Neural and developmental bases of processing language outside the here-and-now

by

Maxime Alexandra Tulling

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

May, 2022

Ailís Cournane

Liina Pylkkänen

DEDICATION

Tijd is iets en tijd is niets. ‘t is een vliegtuig of een fiets.

“Time is something and time is nothing. It’s an airplane or a bike.”

-- Toon Hermans met een vleugje Vonke

ACKNOWLEDGMENTS

I do not take for granted the amount of joy, laughter, and warmth I have experienced over the last six years, all of which I owe to the wonderful people that have surrounded me during this journey. First and foremost, I would like to thank my co-advisors Ailís Cournane and Liina Pylkkänen for their support and guidance throughout this PhD. I first met Ailís when we both started at NYU, and what started out with a short chat ended up in a valuable mentorship. Ailís fueled my interest in child language acquisition, and her enthusiasm, optimism and creativity have been a huge support and inspiration. Ailís, thank you so much for the countless hours of meetings, amazing stories, Dutch word jokes, dinosaur cookies, and endless generosity. I feel very lucky to have been there for the birth of the Child Language Lab and cannot wait to see how it grows. I am equally grateful to my other advisor Liina Pylkkänen, for welcoming me into her lab. Liina always impressed me with her hawk eye for typos and inconsistencies, but her most impressive skill must be her ability to see potential — in ideas, opportunities, and people. Liina, you have always encouraged me to shoot higher than I dared to aim for, and I am extremely grateful for your belief and support over the last six years. I am also grateful for the advice and support I received from Sudha Arunachalam. Thank you, Sudha, for asking tough questions and providing feedback on numerous (revised) experimental designs. Many thanks as well to Alec Marantz and Ana Arregui, for being on this dissertation committee, and sharing their valuable comments and suggestions. I also thank Alec for all the fun facts shared over the years and his very decent taste in licorice.

I want to thank the other professors at NYU for teaching me: Anna Szabolcsi, Lucas Champollion, Chris Collins, Maria Gouskova, Gregory Guy, Stephanie Harves, Richard Kayne and Gary Thoms. Also, thanks to Doug Whalen and Ofer Tchernichovski for their inspiring course "Birds and Babies" at CUNY. I am happy to have worked some with Valetine Hacquard and want to thank Brian McElree and Angelika Kratzer for the feedback they provided as committee members on prior QP/proposal stages of this work. I am also still grateful for all the people who taught me during my undergraduate and master at Leiden University, especially Lisa Cheng, Leticia Pablos-Robles, Stella Gryllia, Yiya Chen, Claartje Levelt, Niels Schiller and Jos Pacilly, many lessons I learnt from you I have applied and passed forward during my PhD.

I owe my thanks to all the members of the NYU Linguistics Department for creating a working environment that felt safe, collaborative, quirky and warm. Thank you for your lovely company during awkward elevator rides, spontaneous kitchen encounters, snack-and-shares, LANYU events and after-colloquium receptions.

This starts with the remaining faculty, Chris Barker, Renée Blake, Sam Bowman, Lisa Davidson, Gillian Gallagher, Laurel MacKenzie, Juliet Stanton and John Singler, but also pertains to the administrative staff, Teresa Leung, Teresa Colaizzo, Star Jimenez, Seena Berg, Hannah Katz, and security guards Mike, Eddie, Robbie and Daryl. Our everyday encounters were pleasant and warm, and I am thankful for all the help that you offered throughout the years.

Moreover, I want to thank the PhD-students that started before me for welcoming me into the department — inviting me to celebrations, trivia nights and game nights — but also for sharing your experiences, knowledge and resources. Thank you Dunja, Yohei, Suzy, Isaac, Zach, Hoaze, Becky, Daniel, Linmin, Natalie, Sonia, Allison, Itamar, Nicole, and especially Yining, Ildi, Masha, Mary, Dan and my mentor Adina. Special shout-out to Paloma for being an amazing collaborator and person, and Sheng-Fu for documenting linguists enjoying life and providing solid, often much-needed company in quiet hours. Thank you, Maria, Woojin and Songhee for welcoming me into the 603 office space, I especially enjoyed when the door closed, and the juicy stories started. Of course, I also want to thank the linguists I got to welcome to the department myself: Omar, Kate, Guy, Anna, Jai, Kimberley, Zhuoye, Soo-Hwan, Özlem and Selikem. I specifically want to thank Omar, Nigel, Alex and Buddy for being wonderful office mates, and Alicia Chatten for her friendship, help and exciting snack escapades. Thanks to José, Simone, Kaustubh, Ross, Cara and Alden for being an awesome batch of first years. I am very happy we overlapped for during our studies and wish you all the best for the years to come!

Next, I want to thank the undergraduate students that I got to teach, mentor and work with. Thank you Maya Orey, Mark Bacon, Yufei Li and Johanna Bunn for working on the counterfactual project, and Michael Marinaccio for working with us on natdisc, it has been a joy. Thank you Starr Sandoval for staying in touch after graduation, looking forward to visiting you at UBC!

I have also enjoyed getting to know several visiting students. Thank you, Kata, Karoliina, Yuchen, Paula, Thom and Steven for all the interesting discussions. I have especially enjoyed my conversations about syntax, scope and life with Hongchen Wu, and hope we will continue them in the future with some more boba. Relatedly, I had the pleasure to be a visitor myself at the University of Maryland. Thank you UMD ModSquad, for taking me in during my visists: Anouk, Mina, Nick, Gesoel and Yuhan, and especially my collaborator Annemarie van Dooren for her hospitability and gezelligheid. I am also especially grateful for the hours and hours spent talking to Yu-an Yang, and fondly remember singing Dutch and Mandarin songs in the Boston rain.

It is hard to imagine how grad school would have been without the support of my peers. First, I would like to thank all past and current members of the Neuroscience of Language lab for their suggestions and support, especially Alicia Parish, Sarah Phillips, Ellie Abrams, Graham Flick, Laura Gwilliams, Marco Lai, Ebony Goldman, Jacqui Fallon, Alicia Mason, Marianne Azar and Aline-Priscillia for great work outings and long nellunches. Special thanks to Julien Dirani, for being the best of work friends. I am also very grateful for the time spent in Abu Dhabi with Alhanoof Albastaki, Ben Lang, Dustin Chacón, Jixing Li, Samantha Wray and my lovely collaborator Ryan Law, who always impresses me with his aptitude and ambition. Ryan, Ben, Alhanoof and Sam: thank you so much for making me laugh, cry and laugh even more with your silly puns and slack responses. Second, I like to thank all the members of the Child Language Lab and LEARN lab for their suggestions and support. Special thanks to Naomi Lee, Ioana Grosu, Sarah and Alicia for thoughtful discussions on experimental design, and Chiara Repetti-Ludlow for her help with participant recruitment and testing. Most of all, I want to thank Vishal Arvindam for his contagious enthusiasm and all the time spent working and not working together.

Last, I am immensely grateful for the support from my (adopted) cohort: Hagen Blix, Sora Yin, Maddie Gilbert, Suhail Matar and Sandy Abu El Adas.

Thank you, Hagen, for procrastinating with me, letting me celebrate your birthday, and teaching me how to break into places. Sora I would like to thank for all the shows watched, games played and dumplings burned. While Maddie and I started out with staring matches trying to figure each other out, I love that we ended up being such a good team. Maddie, I really appreciate all our good times spent and all your help throughout the years (setting up Dutch printers, rating and recording stimuli, and reminding me of most deadlines). I cannot imagine having clicked any of the hard and important buttons without you. Next, I have an especially warm place in my heart for Suhail, who I really believe I could not have done any of this PhD without, or if I could have, I would for sure not have enjoyed the process as much. Suhail, I cannot express enough how grateful I have been for your patience, generosity, and kindness. Thank you and Ali for all the food, jokes and comfort shared over the years. As for Sandy, besides being an incredibly warm and generous friend, she has also functioned as my unofficial third advisor. Sandy, we have had many long walks where you listened to my struggles and uncertainties, and helped me formulate plans and make decisions. You have always cheered me on, and I am very thankful for your energy and love. I also feel very lucky that we got to spend the last year living together, it has truly been a blast.

As for my other roommates at 265 Clifton Place, thank you so much for making this house feel like a home away from home. This starts with the original roommate, Francesco Preta, known by some for squatting the 6th floor of the linguistics department, who I got to experience life's ups and downs with for over 5 years. I have also been very grateful for the company of Konrad Sierzputowski: the most permanent type of temporary subletter, Miriam Hauptman: a wonderful roommate, colleague, and friend, and Vishakha Shukla, who I have really appreciated getting to know over the past year. I have truly enjoyed our shared meals, binge-watching, and long and interesting conversations (extra-long with Miriam) and could not have hoped for better roommates during peak pandemic times. Last but not least, the honorary member of the Clifton gang: Esti Blanco-Elorrieta. Esti, like Shadow found a way into our house, you suddenly found your way into my heart for an unexpected friendship that feels like an everlasting conversation. You continuously make me reconsider what I think to know about life, love and morality, and I am extremely grateful for your stable presence, inspiring bravery and emotional support. Thank you for being you.

