yatsushiro (2008a), there were 5 critical items for each of the two presuppositions of *'jed-' (each)*, but only three testing the trigger *'beid-' (both)*. All these reductions of items might be due to the experiment design, namely, with more than one parameter that has more than one level, the multiple of the number of condition combinations is needed, so that the sum of all items could be too many for young children if every combination is tested with 5 items, especially if fillers are also used.

In Exp.I, we used a different strategy to achieve the balance. Instead of reducing the number of critical items, we excluded the filler items for children, based on the assumption that there are enough items to distract their attention and that their understanding of the task can be ensured by checking their answers to the three warm-up items. Moreover, during the test, we told the children that they could always ask for a break if they needed one, and as many times as they wanted. Pauses were also added if we observed that a child reacted obviously more slowly than before or could not focus on the task anymore. The majority of our preschool participants was able to finish the task with 30 items with no more than two breaks, some of them even without any pause. Still, there were 3 children who failed to finish the task, so that their ratings had to be excluded for the statistic analyses. Additionally, we also observed that even among the children who finished the task, some of them had concentration difficulties at the end.

Therefore, we are allowed to assume that even with breaks, rating 30 items is approaching the limitations of preschooler's cognitive resource or their maximum attention span. In other words, if enough items per condition combination and less pressure on children are both desired, there should ideally be no more than 5 combinations for a test involving preschool children. That is, for the item drafting, the design should not be more complex than 2×2, or if there is only one parameter, it should have no more than 5 levels. If this is not possible, as more levels or more parameters have to be tested, then enough breaks have to be included. Otherwise, more children have to be tested, if the experiment needs to be implemented with a reduced number of critical items.

Secondly, the perspective from L2 acquisition, or the contrast between native and nonnative speakers, is more useful and convenient for research on pragmatic features than has been expected and discussed so far.

Regarding the usefulness, take Exp.II for example. Different from previous studies focusing on classical teaching or examining strategies, this experiment offers crucial clues to the complexity of presuppositions and their heterogeneity. Regardless of whether the nonnative speakers' deviation from the native speakers is due to the acquisition of certain pragmatic skills or constraints in L2 and/or nonnative speakers' pragmatic tolerance, the deviation itself not only confirms the distinction of NRRCs from presuppositions, but also challenges the uniform account of presuppositions and scalar implicatures (see e.g. Chemla 2008; Romoli 2014). Of course, based only on Exp.II, and perhaps also on my previous work on scalar implicatures (Y. Chen 2019), these two explanations are still rather speculations. Thus, to verify their accuracy and influence, more research with L1-L2 contrast on various pragmatic phenomena is needed (for more discussion on this point, see Section 8.3 and 11.2 above).

The convenience of the L1-L2 contrast becomes especially obvious when comparing it with the adults-children test. In contrast to children at preschool or early school age, adult nonnative speakers can be generally assumed to have abundant cognitive resources and sufficient knowledge of the world and communication in general. Therefore, several constraints on the item design can be removed, such as the maximum number of items discussed above, or the careful choice of expressions, events and contexts that children are familiar with, or restrictions on the duration of the test and the presentation form of the items. Moreover, multiple impacts related to development processes at a younger age can also be excluded, as the participants are adults, while at the same time, the effects of acquisition can still be observed, although they can be different from those in L1 acquisition. It has to be emphasized that the L2-perspective or contrastive pragmatics involving nonnative speakers is not a replacement of L1 acquisition research, but rather a very important and convenient complementary strategy for empirical pragmatic research.

Note, however, that with a L2 group in the test, numerous effects need to be further controlled or considered, for instance the influence of their first language (see e.g. Taguchi 2013), their foreign language level, or their exchange experience and residence duration in the foreign country. Furthermore, potential differences and similarities between those societies or cultures, in which their L1 and L2 are used, can also affect their pragmatic interpretations or strategies, as has been observed in previous studies reviewed in Section 5.2. Thus, regarding the contrast between languages, not only the pragmatic differences are relevant, but the similarities and stable behaviors are meaningful as well, as they are the baseline for further contrasts.

The last point I want to shortly mention here is the equivalence between the two forms of clefts in Chinese: '是... 的 *(shi...de)*' and '... 的是 *(...de shi)*'. With reference to the controversy about whether they are both counterparts of *it*-clefts or the '... 的是 *(...de shi)*'-structure is rather comparable with pseudo-clefts (*wh*-clefts), proposals and claims from the syntactical perspective have already been briefly reviewed in Section 9.2. In Exp.III, these two structures were compared from a different perspective, namely in the context of presuppositions. Results show that both structures can trigger presuppositions with similar, in fact almost identical projection strength. Note, however, that this observation does not contribute to the discussion on whether the '... 的是 *(...de shi)*'-structure is rather compatible with pseudo-clefts (*wh*-clefts) or *it*-clefts, as both structures can trigger presuppositions (see e.g. Halvorsen 1978; Prince 1978; Atlas and Levinson 1981; Levinson 1983c).

Nevertheless, this observation can be helpful for further pragmatic studies on clefts in Chinese, as the two forms are differently suitable for marking elements in sentences, see e.g. (138e) and (140b), repeated below in (161). Moreover, as discussed in Section

9.2, they can also differ in their degree of naturalness and completeness of their structure once the sentence is embedded under operators like questions. Thus, the experiment confirms that at least with regard to presuppositions and information structure, they are similar or equivalent variations and both are comparable to clefts (no matter which one). Consequently, researchers have more options and flexibility during their item drafting.

(161) a. # 皮卡丘昨天在市中心看见了**是**喵喵**的**。

Pikaqiu zuotian zai shizhongxin kanjian le **shi** miaomiao **de**.
Pikachu yesterday in city-center see PST-particle **shi** Meowth **de**.
'It was Meowth whom Pikachu saw in the city center yesterday.'

b. 皮卡丘昨天在市中心看见了**的是**喵喵。

Pikaqiu zuotian zai shizhongxin kanjian le **de shi** miaomiao.
Pikachu Yesterday in city-center see PST-particle **de shi** Meowth.
'Whom Pikachu saw in the city center yesterday was Meowth'/'The person Pikachu saw in the city center yesterday was Meowth.'

In sum, based on the observations of all three experiments, some valuable and interesting methodological aspects can be summarized as follows:

- First, for tests involving young children, the balance between a representative number of critical items and the limitations of cognitive resources of children needs to be considered. When running Exp.I, we observed that a task with 30 items and two breaks can still be challenging for some children between 4 and 6 years of age. Assuming that each condition combination should have at least 5 items, for item drafting, a 2×2 design or a design with only one parameter that has no more than 5 levels is highly recommended.
- Second, the L2-perspective, or contrastive pragmatics involving nonnative speakers, is a useful and convenient complementary strategy for empirical linguistic research on pragmatic features. In order to control and investigate the effect of the variables in such L1-L2 contrasts, the comparison between both languages are very important as well, as differences can help to illustrate influences from the mother tongue of the L2 group, while similarities and stable behaviors can be the baseline for further comparisons.
- Third, at least with regard to presuppositions and information structure, the two different structures for clefts, namely '是... 的 *(shi...de)*' and '... 的是 *(...de shi)*', are similar or even equivalent. This similarity can offer researchers more options and flexibility for their item drafting.

12 Conclusion

Considering presuppositions as a phenomenon at the semantics-pragmatics interface, my work mainly concentrated on three main research concerns: the heterogeneity of presuppositions, which leads to the classification issue; the acquisition of presuppositions, which was investigated with regard to aspects from both first and second languages (L1 and L2); and the cross-linguistic comparison of presuppositions, which verified the stability of presupposition projection across languages. The three main research concerns were divided into five main research questions and then investigated in three experimental studies. A brief summary of relevant observations and interesting findings can be seen in the mind map of the dissertation in Figure 19 at the end of this chapter.

Traditionally, in semantic accounts, presuppositions are widely considered as truth value preconditions (see e.g. Frege 1892; Strawson 1950; Van Fraassen 1966; Heim 1991). While this notion fails to explain several observations, such as the filtering and heterogeneous projection behaviors of presuppositions shown in (21) and (22), pragmatic notions take the felicity of conversations and constraints on the common ground into account (see e.g. Stalnaker 1972, 1977, 1998 and summaries in Levinson 1983c; Beaver et al. 2021) and explain the variety or heterogeneity of presuppositions mainly by two algorithms: one branch proposes a uniform, conventional triggering process of presuppositions with certain cancellation mechanisms (e.g. Gazdar 1979a,b; Van der Sandt 1988; Simons et al. 2010; Tonhauser et al. 2018), whereas the other branch assumes different triggering processes between triggers, such as a conventional-conversational or semantic-pragmatic contrast, and claims the possibility and necessity of a classification of presuppositions and their triggers (see e.g. Abusch 2002; Charlow 2009; Abusch 2010; Romoli 2014, among others). Thus, addressing the first main research concern of my study, namely the heterogeneity and classification of presuppositions, two assumptions have been considered more relevant and were empirically investigated: the first one is the soft-hard split of triggers from the later branch, which is originally proposed by Abusch (2002, 2010) and used in a modified and wider sense in this work. The other one rather belongs to the former branch and explains the projection of presuppositions and their variation with regard to at-issueness (Simons et al. 2010; Tonhauser et al. 2018), a concept rooted in the question-driven discourse model (von Stutterheim and Klein 1989; Roberts 1996). The empirical studies aim at two particular questions: verifying and applying the classification, that is, whether the soft-hard split of triggers can be reflected in terms of their potential to express at-issue

contents, and whether the typical soft and hard triggers can be used as anchors to measure and classify other triggers.

The second main research concern focuses on presupposition acquisition, in particular on the acquisition process in both L1 and L2. As for L1 acquisition, comparing the findings of previous experiments (see Table 3 for a summary), the preschool children between 4 and 6 years of age are considered critical in order to study the developmental process. Thus, they were chosen as the critical group in the first experimental study (Exp.I, see Chapter 7). Moreover, in contrast to prior studies, in which children's awareness of presuppositions is mainly examined by means of the projection of presuppositions or speaker commitment, in Exp.I, the Not-At-Issueness Constraint on presuppositions proposed by Aravind and Hackl (2017), see (55), was used to test their sensitivity. Additionally, note that the L1 acquisition process is also a crucial aspect for the classification issue, because "if each presupposition trigger belongs either to the hard or to the soft category, how do children eventually figure out which box a specific expression should go into?" (Zehr and Schwarz 2018, p. 479). The L2 aspect is considered relevant and worth investigating mainly due to the following reasons: first, Carrell's (1984) experiment on presuppositions in L2 shows that presuppositions might not be acquired automatically in L2. Second, in my previous study on scalar implicatures (Y. Chen 2019), I observed that nonnative speakers might be more tolerant towards pragmatic violations than native speakers. Thirdly, since Carrell (1984), there has hardly been any empirical research on presuppositions with a L2-aspect. With reference to all these points and the potential weaknesses and limitations of Carrell's (1984) study, nonnative speakers were chosen as the critical group in the second experiment (Exp.II, see Chapter 8). It has been further expected that the L2-aspect can also shed light on the heterogeneity and classification concern.

The third main research concern targets the universality of presuppositions, that is, whether presuppositions' properties, like projection behaviors and their classification, are stable across languages. This issue needs to be investigated by means of cross-linguistic comparisons, which have rarely been made in recent studies and have mainly been restricted to European languages (see Amaral and Cummins 2015; Schwarz et al. 2020; Reins et al. 2021, among others). However, pragmatic principles or maxims might not be as general or neutral as has been assumed in traditional theories (see e.g. Grice 1989; Sperber and Wilson 1996 for their assumptions on the Cooperative Principle and the Cognitive Principle of Relevance, respectively), but could instead differ between Western and non-Western societies (see the review in Section 5.2). Therefore, the universality of presuppositions and their classification can not be just assumed either, but needs to be empirically verified, too.

Based on the theoretical and empirical research background, in order to address the three main research concerns, five main research questions were raised and listed in (85), repeated below again:

(162) a. The classification issue:
Research question 1: verifying the classification: can the soft-hard dichotomy, a classification based on the projective strength, also be reflected in terms of the ability of different triggers to express at-issue content?
Research question 2: applying the classification: choosing certain typical presuppositions or triggers as markers for certain properties, can they function like a thermometer that measures and sorts other triggers into certain subclasses?

b. The acquisition issue:
Research question 3: acquisition in first languages: are children aware of the difference between trigger types? Do children and adults respond differently once the presuppositions are forced to be at-issue? Are the different trigger types acquired at different ages?
Research question 4: acquisition in second languages: are nonnative speakers as sensitive as native speakers once the presuppositions of different triggers (or trigger types) are forced to be at-issue? If not, how can the difference be explained?

c. The universality issue:
Research question 5: cross-linguistic comparison of presuppositions: can a presupposition be triggered steadily across languages, as long as an adequate and equivalent translation of its trigger can be found? Are their properties like projection behavior or information status independent from linguistic and cultural influence?

These five main research questions were investigated by means of three experiments. Exp.I targets Research question 1 and 3, testing the soft-hard split from the perspective of at-issueness and L1 acquisition. Thus, it had a $3 \times 2 \times 2$ design with the factors TRIGGER (non-restrictive relative clauses (NRRCs) vs. hard trigger vs. soft trigger; between-item and within-subject), ISSUENESS (assertion at-issue vs. presuppositions/NRRCs at-issue; within-item and within-subject), and AGE (children vs. adults; within-item and between-subject). In this experiment, a modified acceptability judgment task with video stimuli was used, and the comparison was made between German native adult controls and monolingual preschool children between 4 and 6 years of age.

The observations show firstly that the heterogeneous projection behavior of presuppositions, which is assumed to be due to a soft-hard trigger distinction by several researchers

(see Section 3.1), can indeed be reflected in terms of their trigger's ability to express at-issue contents, namely via local accommodation. This relation contributes to the debate on the heterogeneity of presuppositions by illustrating that in order to explain the difference within the category of presuppositions, the main assumptions from both pragmatic branches can be or even need to be combined.

Secondly, regarding the acquisition issue, the results illustrate that preschoolers are not only able to understand the backgroundedness of presuppositions and the at-issueness violations, but also are generally aware of the difference in at-issueness between trigger types, although they are still less sensitive than adults. These findings confirm that the preschool age between 4 and 6 years is indeed a crucial stage for the acquisition of pragmatic skills and pragmatic sensitivity in L1, including but not restricted to presuppositions. Moreover, these observations may also provide empirical evidence from L1 acquisition for the distinction or classification of presupposition triggers. That is, already at preschool age, speakers are aware of the heterogeneity within the category of presuppositions, and they are in the midst of the process of "figur[ing] out which box a specific expression should go into" (Zehr and Schwarz 2018, p. 479).

Methodologically, considering that for some preschool children, a task with 30 items could be challenging or perhaps exceed the limitation of their cognitive resources, as was observed in Exp.I. On the other hand, a representative amount of critical items requires at least 5 items per combination of each level of the parameters, as suggested by Mahowald et al. (2016). In order to achieve a balance between these two conditions, for the item drafting, a design including two parameters with two levels each, i.e. a 2×2 design, or with only one parameter that has no more than 5 levels is recommended.

Addressing Research question 1 and 4, in Exp.II, the same design and materials from Exp.I were applied to another group of participants: adult Chinese learners of German with an advanced level of German language. Thus, the soft-hard split of presupposition triggers and their at-issueness was explored from a L2-perspective, that is, by means of a comparison between German native adults, the same control group as in Exp.I, and nonnative speakers.

Regarding the classification issue, the results of Exp.II show that the relation between the soft-hard split of triggers and their different ability to express at-issue contents can also be perceived by nonnative speakers, confirming the possibility or the necessity of combining the two pragmatic accounts again.

Regarding presuppositions in L2 acquisition, the ratings illustrate that the nonnative speakers are less sensitive to or more tolerant towards the pragmatic violations caused by at-issue presuppositions or NRRCs than the native speakers. This discrepancy might have to do with the acquisition process of pragmatic competence in L2 and/or nonna-

tive speakers' pragmatic tolerance, analogical to children's pragmatic tolerance in first language acquisition (Katsos and Bishop 2011). Moreover, note that such insensitivity or tolerance was mainly observed for the hard triggers, but not among the soft triggers, which are assumed to be more pragmatic in some theories (see e.g. Abusch 2002; Chemla 2008; Abusch 2010; Romoli 2014), nor among NRRCs, which might not exist in the participants' L1, Chinese, at least not in the same syntactical form (see e.g. N. Zhang 2001; Del Gobbo 2005; Lin and Tsai 2015; Del Gobbo 2017). This contrast provides interesting clues to the distinction between presuppositions of soft triggers and scalar implicatures on the one hand, and the boundary between presuppositions of hard triggers and NRRCs on the other. More importantly, the observations methodologically highlight the valuable contribution to presupposition research offered by the L2-perspective. That is, besides being important in the traditional contexts of developing L2 teaching and learning strategies, the native-nonnative comparison is both relevant and useful for research on classical issues like delimitation and classification in the pragmatic area, too.

Research question 2 and 5 were mainly investigated by means of the third experiment (Exp.III, see Chapter 9), which was split into two sub-experiments. The first sub-experiment compared four triggers between two very different languages: Chinese and German, and thus had a 4×2 design with the factors TRIGGER (clefts vs. *win* vs. *regret* vs. *discover*, within subjects and between items) and LANGUAGE (Chinese vs. German, within items and between subjects). Of the four triggers, clefts are considered to be typical hard and *win* to be typical soft. They are the anchors for measuring in the test. Using an online paper-and-pencil questionnaire, the projection strength of the presuppositions was rated by assessing the speaker's certainty, adopting the certainty judgment task used in Tonhauser et al. (2018).

The results show that first, once equivalent counterparts of the triggers can be found in both languages and the translation is properly controlled, the cross-linguistically stable projection behavior of their presuppositions can also be expected. Moreover, the significant difference between the soft and hard anchors is also stable between Chinese and German. Thus, the question about the universality of presuppositions and their properties can be partially answered: the presuppositions and their projection strength is cross-linguistically stable – at least between German and Chinese – as long as equivalent counterparts can be found in both languages and no pragmatic features that may cause language specific variations, such as politeness, are involved.

Second, measuring projection strength of a presupposition with certain triggers as anchors or markers seems possible. However, while the emotive factive verb *regret* was rated as highly projective and therefore clearly hard, which is in line with the expectations in previous studies (see e.g. predictions in Simons 2001; Abbott 2006; Simons 2007; Abusch

2010), this was not the case with the cognitive factive verb *discover*, which is generally assumed to be soft (see e.g. Karttunen 1971b; Hooper and Thompson 1973; Abusch 2002, 2010): the ratings did not pattern with the soft anchor *win* but rather with *regret* and the hard anchor clefts in both languages. Combining this result with the observations made with regard to the soft triggers in Exp.I and Exp.II, the variation of *discover* and its contrast with *win* possibly illustrate that the soft triggers are more heterogeneous than has been predicted so far: while the hard triggers are similarly hard, each soft trigger might be soft in its own way. In light of this, it seems reasonable to reconsider the soft-hard distinction in the sense that it perhaps rather describes a distinction between less typical and typical presuppositions or triggers, and the less-typicality can result from several distinct properties or factors.

The second sub-experiment investigated the two different structures of clefts in Chinese, '是... 的 *(shi...de)*' and '... 的是 *(...de shi)*', of which the former is a widely accepted counterpart for *it*-cleft, while the later is rather considered to be a counterpart for pseudo-cleft (*wh*-cleft). The results prove that they can both trigger presuppositions with a similar high projectivity despite of their syntactical differences. As these two structures can only mark focus on restricted positions, which might lead to difficulties during the item drafting, this comparison contributes mainly methodologically to further empirical studies on presuppositions or information structure involving Chinese.