Ending with where I started, I want to thank all the people that where there with me for the first steps of this journey. Thank you to my Dutch linguists, Sofja Volkova, Isabella Jordanoska, Marianne de Heer Kloots, Astrid Gilein, Wytze Wolthuizen, Gerard Spaans, Marc Michels, Roger Lo, Gouming Martens, Suzanne van de Meer, Farida van Eer and Ruby Sleeman. Thank you so much for making me feel at home whenever I return to the motherlands, visiting me here and following my journey from a distance. Special thanks to Sofja, for sharing life, love, tears and laughter, and preparing me for dealing with all the scorpios in my life. While our work skypes are not always that productive, they are always invigorating and good to the soul. I also want to thank Kelsey Kolebrander and Sophie van Dorp, who go many years back with me all the way to Vlaardingen. Thank you for keeping your houses, cars and hearts open to me.

Then finally and most importantly, I want to thank my mother, Joke Kool, and sister, Tamara Tulling, from the bottom of my heart. They have provided me with unconditional love and support, and taught me the strength of determination and resilience. Dankjulliewel voor alles wat jullie hebben gedaan om mij een gelukkige en veilige jeugd te geven. Tamara, dankjewel voor al je ondersteuning en zorg, en Mama, dankjewel voor al je aanmoediging en kracht. Boven alles, dankjewel voor de vrijheid die je me hebt gegeven om te gaan waar het leven me naar toe waait.

ABSTRACT

This dissertation investigates the neural bases and development of displacement, which is a language property that allows us to communicate about situations outside the *here-and-now*. One way to displace from our immediate environment is to project ourselves into the here-and-now point of an alternative actuality. Another form of displacement (modal displacement) involves the postulation of possibilities compatible or incompatible with the actual situation. In this dissertation I report on three studies investigating the neural and developmental bases of modal displacement. The first study consisted of two experiments using magnetoencephalography (MEG) to investigate the neural mechanisms underlying factual and modal language comprehension. The combination of the results from these two experiments suggests that the brain is sensitive to the contrast between fact and possibility rapidly after it is presented, and that discourse situation updating only takes place for factual information. The second part of this dissertation investigated children's developing ability to process counterfactual language, looking at spontaneous production and comprehension. Specifically, I compared the acquisition of counterfactual conditionals with that of counterfactual wishes, as they differ in linguistic complexity. The results of these studies suggest that challenges involved with the form-to-meaning mapping of counterfactuality impact children's performance. Children start to produce the linguistically less complex wishes before counterfactual conditionals, and perform better on wishes in comprehension tasks. This dissertation ends with a discussion illustrating how the fields of cognitive neuroscience and first language acquisition can inform each other and help us build towards a broader understanding of the cognitive ability of human language displacement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

This dissertation is about using and understanding utterances that describe situations outside the *here-and-now,* such as hypothetical scenarios that are possible ("*maybe* it's a pig") or impossible ("*if only* pigs could fly"). Our ability to communicate any imaginable situation, even when completely dislocated from one's immediate surroundings is called displacement (Hockett, 1959). Displacement is one of the core properties of human language, that distinguishes our language from other animal communication systems (W. A. Roberts, 2006; Shi & Zhang, 2021; Suddendorf & Corballis, 2007; Tamura & Hashimoto, 2012). While some social insect species can communicate about things that are spatially dislocated (von Frisch, 1967; Wilson, 1962), only human language seems to allow for complete displacement in both space and time. For example, I could describe to you the building I lived in as a child, and you would be able to form a representation of this building despite the fact that you have never seen or visited it (Tamura & Hashimoto, 2012). Moreover, the situation I talk about can be completely made-up, e.g., I could describe to you how this building was located on Mars and that I grew up playing hide-and-seek with aliens, and you would still be able to represent this fictive event, keeping it separated from what you know to be real. The seemingly effortless human capacity of displacement allows us to effectively transmit knowledge and ideas beyond our individual experience, which has been essential in the formation of modern-day human societies. So, what underlies our language capacity of displacement?

In this dissertation, I address this fundamental question by conducting research in the cognitive neuroscience of language and first language development. Research in cognitive neuroscience provides us with insight into what the neural bases supporting language referring to situations outside the here-and-now are. Research in first language development helps us understand when this ability develops and what developmental stages there are. Combined insights of what the neural bases supporting displacement are and how this capacity develops over time will pave the way for future research on the development of these neural mechanisms and could eventually provide insight into the development of displacement from an evolutionary perspective.

Before I go into detail about the specific questions addressed in this dissertation, I will provide an overview below on the core topics discussed in this work and define relevant notions. An overview of the key concepts and definitions used throughout this dissertation can be found in Table 1.1. First, I introduce different ways one can displace from the here-and-now and discuss how we represent displaced situations. Then, I focus specifically on our ability to displace from any actuality (modal displacement). In particular, I discuss two types of language expressions that allow us to talk about possibilities compatible and incompatible with the actual situation: modal and counterfactual expressions. I will discuss the linguistic properties of such expressions and provide an overview of prior research investigating their neural processing and development. This introductory chapter will end with an overview of the questions asked in this dissertation and the studies conducted to address them. The remaining chapters of this dissertation correspond to separate experimental studies that stand on their own and can be read out of order if one desires to do so. I conclude this dissertation with a general discussion, tying together the results of the different studies, their implications and limitations, and suggestions about how to go forward.

Table 1.1 Key Concepts and Definitions as Used Throughout this Dissertation

Key Concepts	Definitions
Actuality	The current state of a real or imagined situation
Counterfactuality	Subcategory of modality used to discuss alternative ways the world could be or could have been, e.g., expressed through conditional: *if the monster were big, it wouldn't fit under the bed* or wish: *I wish I was braver…*
Discourse Updating	Updating an existing situation model when the situation's here-and-now changes, e.g., change in protagonist, goal, location, event or time
Factuality	Language category concerning what is known to be true or false in a situation
"Fake" Past	The mismatch between the tense morphology present in a counterfactual construction and its expressed temporal orientation, e.g., a counterfactual about the present contains past tense (bolded): *If I* ***had*** *money right now, I would buy a car*
Hypothetical Scenario	Situation that is temporarily stipulated to be true that may or may not conflict with what is accepted as true about the world
Modal Base	The grounds on which the likelihood of a possible or hypothetical scenario is determined, i.e., based on what you know (KNOWLEDGE-BASED/EPISTEMIC) or on what the circumstances are, e.g., rules and norms (RULE-BASED/DEONTIC)
Modal Displacement	An operation that shifts our perspective from the immediate present (*here-and-now*) to a possible or hypothetical scenario
Modal Force	The degree of certainty for a possible or hypothetical scenario x to be true, i.e., whether it is POSSIBLE (of all accessible possibilities there is at least one in which x is the case) or NECESSARY (x is the case for all accessible possibilities)
Modality	Language category used to discuss possibilities, e.g., *may*, *must*, *might*
Non-Actuality	Language category describing situations that are stated to be untrue or not known to be true in the actual world, includes negated expressions (*pigs do not have wings*), counterfactuals (*I wish pigs had wings*) and modals (*pigs may have wings*)
Presupposed Content	Information that is taken for granted within the discourse, e.g., '*since the monster is big'* presupposes prior knowledge of the existence of a big monster

Reality	The actual world that we all share and exist in
Simulation	Process of generating new representations of outcomes from a simplified model of the situation, used to make predictions and weigh probabilities
Situation Model	Mental representation of a situation, tracking events, actions and persons related to the here-and-now of that being discussed
Temporal Orientation	The time of the described event (present/past/future) relative to the utterance time
Theory of Mind (ToM)	The ability to reason about mental states and represent the belief state of others separate from our own

1. DISPLACEMENT FROM THE HERE-AND-NOW

In his effort to set human language apart from other animal communication systems, Hockett (1959) listed 'displacement' as one of 16 design features that characterize human language. While displacement originally was defined as "*being removed in space and time from what we are talking about*" (Hockett, 1959, p. 36), subsequent discussions arose about what counts as displacement and whether this is a property we share with some species of animals (Liszkowski et al., 2009; W. A. Roberts, 2006; Tamura & Hashimoto, 2012). While displacement is sometimes defined as being either spatial or temporal (Hockett, 1959; Riggs et al., 1998; W. A. Roberts, 2006), Tamura & Hashimoto (2012) distinguish between memorized and unexperienced displacement. While certain animal species display behavior suggesting displaced reference and understanding, this displacement is based on prior experience with the referent. For example, bees can communicate the location of a previously visited food source by performing a waggle dance (von Frisch, 1967), and ants can recruit other workers to retrieve food sources that are too heavy to carry alone (Wilson, 1962). In contrast, unexperienced displacement is unique to humans (Tamura &

Hashimoto, 2012). People can talk about referents that are not currently present, and listeners can form a representation of this referent even if they have never encountered it themselves. However, displacement is not limited to our actuality. When we combine displacement with creativity we can talk about any conceivable scenario and place ourselves in imaginary worlds and situations. On top of this, we can displace from the here-and-now of any actuality (real or imagined) by talking about possibilities and other hypothetical situations (e.g., using modal expressions such as *maybe*, *must* and *would*). Semanticists have coined the term "modal displacement[1]" to refer to this displacement from the here-and-now by talking about possibilities. Thus, there are different ways in which one can displace from the here-and-now, and while some types of lower-level dislocation appear to be shared with other animal species (such as spatial displacement), unexperienced and modal displacement seem to be unique to human language. In what follows, I focus on complex displacement: unexperienced displacement and modal displacement, discussing for each type separately theories on how our minds achieve this, and when this ability is thought to develop.