In sum, the empirical data and the assumptions based on them (see also discussion in Chapter 11) not only contribute to the debate on the heterogeneity of presuppositions with evidence from both language acquisition and cross-linguistic comparison, but they also highlight the relevance and usefulness of contrastive pragmatics and provide a cross-linguistic baseline for intercultural pragmatic studies in the future. Moreover, they also provide further methodological indications for the item drafting, especially for experiments involving children as participants.

Unsurprisingly, there still remain some open questions in my study, such as whether hard and soft triggers are acquired in different ways and at different ages, or whether at-issueness in general is acquired before the trigger distinction, as discussed in Chapter 7. Furthermore, if the soft-hard distinction indeed represents a contrast between less typical and typical presuppositions, how many criteria need to be taken into account and how can they be combined? As for L2 acquisition, it remains unclear why the nonnative speakers' pragmatic sensitivity did extremely (at least numerically) differ from that of the native speakers with regard to triggers from A1-level, which should have been learned at the very beginning of their L2 courses? In particular, why did this deviation also occur in question-answer pairs without an at-issueness violation? Moreover, in contrastive pragmatic studies, not limited to native-nonnative and cross-linguistic comparisons, how can

we achieve a balance between controlling and comparing? Regarding the cross-linguistic comparison, which further elements could be influential and restrict the universality of presuppositions? All these questions, and many more that are not mentioned here, require more empirical and theoretical research on presuppositions in the future.

I will end my conclusion with a minor but interesting question that relates mainly to the last open question mentioned above but may also target all three main issues of my research, by reconsidering the very first example of this work. The sentence uttered by Louise in (1) and its presuppositions (2a)-(2c) are repeated below again:

(163) Louise (mumble): I just realize why my husband left me.[...]

a. Louise had a husband, i.e. she was married.

b. The husband was with her before.

c. The husband then left her.

Now, imagine that her addressee was not Ian but the two aliens (heptapods) who have a non-linear time concept, or even the sentence is not in English but in their language, Heptapod. Would there still be an information clash? That is, would some propositions remain presupposed while they would be considered neither given nor could be assumed in the context? If so, can this difference give another clue to the classification issue or provide another criterion for the typical-less typical distinction? Or do we have to admit that none of these propositions in (163a)-(163c) can be interpreted as a presupposition anymore because if there is no linear conception of time, then generally, nothing needs to or can be *pre*-supposed? If so, which conditions or constraints on language, speaker, culture etc. should be further considered, so that the concept *presupposition* itself can exist in a certain language? In other words, what do we actually presuppose when we talk about presuppositions?

Considering that our understanding of the universe and humans, and of language and communication will continue to develop, presuppositions will remain a highly interesting phenomenon at the semantics-pragmatics interface and will perhaps become more research-demanding in the future, while my dissertation may be a small step of our long journey towards further investigations on them.

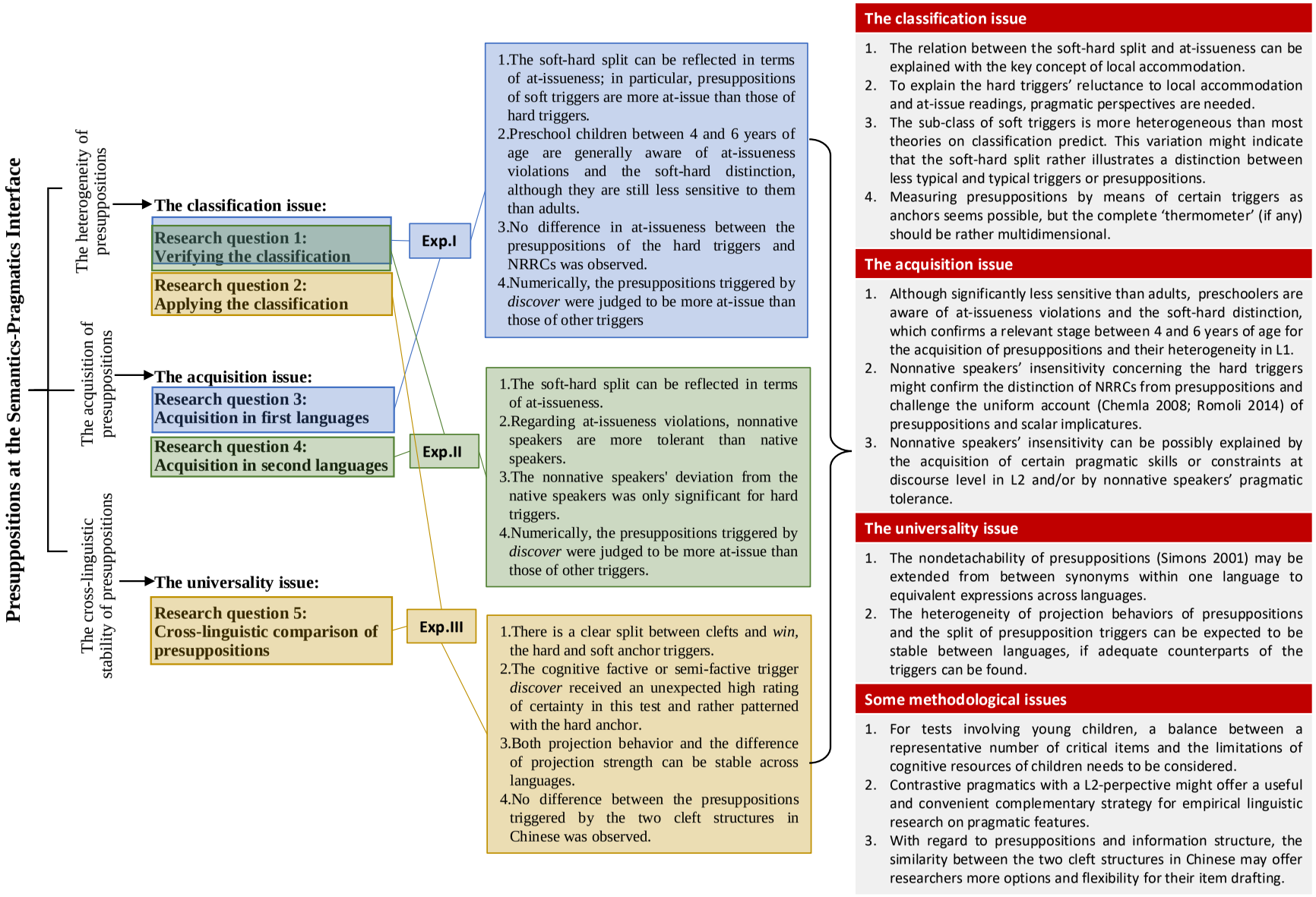


Figure 19: The mind map of the dissertation.

Bibliography

Abbott, Barbara. (2000). Presuppositions as nonassertions. *Journal of Pragmatics*, *32*(10), 1419–1437.

Abbott, Barbara. (2006). Where have some of the presuppositions gone. In B. Birner & G. Ward (Eds.), *Drawing the boundaries of meaning: Neo-gricean studies in pragmatics and semantics in honor of laurence r. horn* (pp. 1–20). Amsterdam: John Benjamins Publishing.

Abrusán, Márta. (2011). Predicting the presuppositions of soft triggers. *Linguistics and Philosophy*, *34*(6), 491–535.

Abrusán, Márta. (2016). Presupposition cancellation: explaining the 'soft‑hard' trigger distinction. *Natural Language Semantics*, *24*(2), 165–202.

Abusch, Dorit. (2002). Lexical alternatives as a source of pragmatic presuppositions. *Semantics and linguistic theory*, *12*, 1–19.

Abusch, Dorit. (2010). Presupposition triggering from alternatives. *Journal of Semantics*, *27*(1), 37–80.

Amaral, Patrícia, & Cummins, Chris. (2015). A Cross-Linguistic Study on Information Backgrounding and Presupposition Projection. In F. Schwarz (Ed.), *Experimental perspectives on presuppositions* (pp. 157–172).

Ameka, Felix K. (2006). "When I die, don't cry": the ethnopragmatics of "gratitude" in West African languages. In C. Goddard (Ed.), *Ethnopragmatics* (Chap. 8, pp. 231–266).

Ameka, Felix K. (2017). Meaning between algebra and culture. In H. Reckman, L. L. Cheng, M. Hijzelendoorn, & R. Sybesma (Eds.), *Crossroads semantics: Computation, experiment and grammar* (Chap. 14, pp. 227–248). Amsterdam: John Benjamins Publishing Company.

Ameka, Felix K., & Terkourafi, Marina. (2019). What if…? Imagining non-Western perspectives on pragmatic theory and practice. *Journal of Pragmatics*, *145*, 72–82.

Antomo, Mailin. (2012). Projective meaning and the licensing of Embedded Root Phenomena. *Proceedings of ConSOLE XIX*, *1*, 1–23.

Antomo, Mailin. (2015). *Abhängige Sätze in einem fragebasierten Diskursmodell* (Dissertation, Georg-August-Universität Göttingen).

Antomo, Mailin. (2016). Marking (not-) at-issue content by using verb order variation in German. *Co-and subordination in German and other languages*, 21–54.

Antomo, Mailin, Chen, Yuqiu, & Thalmann, Maik. (2021). (Un)Selbstständigkeit von Sätzen und Main Point of Utterance: Appositive Relativsätze und deren Erwerb. In R. Külpmann & R. Finkbeiner (Eds.), *Neues zur selbständigkeit von sätzen. linguistische berichte, sonderheft 30* (pp. 257–279). Hamburg: Helmut Buske Verlag.

Antomo, Mailin, Müller, Susanne, Paul, Katharina, Paluch, Markus, & Thalmann, Maik. (2018). When children aren't more logical than adults: An empirical investigation of lying by falsely implicating. *Journal of Pragmatics*, *138*, 135–148.

Aravind, Athulya, Freedman, Eva, Hackl, Martin, & Wexler, Ken. (2016). Subject-object asymmetries in the acquisition of clefts. *Proceedings of the 40th Boston University Conference on Language Development*, 1–17.

Aravind, Athulya, & Hackl, Martin. (2017). Factivity and at-issueness in the acquisition of forget and remember Factive predicates. *Proceedings of the 41st annual Boston University Conference on Language Development*, 46–59.

Aravind, Athulya, Hackl, Martin, & Wexler, Ken. (2018). Syntactic and Pragmatic Factors in Children's Comprehension of Cleft Constructions. *Language Acquisition*, *25*(3), 284–314.

Atlas, Jay D, & Levinson, Stephen C. (1981). It-Clefts, Informativeness and Logical Form: Radical Pragmatics. Academic Press.

Austin, John Langshaw. (1962). *How to do things with words*.

Ayodele, Ayo. (2016). OMOLUWABI: Towards a theory of cooperation in Yoruba interactional discourse practices. In K. Traore, M. Solunsa, & A. Ojo (Eds.), *Expressions of indigenous and local knowledge in africa and its diaspora* (Chap. 6, pp. 80–90). Newcastle Upon Tyne: Cambridge Scholars Publishing.

Bachman, Lyle F. (1990). *Fundamental considerations in language testing*. New York: Oxford University Press.

Bachman, Lyle F, & Palmer, Adrian S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

Bade, Nadine. (2016). *Obligatory Presupposition Triggers in Discourse: Empirical Investigations of the Theories Maximize Presupposition and Obligatory Implicatures* (Doctoral dissertation, University of Tübingen).

Barr, Dale J, Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, Douglas, Kliegl, Reinhold, Vasishth, Shravan, & Baayen, Harald. (2015). Parsimonious Mixed Models. *arXiv preprint arXiv:1506.04967*. arXiv: 1506.04967

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Beaver, David. (2012). Anti-matters. *'Questions in Discourse' during DGfS Jahrestagung*, 7–9.

Beaver, David, Geurts, Bart, & Denlinger, Kristie. (2021). Presupposition. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (SP.2021). Stanford: Metaphysics Research Lab, Stanford University.

Beaver, David, & Krahmer, Emiel. (2001). A partial account of presupposition projection. *Journal of Logic, Language and Information*, *10*(2), 147–182.

Berger, Frauke, & Höhle, Barbara. (2012). Restrictions on addition: Children's interpretation of the focus particles auch 'also' and nur 'only' in German. *Journal of Child Language*, *39*(2), 383–410.

Bergsma, Wenda. (2006). (Un)Stressed ook in Child Dutch. In V. van Geenhoven (Ed.), *Semantics in acquisition* (pp. 329–348).

Bernicot, Josie, Laval, Virginie, & Chaminaud, Stéphanie. (2007). Nonliteral language forms in children: In what order are they acquired in pragmatics and metapragmatics? *Journal of Pragmatics*, *39*(12), 2115–2132.

Bill, Cory, Romoli, Jacopo, & Schwarz, Florian. (2018). Processing Presuppositions and Implicatures: Similarities and Differences. *Frontiers in Communication*, *3*.

Bill, Cory, Romoli, Jacopo, Schwarz, Florian, & Crain, Stephen. (2016). Scalar Implicatures Versus Presuppositions: The View from Acquisition. *Topoi*, *35*(1), 57–71.

Bott, Lewis, & Noveck, Ira A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, *51*(3), 437–457.

Brown, Penelope, & Levinson, Stephen C. (1987). *Politeness: Some Universals in Language Usage*.

Büring, Daniel, & Hartmann, Katharina. (2001). The syntax and semantics of focus-sensitive particles in german. *Natural Language & Linguistic Theory*, *19*(2), 229–281.

Carrell, Patricia. (1978). The given-new strategy: testing presupposition and assertion in L1 and L2. *On TESOL*, *78*.

Carrell, Patricia. (1984). Inferencing in Esl: Presuppositions and Implications of Factive and Implicative Predicates. *Language Learning*, *34*(1), 1–19.

Charlow, Simon. (2009). "Strong" predicative presuppositional objects. *Proceedings of ESSLLI*, *109*.

Chemla, Emmanuel. (2008). Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. *Unpublished Manuscript, ENS and MIT*, 1–67.

Chemla, Emmanuel. (2009). Presuppositions of quantified sentences: experimental data. *Natural Language Semantics*, *17*(4), 299–340.

Chemla, Emmanuel, & Bott, Lewis. (2013). Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and cognitive processes*, *28*(3), 241–260.

Chen, Rong. (2010). Pragmatics east and west: Similar or different? In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (Chap. 5, pp. 167–188).

Chen, Xinren. (2019). 'Family-culture' and Chinese politeness: An emancipatory pragmatic account. *Acta Linguistica Academica*, *66*(2), 251–270.

Chen, Yuqiu. (2019). An experimental study on scalar implicatures: Comparing German native speakers and Chinese learners of German. *Proceedings of ConSOLE XXVI*, *1*, 116–138.

Chen, Yuqiu. (2022). Wenn Nicht-Muttersprachler/innen mehr Toleranz zeigen: Eine experimentelle Studie zu Präsuppositionen, At- issueness und DaF. In L. Auteri, N. Barrale, D. B. Arianna, & S. Hoffmann (Eds.), *Wege der Germanistik in transkultureller Perspektive (Bd. 6). Jahrbuch für Internationale Germanistik Beihefte* (pp. 213–227).

Chen, Yuqiu, & Antomo, Mailin. (to appear). Presuppositions cross-linguistically: A comparison of soft and hard triggers in Chinese and German. *Intercultural Pragmatics*.

Chen, Yuqiu, Thalmann, Maik, & Antomo, Mailin. (2022). Presupposition triggers and (not-)at-issueness: Insights from language acquisition into the soft-hard distinction. *Journal of Pragmatics*, *199*, 21–46.

Cheng, Lisa Lai-shen. (2008). Deconstructing the shì ... De construction. *The Linguistic Review*, *25*, 235–266.

Chiang, Ted. (1998). Story of Your Life. In P. N. Hayden (Ed.), *Starlight 2*. New York: Tor Books.

Chierchia, Gennaro, Crain, Stephen, Guasti, Maria Teresa, Gualmini, Andrea, & Meroni, Luisa. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. *Proceedings of the 25th Boston University conference on language development*, 157–168.

Chierchia, Gennaro, Guasti, Maria Teresa, Gualmini, Andrea, Meroni, Luisa, Crain, Stephen, & Foppolo, Francesca. (2004). Semantic and pragmatic competence in children and adults' comprehension of Or. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 283–300).

Chierchia, Gennaro, & McConnell-Ginet, Sally. (2000). *Meaning and Grammar. An Introduction to Semantics*. Cambridge, Massachusetts: The MIT press.

Chomsky, Noam. (1972). *Studies on Semantics in Generative Grammar*. The Hague: Mouton.

Christensen, R H B. (2019). ordinal – Regression Models for Ordinal Data.

Cibelli, Emily, Xu, Yang, Austerweil, Joseph L, Griffiths, Thomas L, & Regier, Terry. (2016). The Sapir-Whorf hypothesis and probabilistic inference: Evidence from the domain of color. *PloS one*, *11*(7), e0158725.

Clark, Herbert. (1977). Comprehension and the given-new contract. *Discourse production and comprehension*, 1–40.

Costa, Albert, Foucart, Alice, Hayakawa, Sayuri, Aparici, Melina, Apesteguia, Jose, Heafner, Joy, & Keysar, Boaz. (2014). Your morals depend on language. *PLoS ONE*, *9*(4), e94842.

Council of Europe, Council for Cultural Co-operation. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment.*

Crain, Stephen, & McKee, Cecile. (1985). The acquisition of structural restrictions on anaphora. *North Eastern Linguistic Society (NELS)*, *15*(1), 94–110.

Culpeper, Jonathan, Mackey, Alison, & Taguchi, Naoko. (2018). *Second language pragmatics: From theory to research.*

Cummins, Chris, Amaral, Patricia, & Katsos, Napoleon. (2013). Backgrounding and accommodation of presuppositions: an experimental approach. *Proceedings of Sinn und Bedeutung 17*, *17*, 201–218.

Cummins, Chris, Amaral, Patrícia, & Katsos, Napoleon. (2012). Experimental Investigations of the Typology of Presupposition Triggers. *Humana.Mente Journal of Philosophical Studies*, *23*, 1–15.

De Kadt, Elizabeth. (1998). The concept of face and its applicability to the Zulu language. *Journal of Pragmatics*, *29*(2), 173–191.

Del Gobbo, Francesca. (2005). Chinese relative clauses: restrictive, descriptive or appositive? In L. Brugè, G. Giusti, N. Munaro, & W. Schweikert (Eds.), *The thirtieth incontro di grammatica generativa* (pp. 207–305). Venezia: Libreria Editrice Cafoscarina.

Del Gobbo, Francesca. (2010). On Chinese appositive relative clauses. *Journal of East Asian Linguistics*, *19*(4), 385–417.

Del Gobbo, Francesca. (2017). More appositives in heaven and earth than are dreamt of in your Linguistics. *Glossa*, *2*(1), 1–38.

Dingemanse, Mark, Blythe, Joe, & Dirksmeyer, Tyko. (2014). Formats for other-initiation of repair across languages. *Studies in Language*, *38*(1), 5–43.

Dingemanse, Mark, & Enfield, N. J. (2015). Other-initiated repair across languages: Towards a typology of conversational structures. *Open Linguistics*, *1*(1), 96–118.

Djärv, Kajsa, Zehr, Jérémy, & Schwarz, Florian. (2018). Cognitive vs. emotive factives: An experimental differentiation. *Proceedings of Sinn und Bedeutung*, *21*(1), 367–386.

Dudley, Rachel, Orita, Naho, Hacquard, Valentine, & Lidz, Jeffrey. (2015). Three-Year-Olds' Understanding of Know and Think. In F. Schwarz (Ed.), *Experimental perspectives on presuppositions* (pp. 241–262). Cham: Springer.

Egré, Paul. (2008). Question-embedding and factivity. *Grazer Philosophische Studien*, *77*(1), 85–125.

Fillmore, Charles J. (1971). Verbs of judging: An exercise in semantic description. In C. J. Fillmore & D. T. Langendoen (Eds.), *Studies in linguistic semantics* (pp. 273–289). New York: Holt, Rinehart and Winston.

Finnegan, Ruth. (1969). How to do things with words: performative utterances among the Limba of Sierra Leone. *Man*, *4*(4), 537–552.

Forstmeier, Wolfgang, & Schielzeth, Holger. (2011). Cryptic multiple hypotheses testing in linear models : overestimated effect sizes and the winner's curse. *Behavioral ecology and sociobiology*, *65*(1), 47–55.

Frege, Gottlob. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, *100*(1), 25–50.

Frege, Gottlob. (1948). Sense and Reference. *The Philosophical Review*, *57*(3), 209–230.

Gazdar, Gerald. (1978). Heavy Parentheses Wipe-Out Rules,Okay? *Linguistics and Philosophy*, *2*(2), 281–289.

Gazdar, Gerald. (1979a). A Solution to the Projection Problem. In C.-K. Oh & D. A. Dinneeen (Eds.), *Syntax and semantics 11: Presupposition* (pp. 57–89).

Gazdar, Gerald. (1979b). *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.

Geipel, Janet, Hadjichristidis, Constantinos, & Surian, Luca. (2015). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, *59*, 8–17.

Glaboniat, Manuela, Perlmann-Balme, Michaela, & Studer, Thomas. (2013). *Zertifikat B1 - Prüfungsziele, Testbeschreibung : Deutschprüfung für Jugendliche und Erwachsene. Deutsch als Fremdsprache* (1st ed.). Hueber Verlag.

Glanzberg, Michael. (2005). Presuppositions, truth values, and expressing propositions. In G. Preyer & G. Peter (Eds.), *Contextualism in philosophy: Knowledge, meaning, and truth* (pp. 349–396). Oxford: Oxford University Press.

Goebel, Alexander. (2020). *Representing Context: Presupposition Triggers and Focus-sensitivity* (Doctoral dissertation, University of Massachusetts Amherst).

Grice, Paul. (1989). Logic and Conversation. In *Studies in the way of words* (pp. 22–40). Cambridge, Massachusetts: Harvard University Press.

Gu, Yueguo. (1990). Politeness phenomena in modern Chinese. *Journal of Pragmatics*, *14*(2), 237–257.

Halvorsen, Per-Kristian. (1978). *The Syntax and Semantics of Cleft Constructions*.

Harris, Richard J. (1975). Children's comprehension of complex sentences. *Journal of Experimental Child Psychology*, *19*(3), 420–433.

Heim, Irene. (1982). *The semantics of definite and indefinite NPs* (Doctoral dissertation, University of Massachusetts Amherst).

Heim, Irene. (1983). On the projection problem for presuppositions. In M. Barlow, D. Flickinger, & M. Wescoat (Eds.), *Proceedings of wccfl-2* (pp. 114–125).

Heim, Irene. (1991). Artikel und definitheit. In A. von Stechow & D. Wunderlich (Eds.), *Semantik/semantics: Ein internationales handbuch der zeitgenössischen forschung. an international handbook of contemporary research* (Chap. 22, pp. 487–535). Berlin: De Gruyter Mouton.

Holler, Anke. (2005). *Weiterführende Relativsätze: Empirische und theoretische Aspekte.* Berlin: Akademie Verlag.

Holler, Anke. (2013). d- und w-Relativsätze. In J. Meibauer, M. Steinbach, & H. Altmann (Eds.), *Satztypen des deutschen* (pp. 266–300). Berlin/Boston: Walter de Gruyter.

Holler, Judith, Kendrick, Kobin, Casillas, Marisa, & Levinson, Stephen C. (2016). *Turn-Taking in Human Communicative Interaction.*

Hooper, Joan, & Thompson, Sandra. (1973). On the Applicability of Root Transformations. *Linguistic Inquiry*, *4*(4), 465–497.

Hopmann, Marita R., & Maratsos, Michael P. (1978). A developmental study of factivity and negation in complex syntax. *Journal of Child Language*, *5*(2), 295–309.

Horn, Laurence. (1969). A presuppositional analysis of only and even. *The Fifth Regional Meeting of the Chicago Linguistics Society*, 98–107.

Horn, Laurence. (1976). *On the semantic properties of logical operators in English.* Bloomington: Indiana Univ.

Horn, Laurence. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42). Washington: Georgetown University Press.

Horn, Laurence. (1990). Hamburgers and truth: Why Gricean explanation is Gricean. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, *16*(1), 453–71.

Horn, Laurence. (2004). Implicature. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (Chap. 1, pp. 3–27). Oxford: Blackwell Publishing.

Horton, Diane, & Hirst, Graeme. (1988). Presuppositions as beliefs. *Proceedings of the International Conference on Computational Linguistics*, 255–260.

Huang, Cheng-Teh James. (1982). *Logical relations in Chinese and the theory of grammar* (Doctoral dissertation, Massachusetts Institute of Technology).

Huang, Cheng-Teh James. (1988). 說 [是] 和 [有][On 'Be' and 'Have' in Chinese]. *The Bulletin of the Institute of History and Philology*, *59*, 43–64.

Huang, Yi Ting, Spelke, Elizabeth, & Snedeker, Jesse. (2013). What Exactly do Numbers Mean? *Language Learning and Development*, *9*(2), 105–129.

Hussein, Basel Al-sheikh. (2012). The Sapir-Whorf Hypothesis Today. *Theory and Practice in Language Studies*, *2*(3), 642–646.

Ide, Sachiko. (1989). Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. *Multilingua*, *8*(2-3), 223–248.

Ito, Katsumasa. (2017). The presupposition of exclamatives at the syntax-semantics interface: Evidence from German and Japanese. *ConSOLE XXV*, 108–122.

Jacobsson, Bengt. (1994). Nonrestrictive Relative That-Clauses Revisited. *Studia Neophilologica*, *66*(2), 181–195.

Jayez, Jacques, Mongelli, Valeria, Reboul, Anne, & van der Henst, Jean-Baptiste. (2015). Weak and Strong Triggers. In *Experimental perspectives on presuppositions* (pp. 173–193).

Ji, Anfeng(季安锋). (2009). 汉语预设触发语研究 (Doctoral dissertation, 南开大学).

Jia, Mian, & Yang, Guoping. (2021). Emancipating Chinese (im)politeness research: Looking back and looking forward. *Lingua*, *251*, 103028.

Johnson, Carl Nils, & Maratsos, Michael P. (1977). Early Comprehension of Mental Verbs : Think and Know. *Child development*, *48*(4), 1743–1747.

Jung, Ji-Young. (2002). Issues in acquisitional pragmatics. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, *2005*, 1–34.

Karttunen, Lauri. (1971a). Implicative Verbs. *Language*, *47*(2), 340.

Karttunen, Lauri. (1971b). Some observations on factivity. *Paper in Linguistics*, *4*(1), 55–69.

Karttunen, Lauri. (1973). Presuppositions of compound sentences. *Linguistic inquiry*, *4*(2), 169–193.

Karttunen, Lauri. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, *1*(1-3), 181–194.

Karttunen, Lauri. (2016). Presupposition: What went wrong? *Semantics and Linguistic Theory*, *26*, 705.

Karttunen, Lauri, & Peters, Stanley. (1975). Conversational Implicature in Montague Grammar. *Annual Meeting of the Berkeley Linguistics Society*, 266–278.

Katsos, Napoleon, & Bishop, Dorothy V.M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67–81.

Katz, Jerrold. (1972). *Semantic theory*. New York: Harper & Row.

Katz, Jerrold. (1973). On Defining "Presupposition". *Linguistic Inquiry*, *4*(2), 256–260.

Katz, Jerrold, & Langendoen, Terence. (1976). Pragmatics and presupposition. *Language*, 1–17.

Kay, Paul, & Kempton, Willett. (1984). What is the Sapir-Whorf hypothesis? *American anthropologist*, *86*(1), 65–79.

Keenan, Elinor Ochs. (1976). The universality of conversational postulates. *Language in society*, *5*(1), 67–80.

Kempson, Ruth. (1975). *Presupposition and the Delimitation of Semantics* (1st ed.). Cambridge: Cambridge University Press.

Kiparsky, Paul, & Kiparsky, Carol. (1970). FACT. In M. Bierwisch & K. E. Heidolph (Eds.), *Progress in linguistics: A collection of papers* (pp. 143–173).

Klein, Wolfgang, & Von Stutterheim, Christiane. (1987). Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte*, *109*, 163–183.

Klinedinst, Nathan. (2016). Two types of semantic presuppositions. In K. Allan, A. Capone, & I. Kecskes (Eds.), *Pragmemes and theories of language use* (pp. 601–624).

Kondowe, Wellman, Ngwira, Flemmings Fishani, & Madula, Precious. (2014). Linguistic Analysis of Malawi Political Newspaper Cartoons on President Joyce Banda: Towards Grice's Conversational Implicature. *International Journal of Humanities and Social Science*, *4*(7), 401–451.

Kratzer, Angelika, & Heim, Irene. (1998). More of English: Nonverval Predicates, Modifiers, Definite Descriptions. In *Semantics in generative grammar* (Chap. 4, pp. 61–85). Blackwell Oxford.

Kripke, Saul A. (2009). Presupposition and Anaphora: Remarks on the Formulation of the Projection Problem. *Linguistic Inquiry*, *40*(3), 367–386.

Lakoff, George. (1971). Presupposition and relative well-formedness. In D. D. Steinberg, & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology* (pp. 329–340). Cambridge: Cambridge University Press.

Langendoen, Donald, & Savin, Harris B. (1971). The projection problem for presuppositions. In C. Fillmore & T. Langendoen (Eds.), *Studies in linguistic semantics* (pp. 373–388). New York: Holt, Reinhardt and Winston.

Lenth, Russell V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.0.

Levinson, Stephen C. (1983a). Conversational implicature. In *Pragmatics* (Chap. 3, pp. 97–162). Cambridge: Cambridge University Press.

Levinson, Stephen C. (1983b). *Pragmatics* (1st ed.). Cambridge: Cambridge University Press.

Levinson, Stephen C. (1983c). Presupposition. In *Pragmatics* (Chap. 4, pp. 167–225). Cambridge: Cambridge University Press.

Levinson, Stephen C. (2000). On the Notion of a Generalized Conversational Implicature. In *Presumptive meanings: The theory of generalized conversational implicature* (Chap. 1, pp. 11–72). Cambridge, Mass.: MIT Press.

Lewis, David. (1979). Scorekeeping in a Language Game. *Journal of Philosophical Logic*, 339–359.

Li, Mingyang, Hickman, Louis, Tay, Louis, Ungar, Lyle, & Guntuku, Sharath Chandra. (2020). Studying Politeness across Cultures using English Twitter and Mandarin Weibo. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–15.

Liao, Chao Chih, & Bresnahan, Mary I. (1996). A contrastive pragmatic study on American English and Mandarin refusal strategies. *Language Sciences*, *18*(3-4), 703–727.

Liddell, Torrin M., & Kruschke, John K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*(August), 328–348.

Lin, Jo-Wang, & Tsai, Wei-Tien Dylan. (2015). Restricting Non-restrictive Relatives in Mandarin Chinese. *Chinese Syntax in a Cross-Linguistic Perspective*, 100–127.

Liu, Wei, & Kempson, Ruth. (2018). Chinese cleft structures and the dynamics of processing. *Transactions of the Philological Society*, *116*(1), 91–116.

Lyons, John. (1977). *Semantics: Volume 1 & 2*. Cambridge: Cambridge University Press.

Macnamara, John, Baker, Erica, & Olson, Chester L. (1976). Four-Year-Olds' Understanding of "Pretend", "Forget", and "Know": Evidence for Propositional Operations. *Child Development*, *47*(1), 62–70.

Mahowald, Kyle, Graff, Peter, Hartman, Jeremy, & Gibson, Edward. (2016). SNAP judgments: A small N acceptability paradigm (SNAP) for linguistic acceptability judgments. *Language*, *92*(3), 619–635.

Mao, Lu Ming Robert. (1994). Beyond politeness theory: 'Face' revisited and renewed. *Journal of Pragmatics*, *21*(5), 451–486.

Marcu, Daniel. (1994). *A Formalism and an Algorithm for Computing Pragmatic Inferences and Detecting Infelicities* (Doctoral dissertation, University of Toronto).

Matsumoto, Yoshiko. (1989). Politeness and conversational universals —. *Multilingua*, *8*(2/3), 207–221.

Matsuoka, K., Miyoshi, N., Hoshi, K., Ueda, M., Yabu, I., & Hirata, M. (2006). The acquisition of Japanese focus particles: Dake 'only' and mo 'also'. In C. Zaller, D. Bamman, & T. Magnitskaia (Eds.), *Proceedings of the 30th boston university conference on language development*, Cascadilla Press.

Mercer, Robert. (1992). Default logic: Towards a common logical semantics for presuppositions and entailments. *Journal of Semantics*, *9*(3), 223–250.

Merchant, Jason. (1998). 'Pseudosluicing': elliptical clefts in Japanese and English. *ZAS working papers in linguistics*, *10*, 88–112.

Moore, Chris, Bryant, Dana, & Furrow, David. (1989). Mental Terms and the Development of Certainty. *Child Development*, *60*(1), 167–171.

Moore, Chris, & Davidge, Jane. (1989). The development of mental terms: Pragmatics or semantics? *Journal of Child Language*, *16*(3), 633–641.

Noveck, Ira A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Nwoye, Onuigbo G. (1992). Linguistic politeness and socio-cultural variations of the notion of face. *Journal of Pragmatics*, *18*(4), 309–328.

Obeng, Samuel Gyasi. (1994). Verbal indirection in Akan informal discourse. *Journal of Pragmatics*, *21*(1), 37–65.

Obeng, Samuel Gyasi. (1999). "In future if I buy a dog, I'll call it 'Okyeman-is-ungrateful'." Indirect response to potentially difficult communicative situations: The case of Akan dog names. *International Journal of the Sociology of Language*, *140*(1), 83–104.

Obeng, Samuel Gyasi. (2003). *Language in African social interaction: indirectness in Akan communication*. Hauppauge, New York: Nova Science Pub Inc.

Ohashi, Jun. (2008). Linguistic rituals for thanking in Japanese: Balancing obligations. *Journal of pragmatics*, *40*(12), 2150–2174.

Paul, Waltraud, & Whitman, John. (2008). Shi ... De focus clefts in Mandarin Chinese. *Linguistic Review*, *25*(3-4), 413–451.

Perlmann-Balme, Michaela. (2004). *Start Deutsch • Deutschprüfungen für Erwachsene • Prüfungsziele, Testbeschreibung*. Goethe-Institut.

Perlmann-Balme, Michaela. (2015). *Goethe-Zertifikat A2 : Deutschprüfung für Jugendliche und Erwachsene Prüfungsziele Testbeschreibung*. (1st ed.). Hueber Verlag.

Potts, Christopher. (2005). *The Logic of Conventional Implicatures*.

Prince, Ellen. (1978). A Comparison of Wh-Clefts and it-Clefts in Discourse. *Language*, *54*(4), 883–906.

R Core Team. (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ran, Yongping, & Zhao, Linsen. (2018). Building mutual affection-based face in conflict mediation : A Chinese relationship management model. *Journal of Pragmatics*, *129*, 185–198.

Reins, Louisa M., Wiegmann, Alex, Marchenko, Olga P., & Schumski, Irina. (2021). Lying Without Saying Something False? A Cross-Cultural Investigation of the Folk Concept of Lying in Russian and English Speakers. *Review of Philosophy and Psychology*.

Roberts, Craige. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Ohio State University Working Papers in Linguistics*, *49*, 91–136.

Roberts, Craige, Simons, Mandy, Beaver, David, & Tonhauser, Judith. (2009). Presupposition, conventional implicature, and beyond: A unified account of projection. *Proceedings of the ESSLLI 2009 Workshop New Directions in the Theory of Presupposition*, 1–15.

Romoli, Jacopo. (2012). *Soft but strong. Neg-raising, soft triggers, and exhaustification* (Doctoral dissertation, Harvard University).

Romoli, Jacopo. (2014). The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*, *32*(2), 173–219.

Romoli, Jacopo, & Schwarz, Florian. (2015). An Experimental Comparison Between Presuppositions and Indirect Scalar Implicatures. In *Experimental perspectives on presuppositions* (pp. 215–240).

Rossi, Giovanni, Floyd, Simeon, & Enfield, N J. (2020). Recruitments and pragmatic typology. In S. Floyd, G. Rossi, & N. J. Enfield (Eds.), *Getting others to do things: A pragmatic typology of recruitments* (Chap. 1, pp. 1–16).

Rue, Yong-Ju, & Zhang, Grace Qiao. (2008). *Request Strategies: A Comparative Study in Mandarin Chinese and Korean*. Amsterdam: John Benjamins.

Russell, Bertrand. (1905). On Denoting. *Mind*, *14*(4), 479–493.

Sapir, Edward. (1921). *LANGUAGE: AN INTRODUCTION TO THE STUDY OF SPEECH*. New York: Harcourt, Brace and company.

Sauerland, Uli. (2008). Implicated presuppositions. *The discourse potential of underspecified structures*, 581–600.

Schlenker, Philippe. (2010). Local contexts and local meanings. *Philosophical Studies*, *151*(1), 115–142.

Schlenker, Philippe. (2019). Gestural semantics: Replicating the typology of linguistic inferences with pro- and post-speech gestures. *Natural Language and Linguistic Theory*, *37*(2), 735–784.

Schlenker, Philippe. (2021a). Iconic presuppositions. *Natural Language and Linguistic Theory*, *39*(1), 215–289.

Schlenker, Philippe. (2021b). Triggering presuppositions. *Glossa*, *6*(1), 1–28.

Schwarz, Florian, Djärv, Kajsa, & Zehr, Jérémy. (2020). Do Italian factives entail their presupposition? Yes, but... In R. Bhatt, I. Frana, & P. Menéndez-Benito (Eds.), *Making worlds accessible. essays in honor of angelika kratzer* (pp. 150–166). Amherst, Massachusetts: ScholarWorks@UMass.

Sellars, Wilfrid. (1954). Presupposing. *The Philosophical Review*, *63*(2), 197–215.

Shanon, Benny. (1976). On the two kinds of presuppositions in natural language. *Foundations of language*, *14*(2), 247–249.

Simons, Mandy. (2001). On the Conversational Basis of Some Presuppositions. *Semantics and Linguistic Theory*, *11*, 431–448.

Simons, Mandy. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, *117*(6), 1034–1056.

Simons, Mandy, Tonhauser, Judith, Beaver, David, & Roberts, Craige. (2010). What projects and why. *Semantics and linguistic theory*, *20*, 309–327.

Singmann, Henrik, Bolker, Ben, Westfall, Jake, Aust, Frederik, & Ben-Shachar, Mattan S. (2020). afex: Analysis of Factorial Experiments.

Smith, E. Allyn, & Hall, Kathleen Currie. (2011). Projection diversity: Experimental evidence. *Proceedings of the 2011 ESSLLI Workshop on Projective Content*, 156–170.

Snider, Todd. (2017a). *Anaphoric Reference to Propositions* (Doctoral dissertation, Cornell University).

Snider, Todd. (2017b). At-issueness anaphoric availability. *Proceedings of the Linguistic Society of America*, *2*, 39.

Spenader, Jennifer. (2002). *Presuppositions in spoken discourse* (Doctoral dissertation, University of Stockholm).

Sperber, Dan, & Wilson, Deirdre. (1996). *Relevance: Communication and Cognition* (2nd ed.). Oxford: Blackwell.

Stalnaker, Robert. (1972). Pragmatics. In D. Davidson & G. Harman (Eds.), *Semantics of natural language.* (pp. 380–397). Dordrecht: D. Reidel Publishing Company.

Stalnaker, Robert. (1973). Presuppositions. *Journal of philosophical logic*, *2*(4), 447–457.

Stalnaker, Robert. (1977). Pragmatic Presuppositions. *Proceedings of the Texas Conference on Performatives, Presuppositions and Implicatures.*, 135–148.