1.1. Unexperienced Displacement by Representing Situations

When listening or reading a story, we can get completely engrossed in its content. For this to happen, it does not matter whether the described content is real (e.g., a report about a fascinating event that just happened) or fictional (e.g., a story about a talking cat wearing boots), we are able to follow along without losing track of what we believe to be true about reality. How do we achieve this unexperienced displacement from the actual state of affairs? As it turns out, it is not as much

[1] As far as I can tell, this term can be traced back to von Fintel and Heim's lecture notes on intensional semantics (Kratzer, 2013).

about leaving behind the here-and-now as it is about mentally shifting ourselves into the here-and-now perspective of the described situation.

1.1.1. Situation Models in Discourse Representation

When comprehending or producing discourse, we build a mental representation of the described situation, a process which requires a dynamic interaction between linguistic, pragmatic and world knowledge. This mental representation is often referred to as the *situation model*, containing cognitive representations of events, actions and persons related to the situation discussed (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Crucially, these models represent the situation described by language, rather than the specific wording and propositions itself (Glenberg et al., 1987; van Dijk & Kintsch, 1983), and the situation representation is found to be modality-independent, emerging from written, auditory and visual story comprehension (Gernsbacher et al., 1990). Views differ on whether the nature of this representation is propositional (van Dijk & Kintsch, 1983) or experiential (grounded in perception and action) (Glenberg et al., 1987; Zwaan & Madden, 2004). When constructing the discourse model of a described situation, we displace ourselves from the current state of affairs by shifting into another perspective, place and time (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Relative to the situation's here-and-now perspective, we are found to track locations, events, actions, persons and goals of the situation being discussed (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Speakers are found to be particularly sensitive to the protagonist and the situational goal, tracking the situation's here-and-now even when there is incongruency between the physical and mental here-and-now spaces. For example, participants were found to be better at recalling objects from the room the protagonist was thinking about than from the room the protagonist was actually located in, suggesting a higher accessibility of the mental space compared to the physical one (Morrow et al., 1989).

Situation models are dynamic, and continuously updated as new information comes in. When new information becomes available during language processing, situation models are updated such that information relevant to the situation's here-and-now is foregrounded. As a consequence of situation model updating, 'old' information becomes less accessible as it is no longer relevant to the here-and-now of the story (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004). This reduced accessibility caused by discourse updating is found to be measurable in subtle manipulations and has been replicated in many different forms (Bailey & Zacks, 2015; Barnes et al., 2014, 2014; Burmester et al., 2014; Glenberg et al., 1987; Morrow et al., 1989; Pettijohn & Radvansky, 2016; Schoot et al., 2010). For example, Glenberg et al. (1987) showed that objects that were spatially associated with the protagonist were recognized faster than objects that were spatially dislocated. They used a probe-recognition task, in which they contrasted sentences (in context) where for example a main character either *put on* their sweatshirt (spatially associated) or *took off* their sweatshirt (spatially dislocated) before going jogging. They then provided them with a probe (*sweatshirt*) for which they had to respond whether it had occurred in the story or not. It took participants on average about 100-200 ms longer to respond to probes in the dislocated condition than in the dissociated condition. This was interpreted as an effect of foregrounding: as the protagonist is kept in the foreground, the associated object (*sweatshirt*) stays part of the situation model, while a dislocated object no longer is a part of the mental representation of this character.

To reiterate, we achieve unexperienced displacement from the actual state of affairs by mentally shifting ourselves into the here-and-now perspective of the described situation and modeling the properties related to this situation in a dynamic discourse representation. Discourse processing thus involves representing, maintaining and updating this mental representation as the discourse unfolds and new information becomes available. What are the neural mechanisms underlying these discourse processes during language processing?

1.1.2. Representing and Maintaining Discourse in the Brain

1.1.2.1. Right-lateralization of Discourse Processing?

One of the foundational findings about the neurobiology of language, is the left-lateralization or left-dominance of basic language ability (Broca, 1861; Pascual-Leone et al., 1991; Wernicke, 1874). Patients suffering from brain damage in left inferior frontal and/or superior temporal areas display various deficits with language production and comprehension, such as difficulty with grammatical structure, phonological processing and word retrieval (LaCroix et al., 2021; Ries et al., 2016; Turkeltaub, 2019). In contrast, brain damage in the right hemisphere is much less frequently associated with language impairments such as difficulties with basic speech production and perception (Gajardo-Vidal et al., 2018). While patients with a right hemisphere deficit generally pass simple linguistic tests, they have reported pragmatic deficits in interpreting prosody, humor and the point of complex discourse, especially during oral conversation (e.g., Brownell et al., 1986; Hough, 1990; Johns et al., 2008; Jung-Beeman, 2005; Martin & McDonald, 2003; Wapner et al., 1981). This has led researchers to propose that individuals with damage in the right hemisphere are unable to combine information across sentences (Brownell et al., 1986), suggesting that the right hemisphere plays a special role in discourse processing. However, the findings from studies investigating right hemisphere brain damage often conflict and it is not clear whether the reported impairments are due to a linguistic deficit (e.g., establishing global coherence or managing inferred or implied information), an impaired social skill of representing others mental states (Theory of Mind) or impaired executive functioning (such as working memory or suppressing irrelevant information) (Johns et al., 2008; Minga et al., 2021).

The idea that the right hemisphere plays a special role in discourse processing is persistent, however, especially since a body of neuroimaging studies in healthy adults reached similar conclusions (e.g., Ferstl et al., 2005; Jung-Beeman & Chiarello, 1998; Male et al., 2021; Robertson et al., 2000; Xu et al., 2005). Brain areas engaged during language processing are often activated bilaterally, that is we observe brain activation both on the left side of the brain and in its right-side homolog (anatomical equivalent) (Jung-Beeman, 2005). While most of the time this activation appears to be left-dominant, e.g., for language production, lexical access and visual word recognition (Price, 2012; Pujol et al., 1999), several linguistic processes induce equal amounts of bilateral activation, e.g., early speech perception and conceptual representation (Hickok & Poeppel, 2000; Rice et al., 2015). Right-hemisphere activation has also been observed to increase as linguistic processes get increasingly complex (Jung-Beeman, 2005), which has been suggested to reflect establishing coherence and inference at the propositional level (Xu et al., 2005). Increased activation in the right frontal lobe has been observed when comprehending connected discourse compared to reading unrelated sentences (Caplan & Dapretto, 2001; Robertson et al., 2000), and the right anterior temporal lobe has been found to be more engaged when encountering words that are inconsistent with the prior context (Ferstl et al., 2005). Xu et al., (2005) observed a systematic increase in engagement of the right hemisphere as contextual complexity increased, becoming maximal at the level of discourse when narrative details had to be synthesized into a whole. These results have led to the claim that the right-hemisphere plays an important role in discourse comprehension and integrates perceptual information and meaning into a situation model (Male et al., 2021). However, meta-analyses investigating studies on text comprehension and non-literal language use (e.g., irony or indirect requests) provided no evidence for a special role of the right hemisphere during discourse processing, suggesting that prior results finding special right-hemisphere engagement might have been task-induced (Ferstl et al., 2008; Hauptman et al., 2022).

Adding to this, Jacoby and Fedorenko (Jacoby & Fedorenko, 2020) suggest that brain responses to discourse are dependent on the type of discourse that is being processed. While prior studies detected right temporal parietal engagement during the comprehension of narratives (stories built around characters) (e.g., Xu et al., 2005; Yarkoni et al., 2008), expository texts (constituting facts about the real world) elicited no effect of discourse coherency in this region (Jacoby & Fedorenko, 2020; Lin et al., 2018). This is compatible with findings that suggest that the right temporoparietal junction (rTPJ) is particularly involved in representing the mental state of others (using Theory of Mind) (Mano et al., 2009; Saxe & Powell, 2006; Saxe & Wexler, 2005; Vistoli et al., 2011). Thus, like the conclusion we drew from summarizing studies on patients with right hemisphere impairments, studies on healthy adults also provide mixed results on right hemisphere engagement during discourse processing. While right hemisphere engagement is observed during discourse processing, this activation could also be due to increased processing cost or engagement in social processes such as representing other people's minds.

1.1.2.2. Discourse Processing and the Default Mode Network

A different line of research moves away from discussions about right lateralization of discourse processing, and instead relates various discourse processes to intrinsic activity of the default mode network (Raichle, 2015; Yeshurun et al., 2021). The default mode network (DMN) reflects activity patterns that are decreased in attention-demanding tasks, but active during relaxed non-task states, e.g., when performing self-referential processing and mind-wandering. The default mode network includes the medial prefrontal cortex (mPFC), the posterior cingulate cortex (PCC), medial temporal gyrus (MTG), the adjacent precuneus and the temporoparietal junction (rTPJ) / inferior parietal lobule (IPL) (Raichle, 2015; Yeshurun et al., 2021), see Figure 1.1.