Stalnaker, Robert. (1998). On the Representation of Context. *Journal of Logic, Language, and Information*, *7*(1), 3–19.

Stalnaker, Robert. (2002). Common ground. *Linguistics and philosophy*, *25*(5/6), 701–721.

Stateva, Penka, Stepanov, Arthur, Déprez, Viviane, Dupuy, Ludivine E., & Reboul, Anne C. (2019). Cross-linguistic variation in the meaning of quantifiers: Implications for pragmatic enrichment. *Frontiers in Psychology*, *10*(APR), 1–18.

Stevens, Jon Scott, de Marneffe, Marie-Catherine, Speer, Shari R, & Tonhauser, Judith. (2017). Rational use of prosody predicts projection in manner adverb utterances. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, (5), 1144–1149.

Strawson, Peter Frederick. (1950). On Referring. *Mind*, *59*(235), 320–344.

Strawson, Peter Frederick. (1952). *Introduction to logical theory*. London: Methuen.

Strawson, Peter Frederick. (1964). Identifying reference and truth-values. *Theoria*, *30*(2), 96–118.

Sudo, Yasutada. (2012). *On the semantics of phi features on pronouns* (Doctoral dissertation, Massachusetts Institute of Technology).

Sun, Hao. (2004). Opening moves in informal Chinese telephone conversations. *Journal of Pragmatics*, *36*(8), 1429–1465.

Syrett, Kristen, Kennedy, Christopher, & Lidz, Jeffrey. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics*, *27*(1), 1–35.

Taghizadeh, REZA. (2017). *Pragmatic competence in the target language: A study of Iranian Learners of English* (Doctoral dissertation, University of Salford).

Taguchi, Naoko. (2011). Teaching pragmatics: Trends and issues. *Annual Review of Applied Linguistics*, *31*, 289–310.

Taguchi, Naoko. (2013). Comprehension of conversational implicature. In N. Taguchi & J. M. Sykes (Eds.), *Technology in interlanguage pragmatics research and teaching* (Chap. 2, pp. 19–41).

Tang, Chen Hsin, & Zhang, Grace Qiao. (2009). A contrastive study of compliment responses among Australian English and Mandarin Chinese speakers. *Journal of Pragmatics*, *41*(2), 325–345.

Teng, Shou-hsin(邓守信). (1979). Remarks on Cleft Sentences in Chinese. *Journal of Chinese Linguistics*, *7*(1), 101–114.

Thalmann, Maik, Chen, Yuqiu, Müller, Susanne, Paluch, Markus, & Antomo, Mailin. (2021). Against PCI-GCI uniformity. Evidence from deceptive language in German and Chinese. *Linguistische Berichte*, *267*, 235–285.

The Commercial Press. (2012). *Xiandai hanyu cidian* - 现代汉语词典 (6th ed.) (Zhongguo she hui ke xue yuan Yu yan yan jiu suo Ci dian bian ji shi 中国社会科学院语言研究所词典编辑室., Ed.). Beijing: The Commercial Press 商务印书馆.

Theiler, Nadine. (2021). Denn as a highlighting-sensitive particle. *Linguistics and Philosophy*, *44*(2), 323–362.

Tieu, Lyn, & Kriz, Manuel. (2017). Connecting the exhaustivity of clefts and the homogeneity of plural definite descriptions in acquisition. *Proceedings of the 41st annual Boston University Conference on Language Development*, 651–664.

Tieu, Lyn, Schlenker, Philippe, & Chemla, Emmanuel. (2019). Linguistic inferences without words. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(20), 9796–9801.

Tonhauser, Judith. (2012). Diagnosing (not-)at-issue content. *Proceedings of Semantics of Under-represented Languages of the Americas (SULA)*, *6*, 239–254.

Tonhauser, Judith. (2015). Are 'informative presuppositions' presuppositions? *Language and Linguistics Compass*, *9*(2), 77–101.

Tonhauser, Judith. (2016). Prosodic cues to presupposition projection. *Semantics and Linguistic Theory*, *26*, 934.

Tonhauser, Judith, Beaver, David, & Degen, Judith. (2018). How Projective is Projective Content? Gradience in Projectivity and At-issueness. *Journal of Semantics*, 495–542.

Tonhauser, Judith, Beaver, David, Roberts, Craige, & Simons, Mandy. (2013). Toward a Taxonomy of Projective Content. *Language*, *89*(1), 66–109.

Van der Sandt, Rob. (1988). *Context and Presupposition*. London: Routledge Kegan & Paul.

Van Fraassen, B.C. (1966). Singular terms, truth-value gaps, and free logic. *The journal of Philosophy*, *63*(17), 481–495.

Venhuizen, Noortje Joost, Bos, Johan, & Brouwer, Harm. (2013). Parsimonious semantic representations with projection pointers. In K. Erk & A. Koller (Eds.), *Proceedings of the 10th international conference on computational semantics, iwcs 2013 - long papers* (pp. 252–263). Potsdam.

Venhuizen, Noortje Joost, Bos, Johan, Hendriks, Petra, & Brouwer, Harm. (2015). How and why conventional implicatures project. *Semantics and Linguistic Theory*, *24*, 63.

von Fintel, Kai. (2004). Would You Believe It? The King of France is Back! Presuppositions and truth- value intuitions. In M. Reimer & A. Bezuidenhout (Eds.), *Descriptions and beyond* (pp. 315–341). Oxford: Oxford University Press.

von Fintel, Kai. (2008). What is presupposition accommodation, again? *Philosophical Perspectives*, *22*(1), 137–170.

von Stutterheim, Christiane, & Klein, Wolfgang. (1989). Referential Movement in Descriptive and Narrative Discourse. In R. Dietrich & C. F. Graumann (Eds.), *Language processing in social context* (pp. 39–76).

Whorf, Benjamin Lee. (1940). Science and linguistics. *Technology Review*, *42*(6), 229–231, 247–248.

Wierzbicka, Anna. (1985). Different cultures, different languages, different speech acts: Polish vs. English. *Journal of pragmatics*, *9*, 145–178.

Wilson, Deirdre. (1975). *Presuppositions and non-truth-conditional semantics*. New York: Academic Press.

Wilson, Deirdre, & Sperber, Dan. (1979). Ordered entailments: An alternative to presuppositional theories. In C.-K. Oh & D. A. Dinneeen (Eds.), *Presupposition* (pp. 299–323). Academic Press.

Wilson, Deirdre, & Sperber, Dan. (2004). Relevance Theory. In L. Horn & C. Ward (Eds.), *Handbook of pragmatics* (Chap. 27, pp. 607–632). Oxford: Blackwell Publishing.

Xia, Dengshan, & Lan, Chun. (2019). (Im)politeness at a Chinese dinner table: A discursive approach to (im)politeness in multi-party communication. *Journal of Politeness Research*, *15*(2), 223–256.

Xu, Guangcan (许光灿). (2014). 现代汉语中几类预设触发语研究 (Doctoral dissertation, 南开大学).

Xue, Jingyang, & Onea, Edgar. (2011). Correlation between presupposition projection and at-issueness: An empirical study. *Proceedings of the ESSLLI 2011 workshop on projective meaning*, (January 2011), 171–184.

Yatsushiro, Kazuko. (2008a). German determiner presuppositions in first language acquisition. *Proceedings supplement of BUCLD*, *32*.

Yatsushiro, Kazuko. (2008b). Quantifier acquisition: Presuppositions of "every". *Proceedings of Sinn und Bedeutung*, *12*, 663–677.

Zeevat, Henk. (1992). Presupposition and accommodation in update semantics. *Journal of Semantics*, *9*(4), 379–412.

Zehr, Jérémy, Bill, Cory, Tieu, Lyn, Romoli, Jacopo, & Schwarz, Florian. (2016). Presupposition projection from the scope of None: Universal, existential, or both? *Semantics and Linguistic Theory*, *26*, 754–774.

Zehr, Jérémy, Hirsch, Aron, Bacovcin, Hezekiah A, & Schwarz, Florian. (2017). Priming local accommodation of hard triggers in disjunction. In *Proceedings of nels* (Vol. 47), Amherst, MA: GLSA Publications.

Zehr, Jérémy, & Schwarz, Florian. (2016). Entailed vs. Non-Entailed Presuppositions - An Experimental Assessment. In *Proceedings of nels* (Vol. 46, pp. 319–328).

Zehr, Jérémy, & Schwarz, Florian. (2018). Returning to non-entailed presuppositions again. *ZAS Papers in Linguistics*, *61*, 463–480.

Zhan, Fangqiong, & Sun, Chaofen. (2013). A copula analysis of shì in the chinese cleft construction. *Language and Linguistics*, *14*(4), 755–789.

Zhang, Niina. (2001). On the absence of nonrestrictive relatives in Chinese. *Ms., ZAS, Berlin*.

Zhou, Mengru(周梦茹). (2020). A Pragmatic Study of the Exhaustivity of Chinese "shi" Clefts. *Modern Linguistics*, *8*(1), 6–14.

Zhu, Hua, Li, Wei, & Yuan, Qian. (2000). The sequential organisation of gift offering and acceptance in Chinese. *Journal of Pragmatics*, *32*(1), 81–103.

Al-Zubeiry, Hameed Yahya A. (2020). Presupposition triggers in british and saudi english newspaper opinions. *Theory and Practice in Language Studies*, *10*(7), 734–743.

Appendices

A.1 Experiment I[56]

A.1.1 Warm-Up Items

Table 17: Warm-up items.

	German	English
Infelicitous	Der Frosch spielt im Sandkasten. Danach ist er sehr schmutzig. Er geht ins Zimmer und zieht sich um.	The frog is playing in the sandbox. Afterwards, he is very dirty. He goes to his room and changes clothes.
	Der Clown hat nicht gut aufgepasst und fragt: "Wer ist schmutzig?"	The clown did not pay attention again and asks: "Who is dirty?"
	Der kleine Peter antwortet: "Der schmutzige Frosch zieht sich um."	Little Peter responds: "The dirty frog is changing clothes."
Felicitous	Die Ente hat Ball gespielt und hat jetzt großen Hunger. Sie isst eine große Pizza.	Die duck played ball and is very hungry now. She is eating a large pizza.
	Der Clown hat nicht gut aufgepasst und fragt: "Was isst denn die Ente?"	The clown did not pay attention again and asks: "What is the duck eating?"
	Der kleine Peter antwortet: "Die hungrige Ente isst eine große Pizza."	Little Peter responds: "The hungry duck is eating a large pizza."
Infelicitous	Der Panda ist sehr lange geschwommen und ist nun sehr müde. Er setzt sich auf eine Bank und trinkt Milch.	Der Panda swam for a very long time and is now very tired. He sits down on a bench and drinks milk.
	Der Clown hat nicht gut aufgepasst und fragt: "Wer ist müde?"	The clown did not pay attention again and asks: "Who is tired?"
	Der kleine Peter antwortet: "Der müde Panda trinkt Milch."	Little Peter responds: "The tired panda is drinking milk."

[56]This section is published together with Chapter 7, namely as appendix in the journal article Y. Chen et al. (2022), see https://doi.org/10.1016/j.pragma.2022.06.014.

A.1.2 Critical Items

For all triggers below, the two Questions Under Discussion uttered by the clown represent the NRRC/Presupposition at-issue and Assertion at-issue conditions, respectively.

Table 18: Critical Items with NRRCs.

	German	English
1	Der Panda und die Ente spielen Fußball. Plötzlich rutscht der Panda aus und fällt hin. Sein Knie blutet etwas. Der Panda bekommt ein Pflaster und geht dann in die Küche, um etwas zu trinken. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wie hat sich denn der Panda verletzt?" / "Wo ist denn der Panda?" Der kleine Peter antwortet: "Der Panda, der übrigens ausgerutscht ist, ist in der Küche."	Der Panda and the duck are playing soccer. Suddenly, the panda slips and falls. His knee is bleeding a little. The panda gets a band-aid and then goes to the kitchen to drink something. The clown did not pay attention again and asks: "How did the panda injure himself?" / "Where is the panda?" Little Peter responds: "The panda, who slipped by the way, is in the kitchen."
2	Der Frosch und die Ente spielen zusammen im Sandkasten. Der Frosch baut eine Burg und die Ente einen Tunnel. Die Ente setzt sich kurz unter einen Baum, um sich auszuruhen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wo ist denn die Ente?" / "Was hat denn die Ente gebaut?" Der kleine Peter antwortet: "Die Ente, die sich übrigens gerade unter einem Baum ausruht, hat einen Tunnel gebaut."	The frog and the duck are playing together in the sandbox. The frog is building a sandcastle and the duck a tunnel. The duck sits down underneath a tree to rest. The clown did not pay attention again and asks: "Where is the duck?" / "What did the duck build?" Little Peter responds: "The duck, who is resting underneath a tree by the way, built a tunnel."

Table 18 – *Continued from previous page*

	German	English
3	Der Frosch sammelt im Garten schöne Steine. Er hat schon einen kleinen Haufen an Steinen gesammelt. Dann geht er ins Kinderzimmer, um zu malen. Der Clown hat wieder nicht gut aufgepasst und fragt: “Was hat denn der Frosch im Garten gesammelt?” / “Wo ist denn der Frosch?” Der kleine Peter antwortet: “Der Frosch, der übrigens im Garten schöne Steine gesammelt hat, malt im Kinderzimmer.”	The frog is collecting pretty rocks in the garden. He already collected a small pile of rocks. Then he goes to the children’s room to draw. The clown did not pay attention again and asks: “What did the frog collect in the garden?” / “Where is the frog?” Little Peter responds: “The frog, who collected pretty rocks in the garden by the way, is drawing in the children’s room.”
4	Die Ente hat Husten. Sie bekommt Hustensaft. Dann geht sie ins Bett, um sich auszuruhen. Der Clown hat wieder nicht gut aufgepasst und fragt: “Was hat die Ente denn bekommen?” / “Wo ist denn jetzt die Ente?” Der kleine Peter antwortet: “Die Ente, die übrigens Hustensaft bekommen hat, ist nun im Bett.”	The duck has a cough. She gets some cough syrup. Then she goes to bed to rest. The clown did not pay attention again and asks: “What did the duck get?” / “Where is the duck now?” Little Peter responds: “The duck, who got some cough syrup by the way, is in bed now.”
5	Die Tiere frühstücken zusammen. Der Frosch mag keine Rosinen und lässt deshalb sein Müsli stehen. Er geht nach draußen und spielt im Garten. Der Clown hat wieder nicht gut aufgepasst und fragt: “Warum isst denn der Frosch das Müsli nicht?” / “Wo ist denn jetzt der Frosch?”Der kleine Peter antwortet: “Der Frosch, der übrigens keine Rosinen mag, spielt im Garten.”	The animals are eating breakfast together. The frog does not like raisins and thus leaves his muesli untouched. He goes outside and plays in the garden. The clown did not pay attention again and asks: “Why does the frog not eat his muesli” / “Where is the frog now?” Little Peter responds: “The frog, who does not like raisins by the way, is playing in the garden.”

Table 18 – *Continued from previous page*

	German	English
6	Der Panda malt mit Wasserfarben. Er verwischt dabei aus Versehen sein Bild. Er macht dann erstmal eine Pause und trinkt in der Küche Saft. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was ist denn mit dem Bild vom Panda passiert?" / "Wo ist denn jetzt der Panda?" Der kleine Peter antwortet: "Der Panda, der übrigens aus Versehen sein Bild verwischt hat, trinkt nun in der Küche Saft."	The panda is painting with watercolors. He accidentally smears his painting. Then, he takes a break and drinks some juice in the kitchen. The clown did not pay attention again and asks: "What happened to the panda' s painting?" / "Where is the panda now?" Little Peter responds: "The panda, who smeared his painting by the way, is drinking juice in the kitchen now."
7	Der Panda spielt im Garten und hat eine Hütte aus Zweigen gebaut. Er will sie den anderen Tieren zeigen. Leider stößt der Panda aus Versehen mit dem Kopf gegen die Hütte, und die Hütte geht dabei kaputt. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wo spielt denn der Panda?" / "Was hat denn der Panda kaputt gemacht?" Der kleine Peter antwortet: "Der Panda, der übrigens im Garten spielt, hat aus Versehen die Hütte kaputt gemacht."	The panda is playing in the garden and built a hut with branches. He wants to show it to the other animals. Unfortunately, the panda hits the hut with his head and the hut breaks. The clown did not pay attention again and asks: "Where is the panda playing?" / "What did the panda break?" Little Peter responds: "The panda, who is playing in the garden by the way, accidentally broke the hut."

Table 18 – *Continued from previous page*

	German	English
8	Die Tiere spielen Lego und trinken dabei Saft. Der Panda hat einen hohen Turm gebaut. Plötzlich stößt der Panda aus Versehen sein Glas um und der ganze Saft läuft auf die Legosteine. Der Clown hat wieder nicht gut aufgepasst und fragt: " Was hat denn der Panda gebaut?" / "Warum ist denn das Lego so nass?" Der kleine Peter antwortet: "Der Panda, der übrigens einen Turm gebaut hat, hat sein Glas umgekippt."	The animals are playing with Legos and drinking juice. The panda built a high tower. Suddenly, the panda knocks over his glass and the juice spills over the Lego bricks. The clown did not pay attention again and asks: "What did the panda build?" / "Why is the Lego wet?" Little Peter responds: "The panda, who built a tower by the way, knocked over his glass."
9	Die Ente klettert im Garten auf einen Baum. Plötzlich fällt sie von einem Ast. Ihr Fuß tut weh und die Ente weint. Der Panda kommt vorbei und gibt ihr ein Gummibärchen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Worauf ist die Ente denn geklettert?" / "Was hat die Ente denn bekommen?" Der kleine Peter antwortet: "Die Ente, die übrigens auf einen Baum geklettert ist, hat ein Gummibärchen bekommen."	In the garden, the duck is climbing a tree. Suddenly she falls off a branch. Her foot hurts and the she is crying. The panda comes over and hands her a gummy bear. The clown did not pay attention again and asks: "What did the duck climb?" / "What did the duck get?" Little Peter responds: "The duck, who climbed a tree by the way, got a gummy bear."

10	Der Frosch pflückt einen großen Blumenstrauß im Garten. Dann legt er den Strauß auf die Wiese. Er geht rein, um eine Blumenvase zu holen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was hat denn der Frosch gepflückt?" / "Wohin ist denn der Frosch gegangen?" Der kleine Peter antwortet: "Der Frosch, der übrigens die Blumen gepflückt hat, ist reingegangen."	The frog is gathering a large bouquet of flowers in the garden. Then, he puts the bouquet on the floor. He goes inside to get a vase. The clown did not pay attention again and asks: "What did the frog gather?" / "Where did the frog go?" Little Peter responds: "The frog, who gathered flowers by the way, went inside."

Table 19: Critical items with hard triggers.

	German	English
	***it*-clefts**	
1	Die Tiere wollen heute das neue Schwimmbad in der Stadt ausprobieren. Leider wissen sie aber nicht, wo es ist. Der Panda kennt das neue Schwimmbad schon. Er will die Tiere heute nach dem Mittagessen abholen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wann will denn der Panda die Tiere abholen?" / "Wer will denn die Tiere nach dem Mittagessen abholen?" Der kleine Peter antwortet: "Es ist der Panda, der die Tiere nach dem Mittagessen abholen will."	The animals want to try out the new swimming pool in the city today. Unfornately, they do not know where it is. The panda knows the swimming pool already. He wants to pick up the animals after lunch today. The clown did not pay attention again and asks: "When does the panda want to pick up the animals?" / "Who wants to pick up the animals after lunch?" Little Peter responds: "It is the panda, who wants to pick up the animals after lunch."

Table 19 – *Continued from previous page*

	German	English
2	Die Tiere malen im Kindergarten. Der Frosch malt mit Glitzerstiften. Alle Tiere finden sein Bild am schönsten. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wie hat denn der Frosch das schöne Bild gemalt?" / "Wer hat denn das schöne Bild gemalt?" Der kleine Peter antwortet: "Es war der Frosch, der das Bild mit Glitzerstiften gemalt hat."	The animals are drawing in kindergarten. The frog draws with glittery pens. All of the animals like his drawing the best. The clown did not pay attention again and asks: "How did the frog draw the pretty drawing?" / "Who drew the pretty painting?" Little Peter responds: "It was the frog who drew the painting with the glittery pens."
3	Im Kindergarten ist Sommerfest und alle Tiere bringen etwas mit. Die Ente hat Saft und Wasser mitgebracht. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was hat denn die Ente mitgebracht?" / "Wer hat denn Saft und Wasser mitgebracht?" Der kleine Peter antwortet: "Es ist die Ente, die Saft und Wasser mitgebracht hat."	There' s a summer festival at the kindergarten and all animals bring something. The duck brought juice and water. The clown did not pay attention again and asks: "What did the duck bring?" / "Who brought the juice and the water?" Little Peter responds: "It is the duck who brought the juice and the water."
4	Die Ente und der Panda spielen Lego. Danach räumt die Ente das Lego zurück in die Spielzeugkiste. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wohin hat denn die Ente das Lego geräumt?" / "Wer hat denn das Lego in die Spielzeugkiste geräumt?" Der kleine Peter antwortet: "Es war die Ente, die das Lego in die Spielzeugkiste geräumt hat."	The duck and the panda are playing with Legos. Afterwards, the duck puts the Lego back into the box of toys. The clown did not pay attention again and asks: "Where did the duck put the Legos?" / "Who put the Legos into the box of toys?" Little Peter responds: "It was the duck who put the Legos back into the box of toys."

again

Table 19 – *Continued from previous page*

	German	English
1	Seit einiger Zeit geht der Panda jeden Montag zum Schwimmunterricht. Gestern war Montag und er war wie immer schwimmen. Der Clown hat wieder nicht gut aufgepasst und fragt: "War der Panda gestern zum ersten Mal beim Schwimmunterricht?" / "War der Panda gestern beim Schwimmunterricht?" Der kleine Peter antwortet: "Der Panda war gestern wieder beim Schwimmunterricht."	For some time, the panda has been taking swimming lessons every Monday. It was Monday yesterday and he went swimming as usual. The clown did not pay attention again and asks: "Did the panda take swimming lessons for the first time yesterday?" / "Did the panda take swimming lessons yesterday?" Little Peter responds: "The panda took swimming lessons again yesterday."
2	Heute im Kindergarten ist der Frosch beim Spielen eingeschlafen, weil er sehr müde war. Das ist schon öfter passiert. Der Clown hat wieder nicht gut aufgepasst und fragt: "Ist der Frosch heute zum ersten Mal im Kindergarten eingeschlafen?" / "Ist der Frosch heute im Kindergarten eingeschlafen?" Der kleine Peter antwortet: "Der Frosch ist heute wieder im Kindergarten eingeschlafen."	Today in kindergarten, the frog fell asleep while playing because he was very tired. That has happened before. The clown did not pay attention again and asks: "Did the frog fall asleep in kindergarten for the first time today?" / "Did the frog fall asleep in kindergarten today?" Little Peter responds: "The frog fell asleep again in kindergarten today."

Table 19 – *Continued from previous page*

	German	English
3	Der Panda ist im Fußballverein. Er spielt seit einiger Zeit besonders gut und hat in den letzten Wochen schon viele Tore geschossen. Gestern war ein Fußballspiel und der Panda hat ein Tor geschossen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat der Panda gestern zum ersten Mal ein Tor geschossen?" / "Hat der Panda gestern ein Tor geschossen?" Der kleine Peter antwortet: "Der Panda hat gestern wieder ein Tor geschossen."	The panda is in the soccer club. He has been playing very well for some time and scored a lot of goals in past weeks. Yesterday, there was a game and the panda scored a goal. The clown did not pay attention again and asks: "Did the panda score a goal for the first time yesterday?" / "Did the panda score a goal yesterday?" Little Peter responds: "The panda scored a goal again yesterday."
	too	
1	Die Tiere machen ein Picknick. Die Ente hat die Lieblingskekse vom Frosch dabei. Beim Picknick isst der Frosch viele Kekse und Würstchen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat der Frosch beim Picknick nur Kekse gegessen?" / "Hat der Frosch beim Picknick Kekse gegessen?" Der kleine Peter antwortet: "Der Frosch hat auch Kekse gegessen."	The animals are having a picknick. The duck brought the frog's favorite cookies. During the picknick, the frog eats many cookies and sausages. The clown did not pay attention again and asks: "Did the frog only eat cookies at the picknick?" / "Did the frog eat cookies at the picknick?" Little Peter responds: "The frog ate cookies too."
2	Der Frosch hat heute Geburtstag. Deshalb bringt er Kuchen und Schokolade mit in den Kindergarten. Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat der Frosch heute nur Kuchen mitgebracht?" / "Hat der Frosch heute Kuchen mitgebracht?" Der kleine Peter antwortet: "Der Frosch hat heute auch Kuchen mitgebracht."	It's the frog's birthday today. Because of this, he brings cake and chocolate to kindergarten. The clown did not pay attention again and asks: "Did the frog only bring cake?" / "Did the frog bring cake today?" Little Peter responds: "The frog brought cookies too today."

Table 19 – *Continued from previous page*

	German	English
3	Die Ente hat diese Woche im Kindergarten Gartendienst. Sie muss jeden Tag die Blumen und die Tomatenpflanzen gießen. Heute hat sie die Blumen und die Tomatenpflanzen gegossen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat die Ente heute nur die Blumen gegossen?" / "Hat die Ente heute eigentlich die Blumen gegossen?" Der kleine Peter antwortet: "Die Ente hat heute auch die Blumen gegossen."	The duck has to help in the garden at kindergarten this week. She has to water the flowers and the tomatoes every day. Today, she watered the flowers and the tomatoes. The clown did not pay attention again and asks: "Did the duck only water the flowers today?" / "Did the duck water the plants today?" Little Peter responds: "The duck watered the flowers too today."

Table 20: Critical items with soft triggers.

	German	English
	find out	
1	Der Frosch und der Panda wollen Flummi spielen, aber sie wissen nicht, wo der Flummi ist. Sie suchen im Kinderzimmer. Da sieht der Panda den Flummi unter dem Tisch. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wo ist denn der Flummi?" / "Wer hat denn den Flummi gefunden?" Der kleine Peter antwortet: "Der Panda hat entdeckt, dass der Flummi unter dem Tisch ist."	The frog and the panda want to play with a bouncy ball but they do not know where it is. They are looking in the kid's room. There, the panda sees the bouncy ball under the table. The clown did not pay attention again and asks: "Where is the bouncy ball?" / "Who found the bouncy ball?" Little Peter responds: "The panda found out that the bouncy ball is under the table."

Table 20 – *Continued from previous page*

	German	English
2	Die Tiere spielen zusammen Ball. Die Ente wirft den Ball mit so viel Kraft, dass der Ball weit weg fliegt. Die Tiere sehen den Ball nicht mehr. Zusammen suchen sie den Ball. Die Ente sieht den Ball unter einem Busch und holt ihn. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wo war denn der Ball?" / "Wer hat denn den Ball gefunden?" Der kleine Peter antwortet: "Die Ente hat entdeckt, dass der Ball unter einem Busch war."	The animals are playing ball together. The duck throws the ball with so much force that the ball flies far away. The animals do not see the ball anymore. Together, they are looking for the ball. The duck spots the ball under a bush and gets it. The clown did not pay attention again and asks: "Where was the ball?" / "Who found the ball?" Little Peter responds: "The duck found out that the ball was under a bush."
3	Die Tiere spielen zusammen Verstecken. Der Panda versteckt sich hinter einem großen Baum. Der Frosch sieht das schwarz-weiße Fell vom Panda und fängt ihn. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wo hatte sich denn der Panda versteckt?" / "Wer hat denn den Panda gefunden?" Der kleine Peter antwortet: "Der Frosch hat entdeckt, dass sich der Panda hinter einem Baum versteckt hatte."	The animals are playing hide-and-seek. The panda hides behind a large tree. The frog sees the panda' s black and white fur and catches him. The clown did not pay attention again and asks: " Where did the panda hide?" / "Who found the panda?" Little Peter responds: "The frog found out that the panda was hiding behind a tree."

manage

Table 20 – *Continued from previous page*

	German	English
1	Der Panda hat auf seiner Geburtstagsfeier ein sehr schwieriges Puzzle bekommen. Die anderen Tiere schauen zu, wie der Panda das Puzzle macht und klatschen, als er fertig ist. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wer hat denn versucht, das schwierige Puzzle zu machen?" / "Wer hat denn das schwierige Puzzle gemacht?" Der kleine Peter antwortet: "Der Panda hat es geschafft, das Puzzle zu machen."	The panda got a very hard puzzle at this birthday party. The other animals are watching the panda puzzle and clap when he finishes. The clown did not pay attention again and asks: "Who tried to finish the hard puzzle?" / "Who finished the hard puzzle?" Little Peter responds: "The panda managed to finish the puzzle."
2	Die Tiere schneiden Sterne aus, um sie ans Fenster zu kleben. Der rote Stern ist besonders klein und deswegen sehr schwierig auszuschneiden. Die Ente strengt sich sehr an und am Ende hat sie den roten Stern ausgeschnitten. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wer hat denn versucht, den roten Stern auszuschneiden?" / "Wer hat denn den roten Stern ausgeschnitten?" Der kleine Peter antwortet: "Die Ente hat es geschafft, den roten Stern auszuschneiden."	The animals are cutting out stars to glue them to the window. The red star is especially small and thus difficult to cut out. The duck tries very hard and, at the end, she has cut out the red star. The clown did not pay attention again and asks: "Who tried to cut out the red star?" / "Who cut out the red star?" Little Peter responds: "The duck managed to cut out the red star."

Table 20 – *Continued from previous page*

	German	English
3	Der Wind heute ist nicht sehr stark aber die Tiere wollen trotzdem einen Drachen steigen lassen. Mit der Schnur in der Hand rennt der Frosch los, und der Drachen fliegt hoch in die Luft. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wer hat denn versucht, den Drachen steigen zu lassen?" / "Wer hat denn den Drachen steigen lassen?" Der kleine Peter antwortet: "Der Frosch hat es geschafft, den Drachen steigen zu lassen."	The wind is not very strong today but the animals still want to fly a kite. With the rope in hand, the frog starts running and the kite sores high into the sky. The clown did not pay attention again and asks: "Who tried to let the kite fly?" / "Who let the kite fly?" Little Peter responds: "The frog managed to let the kite fly."
4	Die Tiere haben in einen großen Eimer ganz viel Wasser gefüllt. Jetzt ist der Eimer richtig schwer. Die Tiere wollen den Eimer jetzt zum Planschbecken tragen. Mit ganz viel Kraft kann der Panda den Eimer hochheben. Er bringt ihn zum Planschbecken. Der Clown hat wieder nicht gut aufgepasst und fragt: "Wer hat denn versucht, den schweren Eimer zu tragen?" / "Wer hat denn den schweren Eimer getragen?" Der kleine Peter antwortet: "Der Panda hat es geschafft, den schweren Eimer zu tragen."	The animals filled a large bucket with a lot of water. Now, the bucket is very heavy. The animals want to carry the bucket to the pool. With a lot of effort, the panda can lift the bucket. He carries it to the pool. The clown did not pay attention again and asks: "Who tried to lift the heavy bucket?" / "Who lifted the heavy bucket?" Little Peter responds: "The panda managed to lift the heavy bucket."

win

Table 20 – *Continued from previous page*

	German	English
1	Im Kindergarten findet ein Sportfest statt. Beim Wettrennen machen der Panda, der Frosch und die Ente mit. Die Ente ist als erste im Ziel.	There' s a festival in kindergarten. The panda, the frog, and the duck join in the race. The duck is the first past the goal post.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat die Ente am Wettrennen teilgenommen?" / "Ist die Ente am schnellsten gewesen?"	The clown did not pay attention again and asks: "Did the duck participate in the race?" / "Was the duck the fastest?"
	Der kleine Peter antwortet: "Die Ente hat das Wettrennen gewonnen."	Little Peter responds: "The duck won the race."
2	Die Tiere spielen Dosenwerfen. Wer alle Dosen mit seinem Ball umwirft, gewinnt. Der Panda und die Ente werfen daneben. Der Frosch wirft alle Dosen um und wird Sieger. Er bekommt eine Krone.	The animals are playing can knockdown. Whoever knocks down all the cans with their ball, wins. The panda and the duck miss. The frog knocks over all cans and is the victor. He gets a crown.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat der Frosch am Dosenwerfen teilgenommen?" / "Ist der Frosch beim Dosenwerfen am besten gewesen?"	The clown did not pay attention again and asks: "Did the frog participate at can knockdown?" / "Was the frog the best at can knockdown?"
	Der kleine Peter antwortet: "Der Frosch hat das Dosenwerfen gewonnen."	Little Peter responds: "The frog won at can knockdown."
3	Der Panda, die Ente und der Frosch machen einen Malwettbewerb. Alle finden das Bild der Ente am schönsten. Die Ente bekommt als Preis eine Krone.	The panda, the duck, and the frog have themselves a drawing competition. Everybody likes the duck' s drawing the best. As a prize, the duck gets a crown.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Hat die Ente am Malwettbewerb teilgenommen?" / "Ist die Ente beim Malwettbewerb am besten gewesen?"	The clown did not pay attention again and asks: "Did the duck participate in the drawing competition?" / "Was the duck the best at the drawing competition?"
	Der kleine Peter antwortet: "Die Ente hat den Malwettbewerb gewonnen."	Little Peter responds: "The duck won the drawing competition."

A.1.3 Filler Items

Table 21: Filler items.

	German	English
	Underinformative Answers	
1	Die Tiere essen zusammen zu Mittag. Der Panda und der Frosch essen Würstchen. Die Ente isst Pizza. Der Clown hat wieder nicht gut aufgepasst und fragt: "Welche Tiere essen Würstchen?" Der kleine Peter antwortet: "Der Frosch isst Würstchen."	The animals are having lunch together. The panda and the frog are eating sausages. The duck is eating pizza. The clown did not pay attention again and asks: "Which animals are eating sausages?" Little Peter responds: "The frog is eating sausages."
2	Die Tiere sind im Garten. Der Frosch und der Panda spielen im Sandkasten. Die Ente spielt auf der Wiese Ball. Der Clown hat wieder nicht gut aufgepasst und fragt: "Welche Tiere spielen im Sandkasten?" Der kleine Peter antwortet: "Der Panda spielt im Sandkasten."	The animals are in the garden. The frog and the panda are playing in the sandbox. The duck is playing ball on the meadow. The clown did not pay attention again and asks: "Which animals are playing in the sandbox?" Little Peter responds: "The panda is playing in the sandbox."
3	Die Tiere malen zusammen. Die Ente stößt aus Versehen den Topf mit der roten Farbe um. Die Farbe läuft auf die Ente und den Frosch. Die Ente und der Frosch sind jetzt beide ganz rot. Der Panda ist sauber geblieben und lacht. Der Clown hat wieder nicht gut aufgepasst und fragt: "Welche Tiere sind jetzt rot?" Der kleine Peter antwortet: "Die Ente ist jetzt rot."	The animals are drawing together. The duck accidentally knocks over the pot with the red color. The color spills onto the duck and the frog. The duck and the frog are both very red now. The panda stayed clean and is laughing. The clown did not pay attention again and asks: "Which animals are red now?" Little Peter responds: "The duck is red now."

Table 21 – *Continued from previous page*

	German	English
4	Die Tiere schneiden Sterne aus. Der Frosch und der Panda schneiden rote Sterne aus. Die Ente schneidet die grünen Sterne aus.	The animals are cutting out stars. The frog and the panda are cutting out red stars. The duck is cutting out green stars.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Welche Tiere schneiden rote Sterne aus?"	The clown did not pay attention again and asks: "Which animals are cutting out red stars?"
	Der kleine Peter antwortet: "Der Frosch schneidet rote Sterne aus."	Little Peter responds: "The frog is cutting out red stars."
5	Die Tiere gießen zusammen die Pflanzen im Garten. Der Panda und die Ente gießen die Tomaten. Der Frosch gießt die Blumen.	The animals are watering the plants in the garden together. The panda and the duck are watering the tomatoes. The frog is watering the flowers.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Welche Tiere gießen die Tomaten?"	The clown did not pay attention again and asks: " Which animals are watering the tomatoes?"
	Der kleine Peter antwortet: "Der Panda gießt die Tomaten."	Little Peter responds: "The panda is watering the tomatoes."
	Overinformative Answers	
1	Die Tiere essen Süßigkeiten. Die Ente und der Frosch essen Gummibärchen. Der Panda isst einen Lolli.	Die animals are eating candy. The duck and the frog are eating gummy bears. The panda is eating a lollipop.
	Der Clown hat wieder nicht gut aufgepasst und fragt: "Was isst denn die Ente?"	The clown did not pay attention again and asks: "What is the duck eating?"
	Der kleine Peter antwortet: "Die Ente isst Gummibärchen, und der Panda isst einen Lolli."	Little Peter responds: "The duck is eating gummy bears, and the panda is eating a lollipop."

Table 21 – *Continued from previous page*

	German	English
2	Die Tiere spielen im Sandkasten. Der Frosch und der Panda haben jeder eine Burg gebaut. Die Ente hat einen Tunnel gebaut. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was hat denn der Frosch gebaut?" Der kleine Peter antwortet: "Der Frosch hat eine Burg gebaut, und die Ente hat einen Tunnel gebaut."	The animals are playing in the sandbox. The frog and the panda each built a sand castle. The duck built a tunnel. The clown did not pay attention again and asks: "What did the frog build?" Little Peter responds: "The frog built a castle, and the duck built a tunnel."
3	Die Tiere frühstücken zusammen. Der Panda und die Ente essen Müsli. Der Frosch isst ein Brötchen. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was isst denn der Panda?" Der kleine Peter antwortet: "Der Panda isst Müsli, und der Frosch isst ein Brötchen."	The animals are having breakfast together. The panda and the duck are eating muesli. The frog is eating bun. The clown did not pay attention again and asks: "What is the panda eating?" Little Peter responds: "The panda is eating muesli, and the frog is eating bun."
4	Die Tiere sitzen im Zimmer und trinken zusammen etwas. Die Ente und der Frosch trinken Orangensaft. Der Panda trinkt Milch. Der Clown hat wieder nicht gut aufgepasst und fragt: "Was trinkt denn die Ente?" Der kleine Peter antwortet: "Die Ente trinkt Orangesaft, und der Panda trinkt Milch."	The animals sit inside and are drinking something together. The duck and the frog are drinking orange juice. The panda is drinking milk. The clown did not pay attention again and asks: "What is the duck drinking?" Little Peter responds: "The duck is drinking orange juice, and the panda is drinking milk."

Table 21 – *Continued from previous page*

	German	English
5	Die Tiere machen einen Ausflug. Der Frosch und der Panda haben jeder einen Rucksack dabei. Die Ente hat einen Regenschirm dabei. Der Clown hat wieder nicht gut aufgepasst und fragt: “Was hat denn der Frosch dabei?” Der kleine Peter antwortet: “Der Frosch hat einen Rucksack dabei und die Ente hat einen Regenschirm dabei.”	The animals are on an excursion. The frog and the panda each have a backpack. The duck has an umbrella. The clown did not pay attention again and asks: “What does the frog have?” Little Peter responds: “The frog has a backpack, and the duck has an umbrella.”