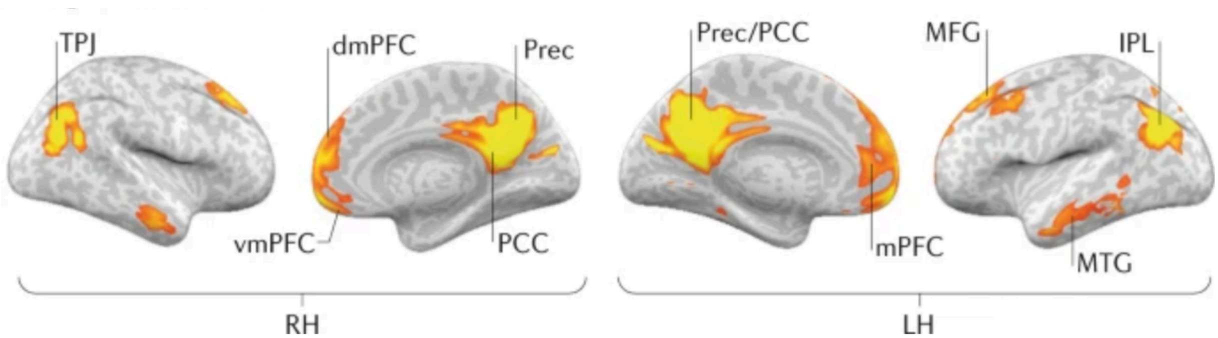

Figure 1.1. Brain regions comprising the Default Mode Network (DMN), including the posterior cingulate cortex (PCC), precuneus (Prec), ventromedial prefrontal cortex (vmPFC) and dorsomedial prefrontal cortex (dmPFC), middle frontal gyrus (MFG), middle temporal gyrus (MTG) and the bilateral temporoparietal junction (TPJ)/ inferior parietal lobule (IPL). LH = left hemisphere; RH = right hemisphere. Taken from (Yeshurun et al., 2021), Adapted by permission from Copyright Clearance Center: Springer Nature, Nature Reviews Neuroscience, *The default mode network: where the idiosyncratic self meets the shared social world,* Yeshurun et al. (2021).

While the default mode network initially has been linked to a variety of internally oriented processes such as daydreaming, and thinking about the past or future (Buckner et al., 2008; Gusnard et al., 2001), it has also been found to be activated during semantic processing (Binder et al., 1999, 2009). While perceptual tasks caused deactivation in the default mode network, a matched semantic decision task did not deactivate these brain regions, suggesting that resting state activity reflects access to conceptual knowledge (Binder et al., 1999). This is supported by the fact that the default mode network is also found to be engaged when listening to narratives. More specifically, it has been proposed that the DMN might be involved in building and maintaining situation models (Ferstl et al., 2008; Morales et al., 2022; Yeshurun et al., 2021). This is supported by studies that find increased DMN activation when contrasting coherent with incoherent speech (Ferstl et al., 2008; Yarkoni et al., 2008). In fact, a recent study shows that left lateralized default mode network (DMN) is similarly active during both discourse comprehension and discourse production, and that this activity is negatively correlated with the level of discourse coherency (Morales et al., 2022). The DMN was more activated for less coherent discourse, which is suggested to be the result of an increased working load to construct situation models, as less coherent discourse requires more frequent updating and reconfiguration. This finding is in contrast with prior literature showing that incoherent discourse (scrambled utterances) engages the default mode network less than coherent discourse (Ferstl et al., 2008; Yarkoni et al., 2008; Yeshurun et al., 2021). Morales et al. (2022) argue this is likely due to the unnaturalness of the incoherent stimuli where participants may abandon any attempt to build a discourse model.

Additional support for the idea that the DMN is involved in maintaining situation models comes from the fact that DMN activation during narrative comprehension is consistent among different modalities (e.g., text, speech, audiovisual movies), different languages, and different levels of abstraction, but dependent on the interpretation of the situation (Dehghani et al., 2017; Nguyen et al., 2019). Activation patterns in the default mode network align more when participants share a similar interpretation of a narrative (Nguyen et al., 2019; Yeshurun et al., 2017). For example, Yeshurun et al. (2017) found that participants' prior beliefs changed the way the DMN seemed to process the same story. When two groups of participants were presented with the same short story where a husband calls his friend at night asking for the whereabouts of his wife, participants who were previously told that the wife was unfaithful before listening to this story responded differently to the narrative than participants who were told that the husband was paranoid and jealous. This was reflected in DMN activity during the story, which showed a stronger alignment of neural responses within groups than between the two groups, even though the exact content of what the participants heard was the same. The default mode network activity thus depends on the content rather than the form of external input.

While the studies described above were focused on language processing, alternative descriptions of the DMN's function are compatible with the overall idea that the default mode network is involved with experienced displacement. Buckner & Carroll (2007) suggested the overarching function of the DMN is "self-projection", *"the ability to shift perspective from the immediate present to alternative perspectives as self-projection"* (p. 49). Hassabis & Maguire (2007) argue this perspective shifting does not have to relate to the self, and instead argue for the broader term "scene construction" to describe the processes of "*mentally generating and maintaining a complex and coherent scene or event*" (p.299). This function would include various cognitive processes that have been shown to engage the default mode network: memory recall, future thinking, navigation, imagination, viewer replay, vivid dreaming and Theory of Mind.

All in all, recent research makes a compelling case that the default mode network is involved with establishing a model of a described situation, incorporating external information with prior knowledge and beliefs, and allowing us to shift into an alternative here-and-now.

1.1.3. Development of Situation Representation and Discourse Updating

Children's initial spontaneous language productions revolve around the here-and-now (Harris, 2001). They refer to people and objects that are in their proximity and express desires and needs. Yet, the first building block of displaced reference, the ability to refer to or understand reference to a familiar absent object, seems to emerge early in life. At the end of their first year, around the time the average child starts producing their first words, infants can gesture for and understand reference to absent objects (Ganea, 2005; Liszkowski et al., 2009; Saylor & Baldwin, 2004). For example, 12-month-old children, but not chimpanzees, can request an absent-referent through pointing at its usual location (Liszkowski et al., 2009), and Ganea (2005) showed that 14-month-olds can understand references to a relatively novel absent object. Verbal reference to an absent new toy brings the toy to the child's mind as long as it is accessible and there is not too much of a delay between the displaced reference and last experience with the object. Saylor & Baldwin (2004) suggest that this early ability is still in development, however. While they found that one-year-old children responded differently to overheard talk about absent and present caregivers, suggesting there is some awareness about the identity of the absent referent, more sophisticated absent reference comprehension skills seem to emerge around age 2, as children started to elaborate verbally about the absent caregiver. Relatedly, Ganea et al. (2007) found that 22-month-olds, but not 19-month-olds, were able to update their mental representation of a character when they were verbally informed about the character's change-of-state happening out of sight (e.g., when the character got wet). However, Ganea et al. (2007) point out that the ability to update a representation could depend on a complex interaction of representational and contextual factors, such as strength of the memory, familiarity with the object and familiarity of the testing environment, and that younger children might be able to update representations under facilitated circumstances.

It is not completely clear what encompasses the initial representations of a situation for young children, and when these representations are first formed. Developmental research does show that children as young as 3 construct rich discourse representations where they take into account the here-and-now perspective of the described situation (Fecica & O'Neill, 2010; Rall & Harris, 2000; Ziegler & Acquah, 2013). Rall & Harris (2000) were the first to find that preschoolers are sensitive to incorrect use of directional movement verbs such *bring/take* and *come/go* if the direction of the movement does not match the perspective of the protagonist. For example, if children were first introduced to Little Red Riding Hood sitting in her bedroom, the utterance "when her mother *went* in" describes a movement inconsistent with the perspective of the protagonist (as from the girl's perspective, her mother *comes* in). Like adults in such paradigms (Black et al., 2018), children had the systematic tendency to misrecall inconsistent verbs, showing they take into consideration the perspective of the protagonist. Similar to adults (Morrow et al., 1989), 4- and 5-year-olds, but not 3-year-olds, also tracked the location a character was thinking about over its physical location (O'Neill & Shultis, 2007). Subsequent research confirms that children indeed seem to project into the space created by the narrative and construct the scene and events from the protagonists' point of view (Ziegler & Acquah, 2013), even though it is not clear at which point in development this becomes fully mature.

All in all, the studies discussed show that the basic building blocks for displacement are in place from a very early age (at the onset of language production), and that at least by age 3 children's discourse representations are sophisticated enough to incorporate rich information about the unfolding discourse situation. In the next section, we turn to a different type of displacement: modal displacement, which allows us to postulate and compare different possibilities.