A.1.4 Further Plots

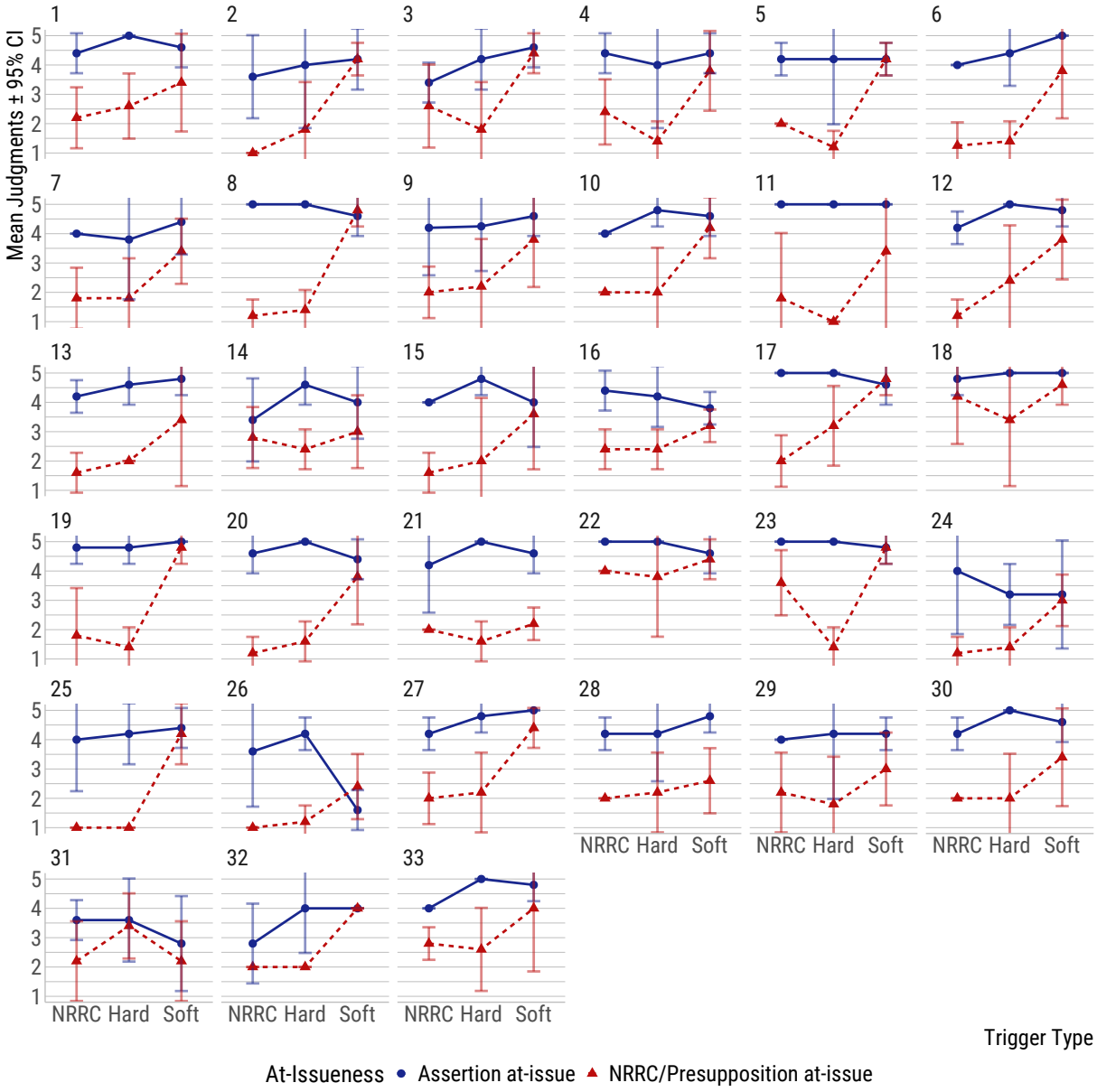


Figure 20: Exp.I: Results by participant for the adult controls.

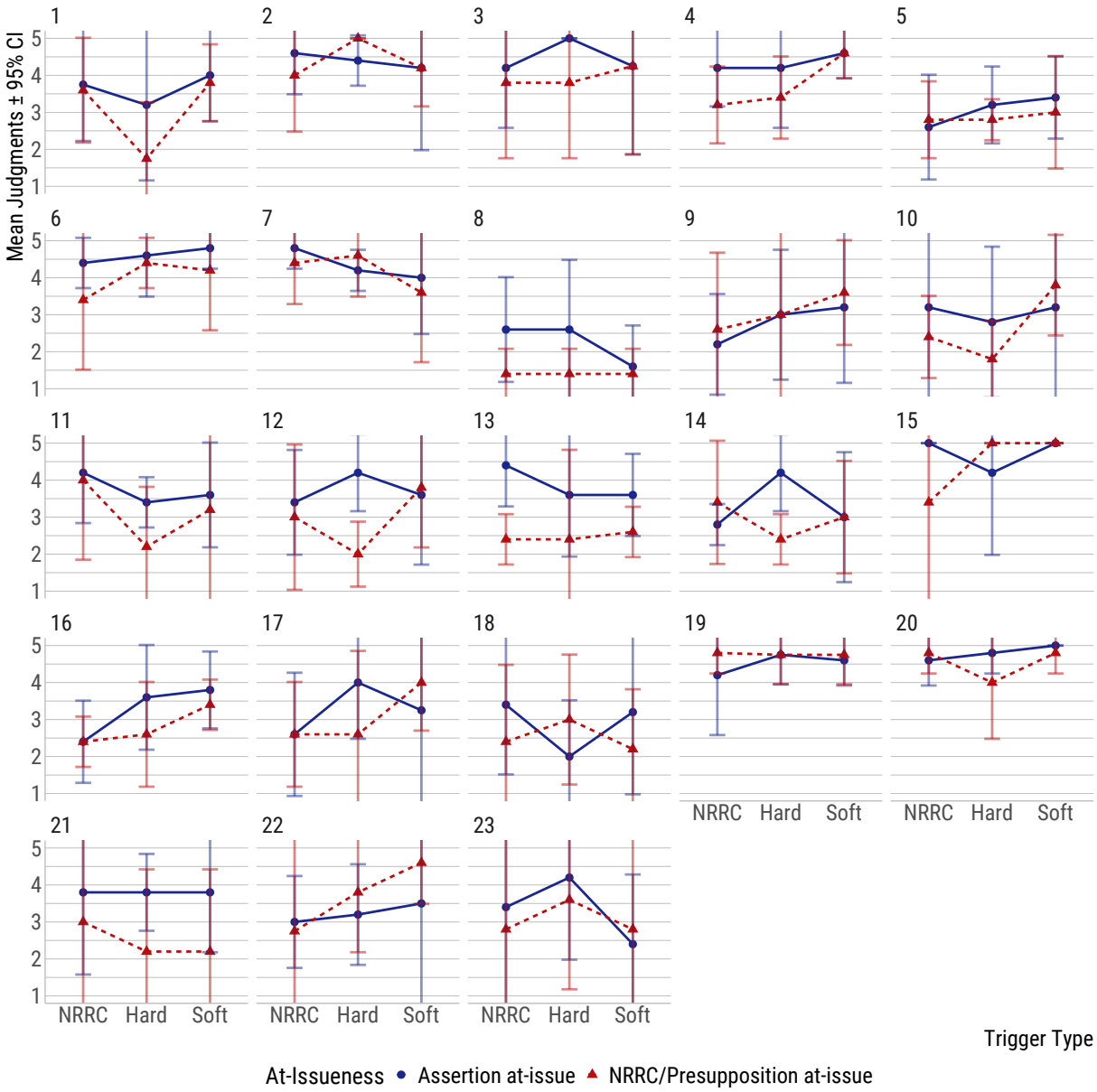


Figure 21: Exp.I: Results by participant for the younger participant group.

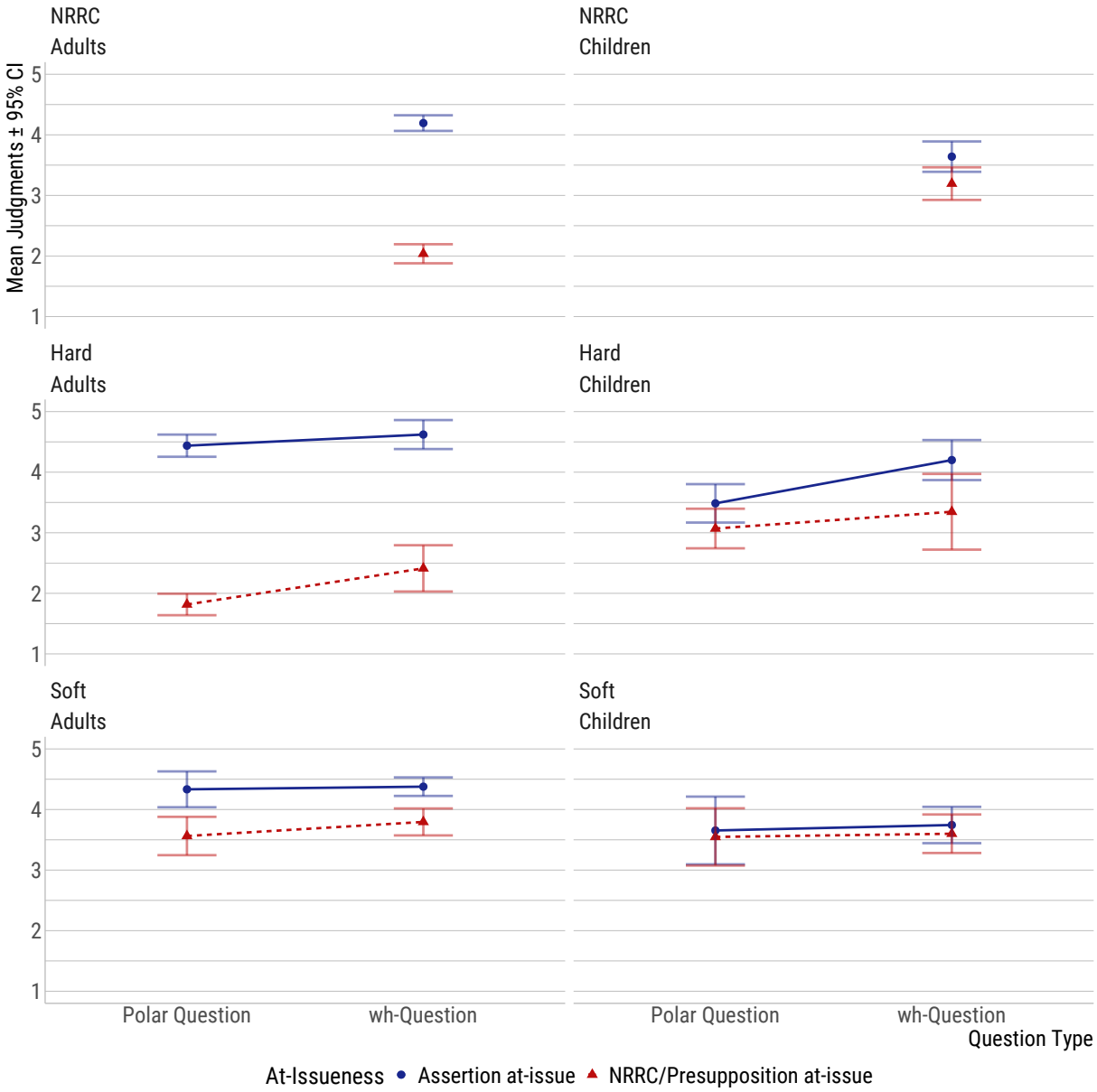


Figure 22: Exp.I: Judgments by QUESTION TYPE, TRIGGER and AGE GROUP.

A.2 Experiment II

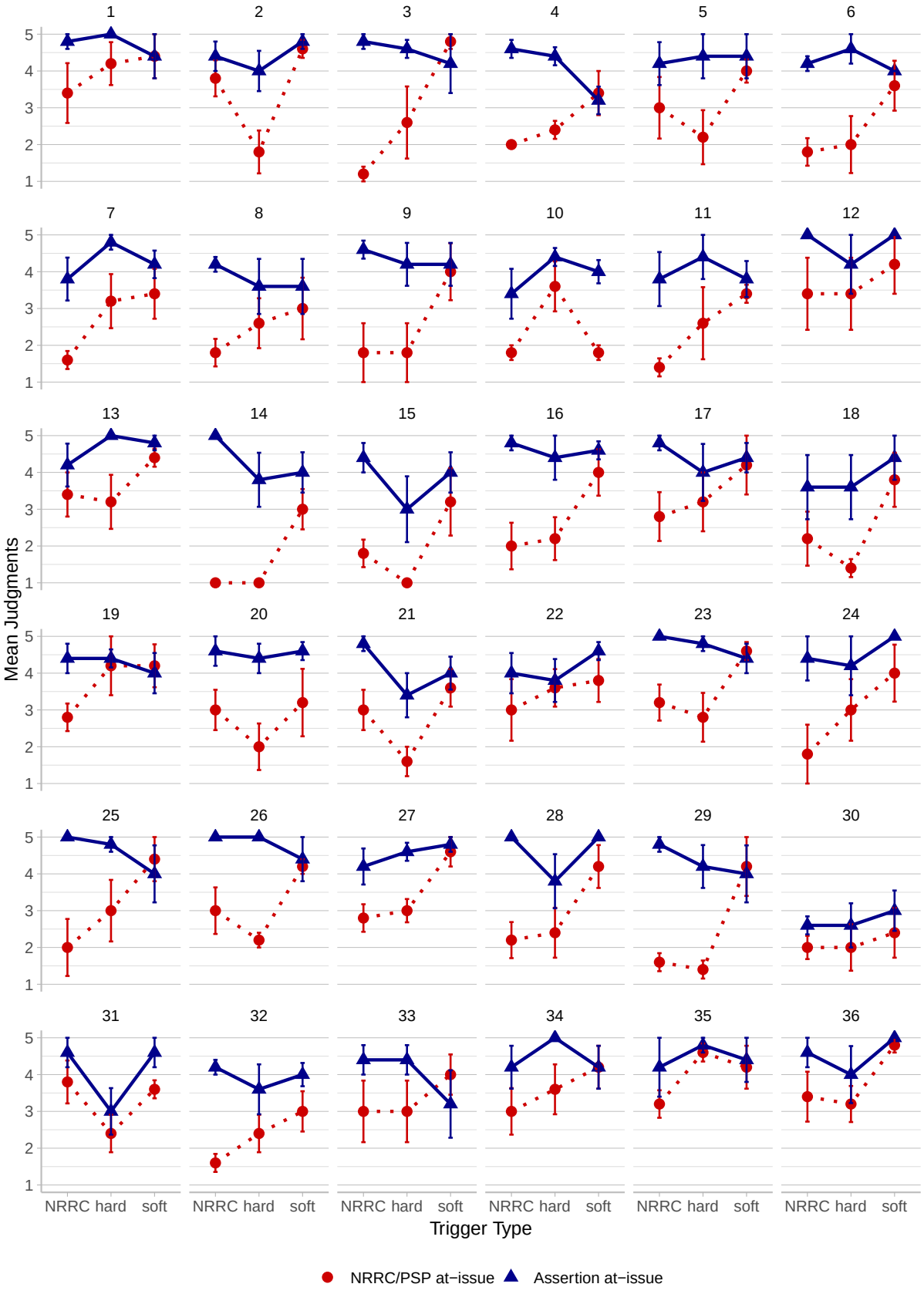


Figure 23: Exp.II: Mean judgments ± standard errors by participant for the DaF group.

A.3 Experiment III[57]

A.3.1 Critical Items

For the items with clefts below, the questions contain the two cleft-structures in Chinese, 是… 的 *(shi...de)*/… 的是 *(...de shi)*, respectively.

Table 22: Critical items in Exp.III.

	Chinese	German	English translation
		clefts	
1	甲问道：是伴娘给新娘送了一条珍珠项链吗？/给新娘送了一条珍珠项链的是伴娘吗？ 甲确定有人给新娘送了一条珍珠项链吗？	Anna fragt: War es die Trauzeugin, die der Braut eine Perlenkette geschenkt hat? Ist sich Anna sicher, dass jemand der Braut eine Perlenkette geschenkt hat?	Anna asks: Was it the maid of honor who gave the bride a pearl necklace? Is Anna certain that someone gave the bride a pearl necklace?
2	乙问道：是伴郎往新娘手花里面插了一朵百合吗？/往新娘手花里面插了一朵百合的是伴郎吗？ 乙确定有人往新娘手花里插了一朵百合吗？	Ben fragt: War es der Trauzeuge, der eine Lilie in den Brautstrauß gesteckt hat? Ist sich Ben sicher, dass jemand eine Lilie in den Brautstrauß gesteckt hat?	Ben asks: Was it the best man who put a lily in the bridal bouquet? Is Ben certain that someone put a lily in the bridal bouquet?
3	丙问道：是厨师女士在下午短暂离开了典礼吗？/下午短暂离开典礼的是厨师女士吗？ 丙确定下午有人短暂离开典礼吗？	Charlotte fragt: War es die Köchin, die die Feier nachmittags kurz verlassen hat? Ist sich Charlotte sicher, dass jemand die Feier nachmittags kurz verlassen hat?	Charlotte asks: Was it the cook who left the party shortly in the afternoon? Is Charlotte certain that someone left the party shortly in the afternoon?

[57]A part of this section is published as appendix in the journal article Y. Chen and Antomo (to appear).

Table 22 – *Continued from previous page*

	Chinese	German	English translation
4	丁问道：是发型师女士在晚饭前和王先生在酒吧见了面吗？/晚饭前和王先生在酒吧见面的是发型师女士吗？ 丁确定有人在晚饭前和王先生在酒吧见面吗？	David fragt: War es die Friseurin, die Herrn Müller kurz vor dem Abendessen an der Bar getroffen hat? Ist sich David sicher, dass jemand Herrn Müller kurz vor dem Abendessen an der Bar getroffen hat?	David asks: Was it the hairdresser who met Mr. Müller at the bar shortly before dinner? Is David certain that someone met Mr. Müller at the bar shortly before dinner?
5	戊问道：是王先生在节目表演过程中多次打电话吗？/节目表演过程中多次打电话的是王先生吗？ 戊确定节目表演过程中有人多次打电话吗？	Emma fragt: War es Herr Müller, der während der Aufführung mehrmals telefoniert hat? Ist sich Emma sicher, dass jemand während der Aufführung mehrmals telefoniert hat?	Emma asks: Was it Mr. Müller who made several phone calls during the performance? Is Emma certain that someone made several phone calls during the performance?
6	己问道：是新郎重新摆放了所有座椅吗？/重新摆放了所有座椅的是新郎吗？ 己确定有人重新摆放了所有座椅吗？	Felix fragt: War es der Bräutigam, der alle Stühle umgestellt hat? Ist sich Felix sicher, dass jemand alle Stühle umgestellt hat?	Felix asks: Was it the groom who moved all the chairs? Is Felix certain that someone moved all the chairs?
		win	
1	甲问道：新娘去年赢了 200 米自由泳的地区赛吗？ 甲确定新娘去年参加了 200 米自由泳的地区赛吗？	Anna fragt: Hat die Braut letztes Jahr das Finale über 200 Meter Freistil im regionalen Wettbewerb gewonnen? Ist sich Anna sicher, dass die Braut letztes Jahr an dem Finale über 200 Meter Freistil im regionalen Wettbewerb teilgenommen hat?	Anna asks: Did the bride win the 200 meter freestyle final in the regional competition last year? Is Anna certain that the bride participated in the 200 meter freestyle final in the regional competition last year?