2. MODAL DISPLACEMENT BY POSTULATING POSSIBILITIES

Human language can express departures from reality, or rather any actuality, by communicating possibilities. In contrast to experienced displacement, where we shift into an alternative here-and-now point, modal displacement allows us to leave the here-and-now point of any actuality (the one we live in, or the one we placed ourselves in) and postulate possibilities that are compatible or incompatible with that actuality. Specialized linguistic expressions called *modal expressions* (such as *maybe*, *probably*, *can* and *must*) allow us to directly encode possibilities that are compatible with the actual situation and indicate how certain we are about this possibility. To talk about possibilities that are incompatible with the actual world, we use a subcategory of modal expressions called *counterfactuals* (such as "If only I was taller", "I wish pigs could fly" and "Had I known this before"). While modal and counterfactual utterances are tightly related and heavily researched as separate topics, they are rarely thoroughly discussed in conjunction, perhaps because of the vastness of the research landscape they both cover. Since modality and counterfactuality arise from the human mind, and what we believe to be possible or impossible, research on these topics is highly interdisciplinary, covering ground in the disciplines of linguistics, philosophy, psychology and cognitive neuroscience. By no means will the following description be a comprehensive overview of all the research and insights gained on modal and counterfactual language and thought, nor will it do justice to the fine details and complexities involved in thinking deeply about these topics. For thorough discussions on modality and counterfactuality, I would like to refer the reader to other detailed overviews on the current state of the linguistic treatment of modals (Kratzer, 2012) and counterfactuals (Arregui, 2020; Iatridou, 2000). For a comprehensive overview of the cognitive processes underlying modal and counterfactual thinking one should consult (Byrne, 2016; Johnson-Laird & Ragni, 2019; Kulakova & Nieuwland, 2016b; Leahy & Carey, 2019; Van Hoeck et al., 2015).

Instead, the goal of this section is to provide a synthesis of different insights and ideas about modality and counterfactuality that have developed from linguistics, psychology and cognitive neuroscience to lay the conceptual groundwork for our discussion on the developmental and neural bases of modal displacement. First, I will lay out some of the core linguistic properties of modals and counterfactuals, then I will discuss theories about how we mentally represent modality and counterfactuality, and end with an overview about insights and open questions pertaining to the neural bases and development of these language categories.

2.1. Linguistic Properties of Modal and Counterfactual Expressions

2.1.1. Modality: Expressing Open Possibilities

Linguistic modality allows us to talk about non-actual[2] situations by postulating open possibilities, commonly described in terms of possible worlds (e.g., Hacquard, 2006; Kratzer, 1981, 2012; Lewis, 1973; Portner, 2009). Possible worlds represent different variants of all the ways the world

[2] Models can be used for some restricted cases of actuality too (actuality entailment) (Bhatt, 1999; Hacquard, 2020), but our focus here is solely on non-actual usages.

could be. A widely accepted approach to the semantic analysis of modal expressions, is to analyze modals as quantifiers scoping over possible worlds, where the quantifier determines the force of the modal (Kratzer, 1977, 2012). Modals come with different forces: necessity (e.g., *must*) and possibility (e.g., *may*). Necessity verbs like *must* or *have to* are associated with a universal quantifier 'for every x' (∀) while possibility verbs like *may* or *can* are associated with an existential quantifier 'for at least some x' (∃). A sentence like "It must be raining" indicates that in all accessible possible worlds, it is raining, while the utterance "It may be raining" indicates that only in a subset of the accessible possible worlds it is raining. Whether a world is accessible or not depends on the conversational background. In the examples above, the conversational background - determining the 'flavor' of the modal - is *epistemic*: it picks out worlds compatible with what is known in the world of evaluation. But modals like *may* and *must* can express multiple modal flavors. If you are the god of rain and have the power to make it rain whenever you please, "it must rain" and "it may rain" take on a different meaning. The modal flavor of *may* and *must* in these utterances are *deontic*, which means they pick out worlds compatible with the rules. Now, "it must rain" indicates that in all of the possible worlds in which the rain god's orders are obeyed, it rains. In contrast, "it may rain" indicates there is at least one possible world in which it is allowed to rain. There are various other modal flavors, such as *bouletic* (picking worlds that are compatible with desires) or *teleological* (picking worlds that are compatible with goals), but most commonly those and other flavors are grouped together with *deontic* modals into one overarching 'root' category (see Hacquard, 2006). This root modal flavor category is contrasted against the *epistemic* (knowledge-based) flavor. Not all possible worlds are as compatible with the actual world as others, instead they are ordered based on restrictions posed by the modal flavor (e.g., the desirability of an outcome or the likelihood of something to happen).

Modals can be expressed through different syntactic categories, both lexical and functional (Kratzer, 1981; Palmer, 2001). In English, the category of functional modals include modal auxiliaries like *can* or *must,* while lexical modals include adverbs like *maybe* and *probably.* While modal auxiliaries are often polysemous (they can take on different modal flavors like *may* and *must*), lexical modals tend to have dedicated meanings that only express one flavor (Cournane, 2021; Hacquard, 2013; Traugott, 2006). It is often thought, that the fact that modal auxiliaries can take on different modal flavors can be derived from their structural position (e.g., Hacquard, 2006). In particular, epistemic modal verbs scope higher than root modal verbs, scoping over tense and aspect, while root modal verbs are thought to be interpreted below tense and aspect.

2.1.2. Counterfactuality: Expressing Alternative Actualities

Being a subset of modality, established theories on the semantics of counterfactual utterances rely on the postulation of possible worlds (Iatridou, 2000; Kratzer, 2012; Lewis, 1973). Counterfactual conditionals, such as "If kangaroos had no tails, they would topple over" consist of two parts: the *if*-clause is called the 'antecedent', while the *then*-clause is called the consequent. Counterfactual antecedent clauses, like "If kangaroos had no tails", have a dual meaning attributed to them. They postulate the possibility of kangaroos not having tails, while at the same time asserting that this possibility is counter-to-fact or conflicting with what we know to be true about the actual world. The consequent describes what the world would be like given that the antecedent is true, e.g., kangaroos would topple over. As with open modal utterances, the domain of possible worlds accessed by counterfactuals is restricted, not all possibilities are considered. Instead, counterfactuals are thought to require consideration of the possible worlds that are most similar to the actual world (Arregui, 2009; Ippolito, 2003; Lewis, 1973; Stalnaker, 1975). For example, "they would topple over" is considered a likely consequence of kangaroos not having tails, while consequents that are too dissimilar from the actual world (e.g., they would carry chairs for resting) is consider a low probability continuation. Another feature of counterfactuals is the presence of the so-called "fake" past tense (Iatridou, 2000). The past tense in counterfactual constructions is called "fake" because there is a mismatch between the counterfactual's morphological tense marking and temporal orientation. Take for example, the utterance "If I **had** money right now, I would buy ice cream". In this example, the expressed morphological tense is past, while the temporal orientation of the utterance is present. Counterfactuals about the past contain two layers of pastness: a real and a "fake" one, resulting in a past perfect construction: "If I **had** bought icecream yesterday, I could have eaten it now". Most theoretical accounts put a lot of explanatory weight on the counterfactual's "fake" past, arguing it plays a direct role in obtaining counterfactual meaning. There are two main approaches to analyzing the semantic role of the counterfactual's past tense morpheme (Bjorkman, 2015; Karawani, 2014; Romero, 2014; Schulz, 2014; von Prince, 2017). Past-as-past (or 'back-shifting') approaches argue that the counterfactual's past tense morpheme fulfills the function of shifting back in time, thus maintaining a true past tense component (Arregui, 2005; Dudman, 1983; Ippolito, 2006; Ippolito & Keyser, 2013; Ogihara, 2000; Romero, 2014). In contrast, past-as-modal ('remoteness-based') approaches argue the counterfactual's past is "fake" in the sense that the morpheme does not make any temporal reference (Bjorkman & Halpert, 2017; Iatridou, 2000; Karawani, 2014; Karawani & Zeijlstra, 2013; Ritter & Wiltschko, 2014; Schulz, 2014). For example, Iatridou (2000) argues that the past tense morpheme is the realization of an 'exclusion' feature, that either scopes over time (excluding the present, resulting in a past tense reading) or over worlds (excluding worlds, resulting in a counterfactual reading).

Most attention in the linguistic literature goes to counterfactual conditionals, but counterfactuality can also be expressed through other constructions. Other constructions in English that express counterfactual meaning are counterfactual wishes (e.g., "I wish I had ice cream") and inversion ("Had I done this sooner, I wouldn't have to rush right now") (Iatridou, 2000; Ritter & Wiltschko, 2014).

2.2. Mental Representation of Possibilities

Our ability to comprehend modal and counterfactual utterances is made possible by modal cognition, our general ability to reason about all types of possibilities (Phillips et al., 2019). As discussed in the previous section, linguistic theories of modality and counterfactuality often involve the postulation of multiple possible worlds (von Fintel, 2006; Iatridou, 2000; Kratzer, 1977, 2012). But how do we mentally represent such possibilities? Some accounts follow the overall insights obtained from Kratzer (2012) and argue that modal cognition involves partitioning the set of relevant possible worlds and ordering them based on probability and moral value (Phillips et al., 2019; Phillips & Knobe, 2018). However, it has been argued that possible world semantics provides no plausible psychological account for modal processing, as the amount of possible worlds considered for any modal statement seems simply too big (Johnson-Laird & Ragni, 2019). Instead, most psychological models of modal and counterfactual thought revolve around the concept of a 'mental model' (Johnson-Laird, 1994).

2.2.1. Mental Model Theory: Possibilities as Finite Mental Models

Mental models are small finite representations that are iconic in the sense that their structure matches the structure of the represented situation as much as possible. Mental models can be construed from perception, imagination, or discourse, and can be either visual or abstract, representing situations that cannot be visualized. Different mental models can be constructed simultaneously to represent different possibilities, but people try to limit this as much as possible (Johnson-Laird & Byrne, 2002). Mental Model Theory differentiates between the mental representation of basic conditionals, which indicate open possibilities (similar to modals like *may*), and counterfactual conditionals. While a basic conditional like "if she played a game then she didn't play music" should strictly give rise to 3 possible mental models (game/no music, no game/music and no game/no music), in practice people would only represent the possibility for which both clauses hold (game/no music). In order to save working memory space, they leave the other possibilities implicit adding a mental 'footnote' to indicate some possibilities are left out (Byrne & Johnson-Laird, 2009; Johnson-Laird & Byrne, 2002). In contrast, counterfactual conditionals like "if she had played a game, she wouldn't have played music" would give rise to two explicit mental models, one that is factual (no game/music) and one that is counterfactual (game/no music). To distinguish between different types of discourse (factual, counterfactual, hypothetical etc.) the theory assumes that symbolic operators attach to the models to mark their status (Johnson-Laird & Ragni, 2019).

Evidence for the representation of dual meaning for counterfactual utterances comes from psycholinguistic studies investigating offline measures of sentence comprehension. After participants read counterfactuals, the dual meaning of counterfactual utterances is accessible

(Fillenbaum, 1974; Thompson & Byrne, 2002). However, it is not clear whether this dual meaning is represented simultaneously or whether factual and hypothetical content is postulated in succession during online processing (Kulakova & Nieuwland, 2016b).

2.2.2. Experiential Simulation Theory: Non-Actuality as an Auxiliary Representation

Two-step processing of dual meaning has been put forward by research focusing on sentential negation. It has been argued that the processing of non-actual information (negation here) involves mental (experiential) simulations of the base proposition. For example, the representation of a sentence with negation (such as "there is no bird in the sky") would involve a first step of simulating a bird being in the sky (Kaup et al., 2006, 2007). This non-actual mental simulation is thought to be created in an auxiliary representational system, which is contrasted against the representation of the actual situation that is created after simulating the non-actual content. Kaup et al. (2006) found that 750 ms after reading a negative sentence, only the non-actual information was available to participants, while at 1500 ms participants were only focused on the actual (negated) representation. The juxtaposition of the actual and non-actual world, which prevents information in the auxiliary system from being integrated into the representation of the described world, gives rise to the mental rejection of the non-actual content. The idea that mental simulations of non-actual information is represented separately, in addition to the regular situation model tracking the here-and-now of a situation, predicts that non-actual information should not be involved in updating of the factual situation model.

This two-step view is compatible with prior work on the integration of counterfactual information (De Vega et al., 2007; 2012). De Vega et al. (2007) modified the probe-recognition task (discussed in section 1.1.1.) to compare the accessibility of events after a factual or counterfactual update. For example, they probed either *TYPE* or *DRINK* after reading the following story: *John was still in the office sitting in front of the computer. He started to type a report that his boss had asked him for*. FACTUAL CONTINUATION: *As he had enough time, he went to the café to drink a beer*. COUNTERFACTUAL CONTINUATION: *If he had had enough time, he would have gone to the café to drink a beer*. De Vega et al. (2007) found that the probe with old information (*TYPE*) was recognized slower in the factual condition than in the counterfactual condition, while the new information (*DRINK*) was equally accessible. They interpreted the longer reaction times in the factual condition to indicate discourse updating, as the action of *typing* had been interrupted and therefore was no longer relevant to the here-and-now, resulting in decreased accessibility of the concept. In contrast, they concluded that the non-actual meaning of counterfactuals did not contribute to the build-up of the discourse representation. In the counterfactual story, the *typing* activity was never interrupted, as the characters stayed where they were in the here-and-now (they did not go drinking). For this reason, discourse updating of the here-and-now action never took place, and no decreased accessibility of the concept *typing* was observed.

The fact that there was no difference in recognition of the new information probe suggests that after reading a counterfactual, a non-actual interpretation of the events was activated and momentarily coexisted with the factual interpretation. The duration of this dual activation is estimated to be somewhere between 500-1500ms, after which the non-actual meaning was found to become inaccessible (de Vega & Urrutia, 2012). In an ERP study, this non-updating of counterfactual discourse integration compared to factual integration was replicated. Urrutia, de Vega, et al. (2012) observed increased negativity for the factual but not for the counterfactual condition starting from 100/200ms after the critical word.

2.2.3. Model Comparison

As an attentive reader may have noticed, there is a lot of overlap between the ideas of situation models (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998) and mental models (Johnson-Laird, 1994). Both are defined as being amodal representations of discourse that can be visual or propositional. In fact, often people will gloss over the difference between the two model types and refer to them in juxtaposition (e.g., Zwaan, 2016, p. 1028). However, there are some differences in the way these representation theories are applied. Situation models are often used by proponents of an experiential view on mental simulation, where language comprehension involves constructing an experiential (perception plus action) simulation of the described situation (Zwaan, 2004). In contrast, mental model theory is often thought of in propositional terms and makes use of symbolic operators (like negation) to mark different types of information (Johnson-Laird & Ragni, 2019). The use of symbolic operators allows for different treatments of negation.

As discussed above, negation in terms of the experiential simulation view involves two representations: the actual representation (in the situation model) and the non-actual representation in an auxiliary representation. For mental model theory, negation can be expressed in one mental model including a negative operator (Johnson-Laird, 1994). For counterfactuals, both theories predict the postulation of two separate models (Byrne & Johnson-Laird, 2009; de Vega & Urrutia, 2011). Although it seems that in mental model theory these models are equal (two mental models), while in the experiential simulation view the non-actual model is auxiliary to the main situation model about the here-and-now. While mental model theory accounts for open (modal) possibilities (representing the stated possibility with a mental 'footnote' to indicate some possibilities that are left implicit), it is not clear how the auxiliary system of situation models would handle possible information. The non-actual meaning of negated and counterfactual utterances is assumed to be temporary and fleeting, never incorporated into the here-and-now model. This makes sense as in both cases there is also actual information that can be incorporated into this model. However, with modal utterances the status of actuality is unknown. It is thus a question whether this information would stay separate from the situation model, and how factual and modal information would be integrated.

All in all, this section has discussed three different approaches to the representation of possibilities: models that simultaneously represent a multitude of alternatives (Phillips et al., 2019; Phillips & Knobe, 2018), mental model theory (Byrne & Johnson-Laird, 2009; Johnson-Laird, 1994; Johnson-Laird & Ragni, 2019), and the experiential approach (de Vega & Urrutia, 2011; Kaup et al., 2007). While these models differ on their assumptions and structure, they all differentiate between different possibilities by postulating separate representations or some type of distinctive marking that allows us to differentiate between what is actual and non-actual. We would therefore expect that comprehending modal or counterfactual utterances involves increased processing costs and/or additional neural mechanisms compared to factual information.

2.3. Possibilities in the Brain

While counterfactual processing in the brain is a hot research topic, there are only a few studies targeting the neural bases of modal processing specifically. However, as counterfactual reasoning includes the process of postulating a non-actual possibility, we expect some shared neural bases between modal and counterfactual processing. In particular, the shared computation of interest is that of ‘modal displacement’. In what follows we will first consider the neural underpinnings of counterfactual reasoning, and then discuss some studies looking at modal processing.

2.3.1. Counterfactual Processing and the Default Mode Network

It has been found that both thinking about the future and remembering past events (in contrast to thinking about factual knowledge[3]) recruits brain regions that overlap with the default mode network (DMN) (Buckner & Carroll, 2007; Spreng et al., 2009). Since counterfactual thinking involves both remembering elements about the past[4] and recombining those elements into a new scenario, it should not be surprising that the default mode network also plays a role in counterfactual reasoning (Van Hoeck et al., 2013). In counterfactual reasoning tasks, participants are cued to reason about how a short scenario or memory could have gone differently. Brain activity associated with counterfactual reasoning is commonly observed in the hippocampal area (involved in forming and recollecting memories), and in regions of the default mode network: the medial frontal and temporal lobes, the posterior cingulate cortex, precuneus, and the lateral parietal and temporal lobes (De Brigard et al., 2013; Kulakova et al., 2013; Nieuwland, 2012; Urrutia, Gennari, et al., 2012; Van Hoeck et al., 2015). In addition to these DMN regions, counterfactual thinking is also found to be correlated with increased activity in brain regions that are known to be involved in conflict detection, adaptive control and causal reasoning: the posterior medial frontal cortex (pmFC) and prefrontal cortex (PFC) (Van Hoeck et al., 2013). The engagement of the

[3] In most studies that report such results, thinking about the future or past is contrasted against a ‘semantic memory’ task that involves retrieving knowledge about facts (e.g., participants are asked to reflect on how to do routine operations, or describe properties of an object or word) (Spreng et al., 2009).
[4] In the case of past counterfactuals. Most neuroimaging research takes past counterfactual conditionals to be the default example of counterfactual thinking.

prefrontal cortex is thought to be induced by the fact that, in order to reason counterfactually, you have to consider two opposing (actual and non-actual) events at the same time. A special role for the prefrontal cortex in counterfactual reasoning is further supported by the observation that patients with brain injuries and other impairments in the prefrontal cortex show an absence of spontaneous counterfactual expressions (overview in Byrne, 2016, p. 143).