Table 22 – *Continued from previous page*

	Chinese	German	English translation
2	乙问道：伴娘 2018 年赢了水球的地区赛吗？ 乙确定伴娘 2018 年参加了水球的地区赛吗？	Ben fragt: Hat die Trauzeugin 2018 den regionalen Wasserball-Wettbewerb gewonnen? Ist sich Ben sicher, dass die Trauzeugin 2018 an dem regionalen Wasserball-Wettbewerb teilgenommen hat?	Ben asks: did the maid of honor win the regional water polo competition in 2018? Is Ben certain the maid of honor participated in the regional water polo competition in 2018?
3	丙问道：酒保先生下午赢了歌唱比赛吗？ 丙确定酒保先生下午参加了歌唱比赛吗？	Charlotte fragt: Hat der Barkeeper nachmittags den Gesangswettbewerb gewonnen? Ist sich Charlotte sicher, dass der Barkeeper nachmittags an dem Gesangswettbewerb teilgenommen hat?	Charlotte asks: Did the bartender win the singing competition in the afternoon? Is Charlotte certain that the bartender participated in the singing competition in the afternoon?
4	丁问道：鼓手女士去年赢了地区音乐大赛吗？ 丁确定鼓手女士去年参加了地区音乐大赛吗？	David fragt: Hat die Schlagzeugspielerin letztes Jahr den regionalen Musikwettbewerb gewonnen? Ist sich David sicher, dass die Schlagzeugspielerin letztes Jahr an dem regionalen Musikwettbewerb teilgenommen hat?	David asks: did the percussion player win the regional music competition last year? Is David certain that the percussion player participated in the regional music competition last year?
5	戊问道：厨师女士下午赢了竞答吗？ 戊确定厨师女士下午参加了竞答吗？	Emma fragt: Hat die Köchin nachmittags das Quiz gewonnen? Ist sich Emma sicher, dass die Köchin nachmittags an dem Quiz teilgenommen hat?	Emma asks: Did the cook win the quiz in the afternoon? Is Emma certain that the cook took the quiz in the afternoon?

Table 22 – *Continued from previous page*

	Chinese	German	English translation
6	己问道：花童们下午赢了拔河比赛吗？ 己确定花童们下午参加了拔河比赛吗？	Felix fragt: Haben die Blumenkinder nachmittags das Tauziehen gewonnen? Ist sich Felix sicher, dass die Blumenkinder nachmittags an dem Tauziehen teilgenommen haben?	Felix asks: Did the flower children win the tug-of-war in the afternoon? Is Felix certain that the flower children took part in the tug-of-war in the afternoon?
		regret	
1	甲问道：伴郎后悔他在茶歇期间吃了四块胡萝卜蛋糕了吗？ 甲确定伴郎在茶歇期间吃了四块胡萝卜蛋糕吗？	Anna fragt: Bereut der Trauzeuge, dass er während der Kaffeepause vier Stück Karottenkuchen gegessen hat? Ist sich Anna sicher, dass der Trauzeuge während der Kaffeepause vier Stück Karottenkuchen gegessen hat?	Anna asks: Does the best man regret that he ate four pieces of carrot cake during the coffee break? Is Anna certain that the best man ate four pieces of carrot cake during the coffee break?
2	乙问道：新娘后悔她邀请了六位花童吗？ 乙确定新娘邀请了六位花童吗？	Ben fragt: Bereut die Braut, dass sie sechs Blumenkinder eingeladen hat? Ist sich Ben sicher, dass die Braut sechs Blumenkinder eingeladen hat?	Ben asks: Does the bride regret that she invited six flower children? Is Ben certain that the bride invited six flower children?
3	丙问道：鼓手女士后悔她选择了一件红衬衫吗? 丙确定鼓手女士选择了一件红衬衫吗？	Charlotte fragt: Bereut die Schlagzeugspielerin, dass sie sich für eine rote Bluse entschieden hat? Ist sich Charlotte sicher, dass sich die Schlagzeugspielerin für eine rote Bluse entschieden hat?	Charlotte asks: Does the percussion player regret that she chose a red blouse? Is Charlotte certain that the percussion player chose a red blouse?

Table 22 – *Continued from previous page*

	Chinese	German	English translation
4	丁问道：烘培师女士后悔她在烘焙时使用了杏仁泥吗？ 丁确定烘培师女士在烘焙时使用了杏仁泥吗？	David fragt: Bereut die Konditorin, dass sie beim Backen Marzipan verwendet hat? Ist sich David sicher, dass die Konditorin beim Backen Marzipan verwendet hat?	David asks: Does the pastry chef regret that she used marzipan in her baking? Is David certain that the pastry chef used marzipan in the baking?
5	戊问道：录影师先生后悔他在饭前和伴娘眉来眼去了吗？ 戊确定录影师先生在饭前和伴娘眉来眼去了吗？	Emma fragt: Bereut der Kameramann, dass er vor dem Essen mit der Trauzeugin geflirtet hat? Ist sich Emma sicher, dass der Kameramann vor dem Essen mit der Trauzeugin geflirtet hat?	Emma asks: Does the cameraman regret that he flirted with the maid of honor before the meal? Is Emma certain that the cameraman flirted with the maid of honor before the meal?
6	己问道：主持人先生后悔他在午餐后又吃了三个冰淇淋球吗？ 己确定主持人先生在午餐后又吃了三个冰淇淋球吗？	Felix fragt: Bereut der Moderator, dass er nach dem Mittagessen noch drei Kugeln Eis gegessen hat? Ist sich Felix sicher, dass der Moderator nach dem Mittagessen noch drei Kugeln Eis gegessen hat?	Felix asks: Does the emcee regret that he ate three scoops of ice cream after lunch? Is Felix certain that the emcee ate three scoops of ice cream after lunch?
		discover	
1	丁问道：伴娘发现了有一只化妆刷不见了吗？ 丁确定有一只化妆刷不见了吗？	David fragt: Hat die Trauzeugin entdeckt, dass ein Make-up-Pinsel verschwunden war? Ist sich David sicher, dass ein Make-up-Pinsel verschwunden war?	David asks: Did the maid of honor discover that a makeup brush was missing? Is David certain that a make-up brush was missing?

Table 22 – *Continued from previous page*

	Chinese	German	English translation
2	戊问道：花艺师先生发现了有许多向日葵在晚餐后不见了吗？ 戊确定有许多向日葵在晚餐后不见了吗？	Emma fragt: Hat der Florist entdeckt, dass viele Sonnenblumen nach dem Abendessen weg waren? Ist sich Emma sicher, dass viele Sonnenblumen nach dem Abendessen weg waren?	Emma asks: Did the florist discover that many sunflowers were gone after dinner? Is Emma certain that many sunflowers were gone after dinner?
3	己问道：录影师先生发现了新娘晚餐前换了发饰吗？ 己确定新娘晚餐前换了发饰吗？	Felix fragt: Hat der Kameramann entdeckt, dass die Braut vor dem Abendessen ihren Haarschmuck gewechselt hat? Ist sich Felix sicher, dass die Braut vor dem Abendessen ihren Haarschmuck gewechselt hat?	Felix asks: Did the cameraman discover that the bride changed her hair ornament before dinner? Is Felix certain that the bride changed her hair ornament before dinner?
4	甲问道：主持人先生发现了有一只话筒在典礼后坏了吗？ 甲确定有只话筒在典礼后坏了吗？	Anna fragt: Hat der Moderator entdeckt, dass nach der Trauung ein Mikrofon defekt war? Ist sich Anna sicher, dass nach der Trauung ein Mikrofon defekt war?	Anna asks: Did the emcee discover that a microphone was defective after the wedding ceremony? Is Anna certain that a microphone was defective after the wedding ceremony?
5	乙问道：厨师女士发现了王先生晚餐期间一口没碰鱼吗？ 乙确定王先生晚餐期间一口没碰鱼吗？	Ben fragt: Hat die Köchin entdeckt, dass Herr Müller beim Abendessen keinen Bissen von dem Fischgericht angerührt hat? Ist sich Ben sicher, dass Herr Müller beim Abendessen keinen Bissen von dem Fischgericht angerührt hat?	Ben asks: Did the cook discover that Mr. Müller didn't touch a bite of the fish dish at dinner? Is Ben certain that Mr. Miller didn't take a bite of the fish dish at dinner?

Table 22 – *Continued from previous page*

	Chinese	German	English translation
6	丙问道：新郎发现了他的父亲和王先生是大学同学了吗？ 丙确定新郎的父亲和王先生是大学同学吗？	Charlotte fragt: Hat der Bräutigam entdeckt, dass sein Vater und Herr Müller Studienkollegen waren? Ist sich Charlotte sicher, dass der Vater des Bräutigams und Herr Müller Studienkollegen waren?	Charlotte asks: Did the groom discover that his father and Mr. Müller were fellow students? Is Charlotte certain that the groom's father and Mr. Müller were fellow students?

A.3.2 Control Items

For the items with clefts below, the questions contain the two cleft-structures in Chinese, 是… 的 *(shi...de)*/… 的是 *(...de shi)*, respectively.

Table 23: Control items in Exp.III.

	Chinese	German	English translation
		clefts	
1	丁问道：是发型师女士点了一份无乳糖饮食吗？/点了一份无乳糖饮食的是发型师女士吗？ 丁确定发型师女士点了一份无乳糖饮食吗？	David fragt: War es die Friseurin, die ein laktosefreies Menü bestellt hat? Ist sich David sicher, dass die Friseurin ein laktosefreies Menü bestellt hat?	David asks: Was it the hairdresser who ordered a lactose-free menu? Is David certain that the hairdresser ordered a lactose-free menu?
2	戊问道：是歌手先生偷偷给新郎拍了照吗？/偷偷给新郎拍了照的是歌手先生吗？ 戊确定歌手先生偷偷给新郎拍了照吗？	Emma fragt: War es der Sänger, der den Bräutigam heimlich fotografiert hat? Ist sich Emma sicher, dass der Sänger den Bräutigam heimlich fotografiert hat?	Emma asks: Was it the singer who secretly photographed the groom? Is Emma certain it was the singer who secretly photographed the groom?

Table 23 – *Continued from previous page*

	Chinese	German	English translation
3	己问道：是新郎为庆典挑选了音乐吗？/为庆典挑选了音乐的是新郎吗？ 己确定新郎为庆典挑选了音乐吗？	Felix fragt: War es der Bräutigam, der die Musik für die Feier ausgesucht hat? Ist sich Felix sicher, dass der Bräutigam die Musik für die Feier ausgesucht hat?	Felix asks: Was it the groom who chose the music for the ceremony? Is Felix certain that the groom chose the music for the celebration?
		win	
1	甲问道：王先生下午赢了扎花束比赛吗？ 甲确定王先生下午赢了扎花束比赛吗？	Anna fragt: Hat Herr Müller nachmittags den Wettbewerb im Blumenstraußbinden gewonnen? Ist sich Anna sicher, dass Herr Müller nachmittags den Wettbewerb im Blumenstraußbinden gewonnen hat?	Anna asks: Did Mr. Müller win the bouquet tying competition in the afternoon? Is Anna certain that Mr. Müller won the bouquet tying competition in the afternoon?
2	乙问道：照相师女士去年赢了环保协会的摄影大赛吗？ 乙确定照相师女士去年赢了环保协会的摄影大赛吗？	Ben fragt: Hat die Fotografin letztes Jahr den Fotowettbewerb des Naturschutzvereins gewonnen? Ist sich Ben sicher, dass die Fotografin letztes Jahr den Fotowettbewerb des Naturschutzvereins gewonnen hat?	Ben asks: Did the photographer win the Nature Preservation Association's photo contest last year? Is Ben certain that the photographer won the Nature Preservation Association's photo contest last year?

Table 23 – *Continued from previous page*

	Chinese	German	English translation
3	丙问道：厨师女士 2015 年赢了厨艺大赛吗？ 丙确定厨师女士 2015 年赢了厨艺大赛吗？	Charlotte fragt: Hat die Köchin 2015 den Kochwettbewerb gewonnen? Ist sich Charlotte sicher, dass die Köchin 2015 den Kochwettbewerb gewonnen hat?	Charlotte asks: Did the cook win the cooking competition in 2015? Is Charlotte certain that the cook won the cooking competition in 2015?
		regret	
1	丁问道：照相师女士后悔她在婚礼前拜访了王先生吗？ 丁确定照相师女士后悔她在婚礼前拜访了王先生吗？	David fragt: Bereut die Fotografin, dass sie vor der Hochzeit Herrn Müller besucht hat? Ist sich David sicher, dass die Fotografin bereut, dass sie vor der Hochzeit Herrn Müller besucht hat?	David asks: Does the photographer regret that she visited Mr. Müller before the wedding? Is David certain that the photographer regrets that she visited Mr. Müller before the wedding?
2	戊问道：伴娘后悔她早上打磨了镜子吗？ 戊确定伴娘后悔她早上打磨了镜子吗？	Emma fragt: Bereut die Trauzeugin, dass sie morgens die Spiegel poliert hat? Ist sich Emma sicher, dass die Trauzeugin bereut, dass sie morgens die Spiegel poliert hat?	Emma asks: Does the maid of honor regret that she polished the mirrors in the morning? Is Emma certain that the maid of honor regrets that she polished the mirrors in the morning?

Table 23 – *Continued from previous page*

	Chinese	German	English translation
3	己问道：花童们后悔他们下午把糖果藏在了鲜花下面吗？ 己确定花童们后悔他们下午把糖果藏在了鲜花下面吗？	Felix fragt: Bereuen die Blumenkinder, dass sie nachmittags Bonbons unter den Blumen versteckt haben? Ist sich Felix sicher, dass die Blumenkinder bereuen, dass sie nachmittags Bonbons unter den Blumen versteckt haben?	Felix asks: Do the flower children regret that they hid candy under the flowers in the afternoon? Is Felix certain that the flower children regret that they hid candy under the flowers in the afternoon?
		discover	
1	甲问道：伴娘发现了她的发夹滑到了餐桌下面吗？ 甲确定伴娘发现了她的发夹滑到了餐桌下面吗？	Anna fragt: Hat die Trauzeugin entdeckt, dass ihre Haarspange unter einen Tisch gerutscht ist? Ist sich Anna sicher, dass die Trauzeugin entdeckt hat, dass ihre Haarspange unter einen Tisch gerutscht ist?	Anna asks: Did the maid of honor discover that her hair clip slipped under a table? Is Anna certain that the maid of honor discovered that her hair clip slipped under a table?
2	乙问道：伴郎发现了他的母亲和王先生大学时是一对儿了吗？ 乙确定伴郎发现了他的母亲和王先生大学时是一对儿了吗？	Ben fragt: Hat der Trauzeuge entdeckt, dass seine Mutter und Herr Müller während des Studiums ein Paar waren? Ist sich Ben sicher, dass der Trauzeuge entdeckt hat, dass seine Mutter und Herr Müller während des Studiums ein Paar waren?	Ben asks: Did the best man discover that his mother and Mr. Müller were a couple during their studies? Is Ben certain that the best man discovered that his mother and Mr. Müller were a couple during their studies?

Table 23 – *Continued from previous page*

	Chinese	German	English translation
3	丙问道：新娘发现了她的母亲在典礼前哭了吗？ 丙确定新娘发现了她的母亲在典礼前哭了吗？	Charlotte fragt: Hat die Braut entdeckt, dass ihre Mutter vor der Trauung geweint hat? Ist sich Charlotte sicher, dass die Braut entdeckt hat, dass ihre Mutter vor der Trauung geweint hat?	Charlotte asks: Did the bride discover that her mother cried before the wedding ceremony? Is Charlotte certain that the bride discovered that her mother cried before the wedding ceremony?

A.3.3 Filler Items

Table 24: Filler items in Exp.III.

	Chinese	German	English translation
1	甲问道：新郎在更衣室里藏了一朵玫瑰吗？ 甲确定新郎在更衣室里藏了一朵玫瑰吗？	Anna fragt: Hat der Bräutigam im Umkleidezimmer eine Rose versteckt? Ist sich Anna sicher, dass der Bräutigam im Umkleidezimmer eine Rose versteckt hat?	Anna asks: Did the groom hide a rose in the dressing room? Is Anna certain that the groom hid a rose in the dressing room?
2	乙问道：录影师先生晚餐前休息了一下吗？ 乙确定录影师先生晚餐前休息了一下吗？	Ben fragt: Hat der Kameramann vor dem Abendessen eine Pause gemacht? Ist sich Ben sicher, dass der Kameramann vor dem Abendessen eine Pause gemacht hat?	Ben asks: Did the cameraman take a break before dinner? Is Ben certain that the cameraman took a break before dinner?
3	丙问道：新娘的父亲典礼期间看了好多次手机吗？ 丙确定新娘的父亲典礼期间看了好多次手机吗？	Charlotte fragt: Hat der Vater der Braut während der Trauung mehrmals auf sein Handy geschaut? Ist sich Charlotte sicher, dass der Vater der Braut während der Trauung mehrmals auf sein Handy geschaut hat?	Charlotte asks: Did the bride's father look at his cell phone several times during the wedding ceremony? Is Charlotte certain that the father of the bride looked at his cell phone several times during the wedding ceremony?

Table 24 – *Continued from previous page*

	Chinese	German	English translation
4	甲问道：歌手先生说了他今天只吃了蔬菜吗？ 甲确定歌手先生今天只吃了蔬菜吗？	Anna fragt: Hat der Sänger gesagt, dass er heute nur Gemüse gegessen hat? Ist sich Anna sicher, dass der Sänger heute nur Gemüse gegessen hat?	Anna asks: Did the singer say that he ate only vegetables today? Is Anna certain that the singer ate only vegetables today?
5	乙问道：花艺师先生说了他上午装饰了舞台吗？ 乙确定花艺师先生上午装饰了舞台吗？	Ben fragt: Hat der Florist gesagt, dass er vormittags die Bühne dekoriert hat? Ist sich Ben sicher, dass der Florist vormittags die Bühne dekoriert hat?	Ben asks: Did the florist say that he decorated the stage in the morning? Is Ben certain that the florist decorated the stage in the morning?
6	丙问道：调酒师先生说了他派对期间只喝了水吗？ 丙确定调酒师先生派对期间只喝了水吗？	Charlotte fragt: Hat der Barkeeper gesagt, dass er während der Party nur Wasser getrunken hat? Ist sich Charlotte sicher, dass der Barkeeper während der Party nur Wasser getrunken hat?	Charlotte asks: Did the bartender say that he only drank water during the party? Is Charlotte certain that the bartender only drank water during the party?
7	丁问道：王先生下午参加了唱歌比赛吗？ 丁确定王先生下午参加了唱歌比赛吗？	David fragt: Hat Herr Müller nachmittags am Gesangswettbewerb teilgenommen? Ist sich David sicher, dass Herr Müller nachmittags am Gesangwettbewerb teilgenommen hat?	David asks: Did Mr. Müller participate in the singing competition in the afternoon? Is David certain that Mr. Müller participated in the singing competition in the afternoon?

Table 24 – *Continued from previous page*

	Chinese	German	English translation
8	戊问道: 主持人先生下午参加了扎花束比赛吗？ 戊确定主持人先生下午参加了扎花束比赛吗？	Emma fragt: Hat der Moderator nachmittags an dem Wettbewerb im Blumenstraußbinden teilgenommen? Ist sich Emma sicher, dass der Moderator nachmittags an dem Wettbewerb im Blumenstraußbinden teilgenommen hat?	Emma asks: Did the emcee take part in the bouquet tying competition in the afternoon? Is Emma certain that the emcee took part in the bouquet tying competition in the afternoon?
9	己问道：花童们下午参加了竞答吗？ 己确定花童们下午参加了竞答吗？	Felix fragt: Haben die Blumenkinder nachmittags am Quiz teilgenommen? Ist sich Felix sicher, dass die Blumenkinder nachmittags am Quiz teilgenommen haben?	Felix asks: Did the flower children take part in the quiz in the afternoon? Is Felix certain that the flower children took part in the quiz in the afternoon?
10	丙问道：鼓手女士认为宾客们更喜欢快歌吗？ 丙确定宾客们更喜欢快歌吗？	David fragt: Glaubt die Schlagzeugspielerin, dass den Gästen schnellere Lieder besser gefallen? Ist sich David sicher, dass den Gästen schnellere Lieder besser gefallen?	David asks: Does the percussion player think that the guests like faster songs better? Is David certain that the guests like faster songs better?
11	戊问道：厨师女士认为大多数宾客更喜欢肉菜吗？ 戊确定大多数宾客更喜欢肉菜吗？	Emma fragt: Glaubt die Köchin, dass die meisten Gäste das Fleischgericht bevorzugen? Ist sich Emma sicher, dass die meisten Gäste das Fleischgericht bevorzugen?	Emma asks: Does the cook think that most guests prefer the meat dish? Is Emma certain that most of the guests prefer the meat dish?