The default mode network does not only engage in active counterfactual reasoning but is also recruited when processing counterfactual sentences. Several studies have reported increased activation in DMN brain regions for counterfactual utterances compared to factual or hypothetical utterances (Kulakova et al., 2013; Nieuwland, 2012; Urrutia, Gennari, et al., 2012). This increased activation for counterfactual utterances is thought to follow from the increased processing costs involved in representing both the actual and non-actual described situation. Increased brain activity has been observed in left superior frontal regions, bilateral hippocampal gyri and the right inferior temporal gyrus when counterfactual conditionals (e.g., "If Pedro had decided to paint the room, he would have moved the sofa") were compared against factual utterances (e.g., "Since Pedro decided to paint the room, he is moving the sofa") (Urrutia, Gennari, et al., 2012). A similar increase in brain activity for counterfactual conditionals ("If the motor had been switched on yesterday, would it have burned fuel?") was found in the right lateralized cuneus and caudate nucleus when contrasted against past hypothetical conditionals (like "If the motor was switched on yesterday, did it burn fuel?") (Kulakova et al., 2013). Right-lateralization was also observed for the processing of inconsistent counterfactual sentences (compared to inconsistent factual sentences) (Nieuwland, 2012). While it has been suggested that this speaks to a specialized role for the right hemisphere during counterfactual processing (Kulakova & Nieuwland, 2016b), increased right hemisphere engagement could also be explained as the result of increased processing costs.

2.3.2. Modal Processing in the Brain

As discussed in the previous section, there has been quite some research on the neurobiology of counterfactual processing (focusing on counterfactual thoughts about the past). However, since counterfactual processing involves keeping in mind both non-actual information and conflicting actual information, it cannot isolate the cognitive ability of modal displacement from other processes such as conflict resolution or updating a situation with the actual information. To learn more about the neural mechanisms underlying modal displacement, we should thus consider the processing of modal utterances, for which the status of the actual world is open.

However, it is not at all clear what the neurobiology is of thinking about open possibilities (Byrne, 2016, p. 151). Dwivedi et al. (2006) found increased brain activity for modal utterances ("it might end quite abruptly") compared to factual utterances ("it ends quite abruptly") and speculated this activity increase reflects the cost of mentally representing and comparing multiple possibilities. However, this study was not controlled for utterance length or complexity, leaving uncertain whether the observed activity increases were really because of the experimental manipulation. Dwivedi et al. (2006) did find the brain to be sensitive to incongruent pronoun reference caused by the modality of utterances. A longstanding observation about pronoun reference is that you can only refer back to an entity introduced by a non-factual utterance, if you stay within the non-actual domain (Roberts, 1987). For example, while 'John might buy <u>a book</u>' can be continued with "<u>It</u> would be a murder mystery", this is not the case for the factual continuation "#<u>It</u>'s a murder mystery". This intuition has been confirmed through self-paced-reading experiments, showing that reading times for factual continuations are longer than for non-factual ones (Claus, 2008; Dwivedi, 1996). Similarly, a P600 ERP component[5] was observed when pronouns in factual utterances referred to a prior hypothetical antecedent, but not when factual pronouns referred to a factual antecedent or when the pronoun was presented in a hypothetical utterance as well (Dwivedi et al., 2006).

Together, these results suggest that elements of reality are available in non-actual situations, but that referents introduced by non-actual discourse cannot be accessed by the factual situation. This supports a view where the situation described by non-actual utterances are mentally simulated and kept separate from factual information (Claus, 2008), although it is far from clear how this is realized in the brain.

2.3.3. Open Question: Modal Displacement in the Brain

So far, our review of past literature on the neural correlates of experienced displacement (Section 1.1.2.) and modal displacement suggests that the default mode network (DMN) plays an important role in our ability to displace from the here-and-now. DMN brain regions become active when constructing stories, mentalizing other people's perspectives, thinking about the future, past and even when considering alternative ways a situation could have played out (Buckner & Carroll, 2007; Van Hoeck et al., 2015; Yeshurun et al., 2021). It has therefore been suggested that the default mode network forms rich context-dependent models of situations as they unfold over time (Yeshurun et al., 2021). It is tempting to hypothesize that the processing of modal utterances,

---

[5] A brain response observed with electroencephalography which indicates structural revision or repair (Burkhardt, 2007).

which postulate open possibilities, might also engage this default network. However, it is important to keep in mind a crucial difference between displacement through mentally shifting to an alternative here-and-now point[6] of a described actuality, and displacement through reasoning about non-actual possibilities compatible or incompatible with the actual situation (modal displacement). Actual information that updates the here-and-now situation seems to be incorporated into a mental model of the situation (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004), but there are reasons to believe that non-actual information is represented in an auxiliary system and not incorporated into the situation model (de Vega et al., 2007; de Vega & Urrutia, 2012; Kaup et al., 2007). If default mode network activity is involved in building and maintaining situation models, it is not obvious it would also be involved in representing an auxiliary system that represents non-actual information. Since counterfactuals have a dual actual and non-actual meaning (e.g., "If only Orpheus hadn't turned around…"), it is also not clear whether brain activity evoked by counterfactual processing reflects postulating a non-actual possibility (Orpheus not turning around) or updating a situation model with the acquired actual information (Orpheus turned around).

If we really want to isolate activity related to modal displacement, we should focus on the processing of modal utterances that indicate open possibilities. While modal and counterfactual utterances both involve the postulation of possibilities, they crucially differ in what they express about the actual world. Counterfactuals indicate that the expressed possibility is contra to the state

[6] During story comprehension you switch into the here-and-now of an alternative actuality, during past and future thinking you temporally displace through mental time travel to the here-and-now of a different point in time, when you mentalize someone's state of mind you consider the here-and-now form their perspective and past experiences.

of the actual world, but modal utterances express that the status of the actual world is open. For example, an utterance like “Maybe Eurydice isn’t following” leaves open whether Eurydice is or is following or not. How do our brains represent uncertain discourse information? Before we dive further into this question, the last section of this literature review will discuss how children learn to process modal and counterfactual utterances.

2.4. Developing Modality and Counterfactuality

To use and understand modal and counterfactual utterances in an adult-like way, children must acquire the ability to postulate possibilities and compare these against the actual world. While much research has looked into the development of modal and counterfactual reasoning and language processing, there are still many open issues about how and when these abilities start to develop and mature. Three factors that contribute to this lack of clarity in the literature are: 1) difficulty separating the influence of linguistic ability, cognitive ability and executive functions (such as inhibition or working memory) on children’s performance; 2) task complexity: difficult tasks might underestimate children’s abilities while easy tasks can be passed with a shortcut or strategy, therefore overestimating children’s abilities; and 3) disagreements about nomenclature and what exactly counts as “true” modal or counterfactual reasoning. In this section, I provide an overview of our current understanding of children’s modal and counterfactual possibility reasoning abilities, pointing out discrepancies and disagreements in the literature when they arise.

2.4.1. The Acquisition of Modal Language and Possibility Reasoning

2.4.1.1. Children's Production of Modal Utterances

Children spontaneously start to produce modal utterances in appropriate non-actual contexts as young as age 2 (Cournane, 2015, 2021; Papafragou, 1998). A well-studied phenomenon regarding children's early modal utterances is that children do not seem to produce different flavors of modal meaning (the modal base) simultaneously. The 'epistemic gap' refers to a period (approximately between age 2-3.5 years) in which children use modal auxiliaries with root meanings (e.g., deontic or ability *can*/*has*/*going to*) but not yet with epistemic meanings (e.g., *might*) (e.g., Shatz & Wilcox, 1991; Papafragou, 1998; Cournane, 2015). This asymmetry used to be explained through the different conceptual demands posed by root and epistemic modals. Epistemic, but not root modals would require a developed Theory of Mind (the ability to attribute to oneself and others mental representations and to reason inferentially about them), which is usually thought to develop between age 3 and 4 (P. Bloom, 2000; Papafragou, 1998, 2002). However, recent work rejects this hypothesis as an explanation for the epistemic gap period. Cournane (Cournane, 2015, 2021) argues that prior work showing an epistemic gap has been biased to report data from one linguistic modal category: auxiliary verbs. This is problematic, as in English (and in most other Indo-European languages) modal auxiliaries are polysemous between root and epistemic meaning and have different underlying structures (as discussed in Section 2.1.1). The polysemous meanings of modal verbs are also not evenly distributed, as most verbs (with the exception of *might* and *must*) have a strong bias towards a root meaning (van Dooren et al., 2017, 2019). For this reason, findings based on the acquisition of modal auxiliaries conflate conceptual, grammatical, and input factors. To untangle these different factors, Cournane (2021) compared the individual development of children's spontaneous productions of epistemic adverbs (e.g., *maybe* or *probably*) and epistemic modal auxiliaries. While children started using epistemic adverbs around age 2 (see also O'Neill & Atance, 2000), epistemic uses of modal auxiliaries generally emerged around age 3. Crucially, children produced epistemic adverbs during their modal auxiliary epistemic gap period, refuting the hypothesis that young children do not have the conceptual capacity to produce epistemic modal expressions. Does this mean that 2-year-olds already have mature concepts of all modal flavors, and entertain multiple possibilities when using them? Production data alone cannot answer that question, as it is impossible to know how the child intended their utterances, even when their uses seem appropriate and adult-like in the context. Studies specifically investigating children's ability to represent possibilities, however, show mixed results.