Table 24 – *Continued from previous page*

	Chinese	German	English translation
12	丁问道：照相师女士认为花童们吃光了巧克力吗？ 丁确定花童们吃光了巧克力吗？	Felix fragt: Glaubt die Fotografin, dass die Blumenkinder die Schokolade aufgegessen haben? Ist sich Felix sicher, dass die Blumenkinder die Schokolade aufgegessen haben?	Felix asks: Does the photographer think that the flower children ate all the chocolate? Is Felix certain that the flower children ate all the chocolate?

A.3.4 Further Plots

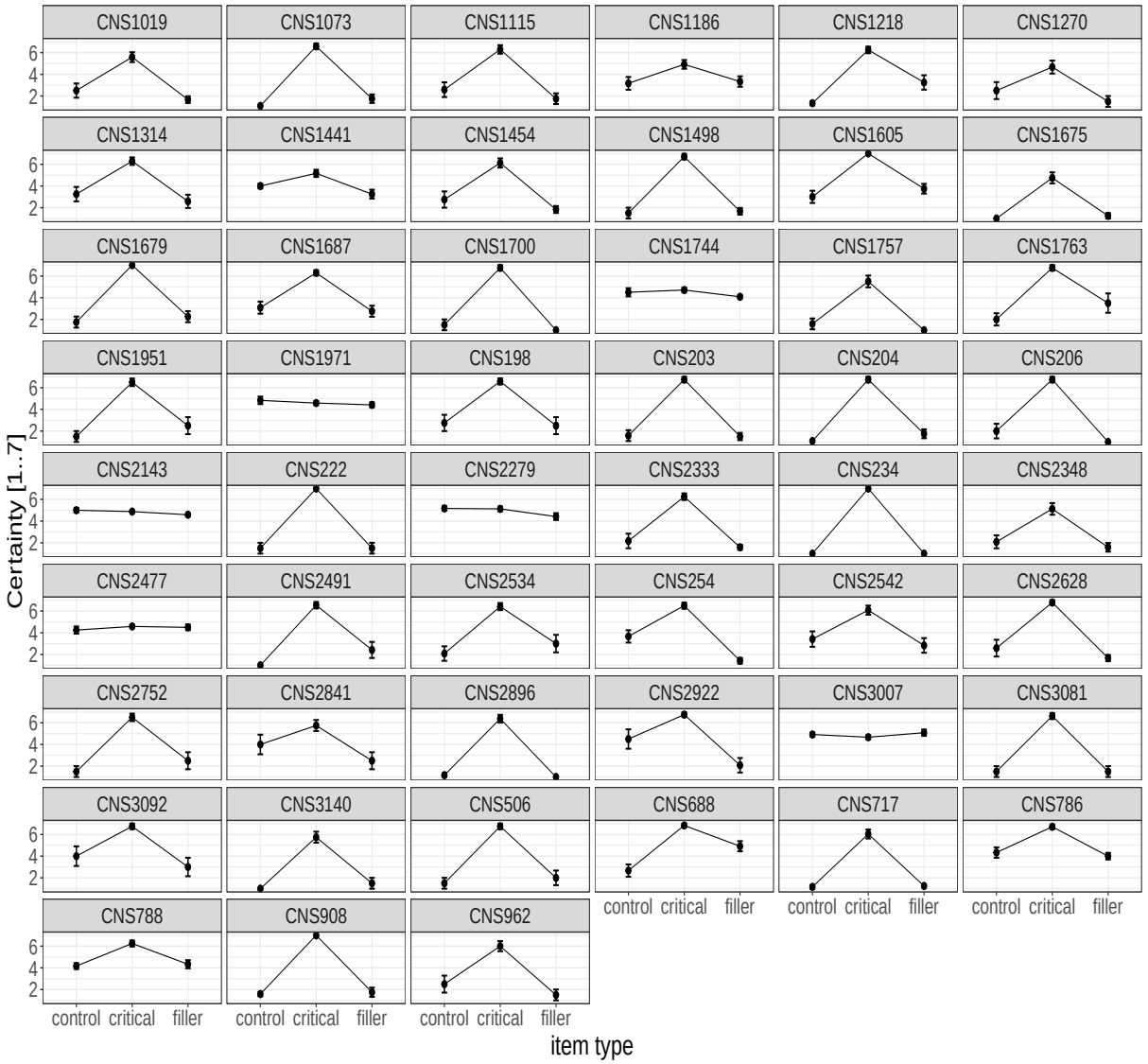


Figure 24: Exp.III: Attention check: mean ratings per item type ± standard errors for the group CNS. n=51, 6 participants discarded as they did not pass the attention and understanding check: CNS1744, CNS1971, CNS2143, CNS2279, CNS2477, CNS3007.

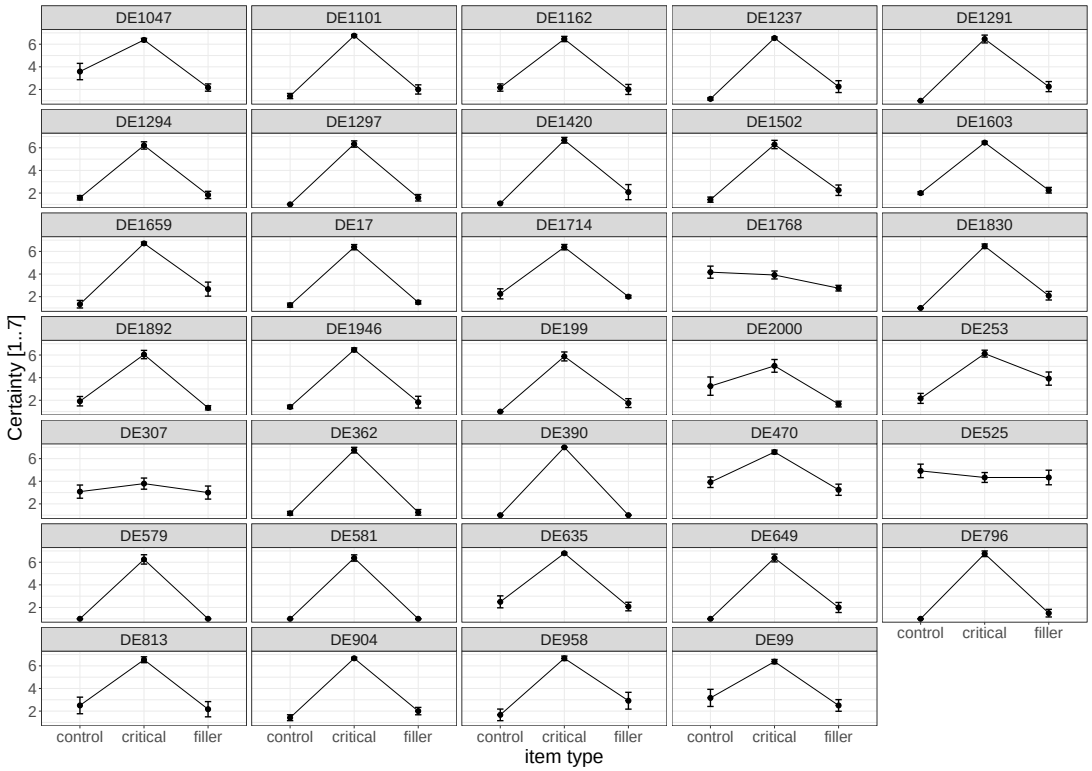


Figure 25: Exp.III: Attention check: mean ratings per item type ± standard errors for the group DE. n=34, 2 participants discarded as they did not pass the attention and understanding check: DE1768, DE525.

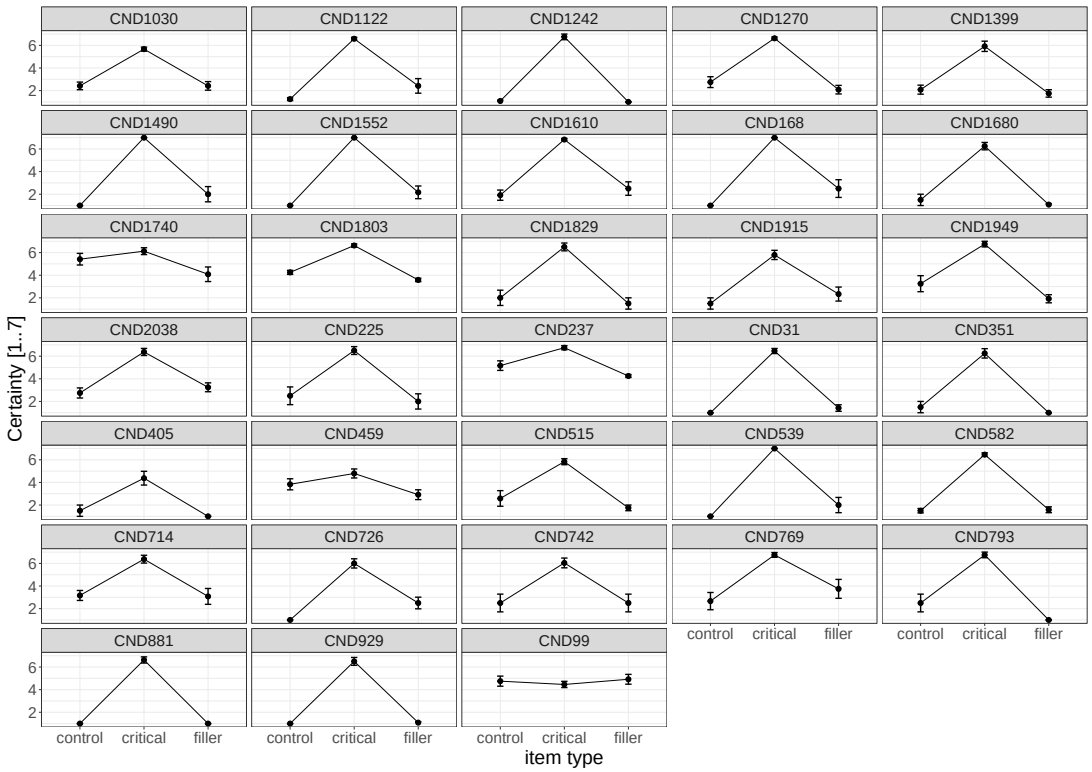


Figure 26: Exp.III: Attention check: mean ratings per item type ± standard errors for the group CND. n=33, one participants discarded as s/he did not pass the attention and understanding check: CND99.

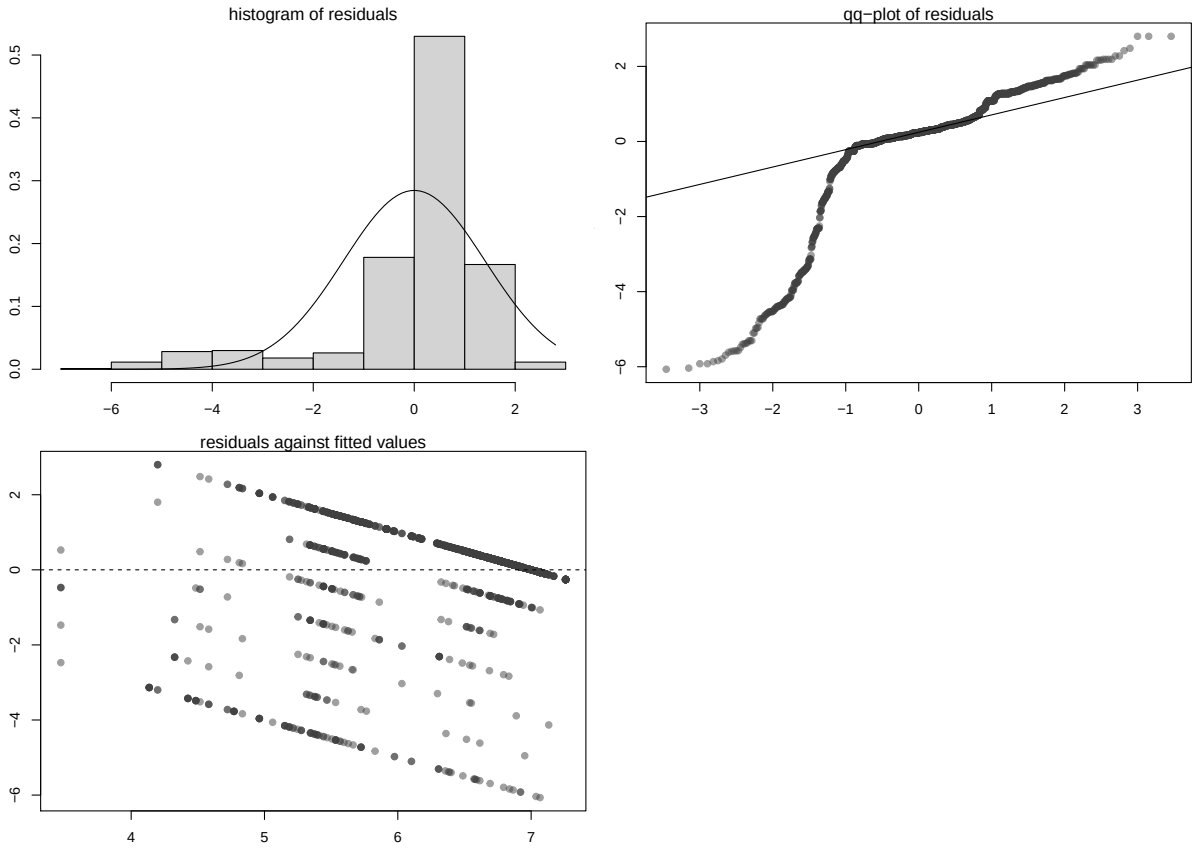


Figure 27: Exp.IIIa: Plots of residuals against the LMM fitting.

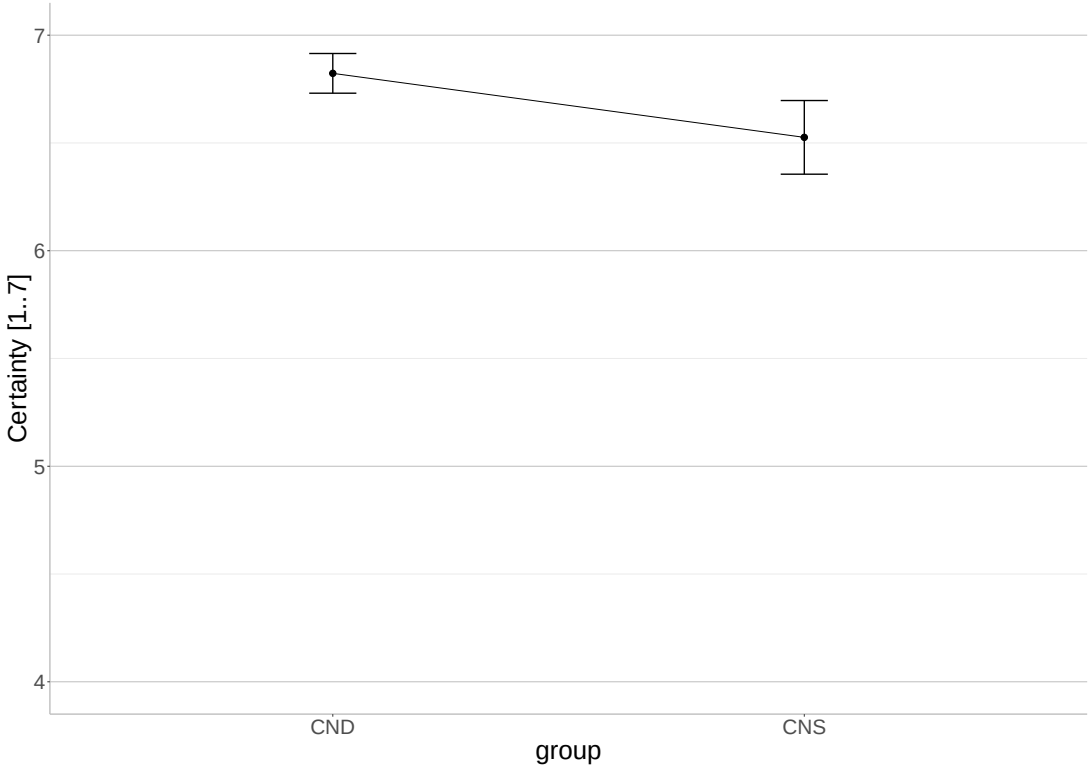


Figure 28: Exp.IIIb: Mean judgments ± 95% confidence interval for the factor Group.

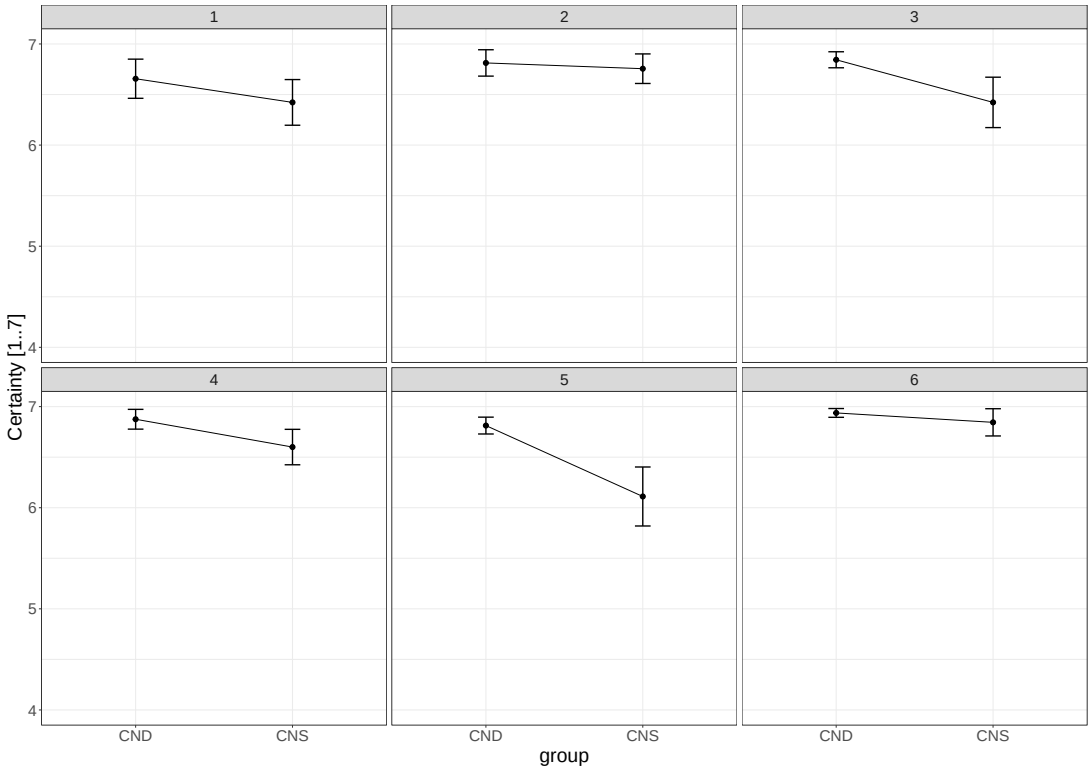


Figure 29: Exp.IIIb: Meaning ratings of each critical cleft items.

Versicherung

Hiermit versichere ich an Eides statt, dass ich die schriftliche wissenschaftliche Abhandlung (Dissertation) selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, und dass anderweitig keine entsprechende Promotion beantragt wurde und hierbei die eingereichte Dissertation oder Teile daraus vorgelegt worden sind.

Des Weiteren erkläre ich hiermit, dass diese digitale Fassung mit der schriftlichen wissenschaftlichen Abhandlung übereinstimmt.