2.4.1.2. Children's Representations of Possibility

Several studies claim infants and even certain animal species (e.g., great apes, monkeys, and birds) have a prelinguistic ability to represent possibilities (e.g., Cesana-Arlotti et al., 2018; Hill et al., 2011; Kovács et al., 2014; Pepperberg et al., 2013; Téglás et al., 2011). For example, when hiding a desirable (food) item under one of two cups, and revealing which cup is empty, infants and some animal species can infer that the desirable item is underneath the other cup. However, Leahy & Carey (2019) argue that tasks like these could be passed with a 'minimal representation' of possibility, and not by considering multiple possibilities in parallel (the modal representation). Supporting this idea is the fact that great apes and children up to age 3 fail to prepare for multiple alternative possibilities (Redshaw & Suddendorf, 2016). When chimpanzees or children are presented with an upside-down Y-shaped tube, and experimenters put in an item at the top, some

3-year-olds and almost all 4-year-olds will hold up one hand at each exit of the tube, to catch the item. This clearly indicates their knowledge that the item could come out either way. Children younger than 4 and chimpanzees do not make the same inference, they will only hold up their hand at one of the exits, guessing where the item will come out. Leahy & Carey (2019) suggest that tasks like these uncover the fact that preschoolers and animals construct a 'minimal representation' of possibility. When presented with uncertainty, they simulate one possibility and commit to that solution. More specifically, they argue that the challenge for young children lies in keeping in mind two alternative outcomes of the same situation. This is somewhat surprising, as children younger than 3 display the ability to entertain two representations, e.g., they are able to distinguish reality from someone else's mental state or a pretend world (e.g., Harris et al., 1993; Onishi & Baillargeon, 2005). However, the ability to represent (and compare) two conflicting possibilities is thought to rely on marking the uncertainty of these competing representations by some type of symbolic operator[7], which would only start to emerge around age 4 (Leahy & Carey, 2019).

2.4.1.3. Children's Comprehension of Modal Utterances

Research directly targeting children's possibility reasoning abilities and their understanding of modal utterances, also suggest that their understanding of modal concepts might lag behind their production abilities. In fact, some have suggested that children younger than 6 have difficulty holding one or more possibilities in mind (Acredolo & Horobin, 1987; Ozturk & Papafragou, 2015;

---

[7] While Leahy & Carey (2019) are not explicit about what type of mental representation of possibility they envision, their use of a 'symbolic operator' to mark modality is compatible with Mental Model Theory (Johnson-Laird & Ragni, 2019).

Robinson, Rowley, Beck, Carroll, & Apperly, 2006). While children show understanding of modal uncertainty when provided with a clear contrast of the modal alternatives (Moscati et al., 2017; Noveck, 2001; Ozturk & Papafragou, 2015), they seem to engage in so called "premature closure" where they commit to a possible but not necessary conclusion before decisive evidence is available (Beck, McColgan, Robinson, & Rowley, 2011; Ozturk & Papafragou, 2015; Robinson et al., 2006). For example, Ozturk and Papafragou (2015) found that 5-year-olds had trouble judging modal sentences about whether an animal could be in one of two boxes when neither of the boxes where opened, and thus the location of the animal was uncertain. Beck et al. (2011) found that children up to age 6 have difficulty with epistemic uncertainty when they have access to the possible outcome. They suggest that children's willingness to imagine a possible outcome causes them to overestimate their knowledge about uncertain events. Moscati et al. (2017) similarly show that 5-year-olds displayed premature closure in a truth-value judgment task. However, eye gaze data from the same children doing the task showed similar looking behavior as adults when processing modal expressions that include *might* or *must*, showing sensitivity to the force meaning difference between the two expressions. It thus seems that children's tendency to engage in premature closure is a consequence of the task demands that require active non-actual reasoning rather than a consequence from immature processing of modal expressions.

All in all, research on the acquisition of modal language and reasoning provides mixed results. Children start spontaneously producing modal utterances around age 2 or 3, suggesting knowledge about the uncertainty modals express. Yet, children up to age 6 struggle with tasks that test comprehension on modal utterances. While some form of possibility reasoning seems to be in place from a very young age, tasks may sometimes overestimate children's abilities. It has been suggested that children under age 4 have a simplified representation of possibilities, which involves the simulation of one option, rather than simultaneously considering multiple conflicting possibilities. Older children sometimes overestimate their knowledge about uncertain events and engage in so-called “premature closure”, although this behavior might be task induced.

2.4.2. The Acquisition of Counterfactual Language and Reasoning

2.4.2.1. Pretense, Possibilities and Counterfactuality – Where to Draw the Line?

Researchers have often remarked on the similarities between counterfactual reasoning and pretend play, as both require temporarily entertaining a situation that is known to be untrue. Children, known for their rich imagination, engage in pretend play from a very young age (Harris et al., 1993). Pretend play involves a variety of different behaviors, including playing with imaginary objects starting between 12 and 18 months (e.g., pretending a banana is a phone), communicating with imaginary social partners (e.g., imaginary friends) and building elaborate unreal scenarios about possible worlds starting around age 3 (Lillard, 2017). During pretend play, children adopt the name of their played character and comment on objects and events that are not present in their actual environment (Harris, 2001). Crucially, even young children (age 3) understand the difference between imagination and reality (Lillard, 2013; Woolley & Wellman, 1993). As with interpreting narratives, pretend play has been argued to rely on the capacity to construct a situation model, and imagine the perspective of an agent within the imagined situation (Harris, 2001, p.257). The existence and prevalence of pretend play has been framed as a puzzle for development: why would children spend so much time and energy engaging with unreal worlds if they have still so much to learn about the real world? While some argue there is no clear answer to this question (Lillard, 2013, 2017), others have suggested that pretend play may be an important precursor to imagining possible worlds (Francis & Gibson, 2021; Gopnik & Walker, 2013). Specifically, pretend play and counterfactual reasoning are thought to rely on the same cognitive abilities: disengaging with current reality, postulating and reasoning about an alternative reality, and keeping the alternative possibility separate from reality (Walker & Gopnik, 2013; Weisberg & Gopnik, 2013). Supporting this view, some studies have found a correlation between children's performance on reasoning tasks that involve pretending and tasks that involve counterfactual reasoning (Buchsbaum et al., 2012; Francis & Gibson, 2021). In fact, Walker and Gopnik (2013) argue that pretending is a form of counterfactual reasoning, and that pretend play provides early opportunities to learn and develop this skill.

The claim that pretend play is a form of counterfactual reasoning has received some pushback. Beck (2016) argues that pretend play and counterfactual reasoning are quantitively different in their relation with reality and the cognitive demands they make. Beck points out that there are two definitions of "counterfactual thinking" used in the literature. The first definition, typically used by social scientists, uses the term counterfactual thinking specifically for thoughts about "what might have been", which is defined as "*thoughts about alternatives to specific elements of the real world*" (real-world counterfactuals). The second definition (general counterfactuals), often used by philosophers, defines counterfactual thinking broadly to include thoughts about pretend, future and fictional worlds as well as the real-world counterfactuals. Beck (2016) claims that real-world counterfactuals are closely tied to reality while pretend play is decoupled from reality, and therefore does not make the same cognitive demands. Citing Lewis' similarity principle (Lewis, 1973), Beck argues that real-world counterfactuals share a decision point with the actual world where the two worlds diverged, and the counterfactual needs to be the nearest possible world to the actual one. Pretend worlds, on the other hand, do not need to be contrasted with the actual world like counterfactuals do. This argumentation is in line with the distinction I draw in this dissertation between experienced displacement (underlying pretending), which is the result of shifting into an alternative here-and-now, and modal displacement (underlying modal and counterfactual reasoning) which requires the postulation of possible worlds (in)compatible with the actual one. Throughout this dissertation, I therefore only refer to counterfactual thinking in the narrow sense.

2.4.2.2. Children's Comprehension of Counterfactual Utterances

One of the main motivations behind Beck's argumentation for a distinction between general-counterfactual thinking and real-counterfactual thinking, comes from developmental studies investigating children's comprehension of counterfactual conditionals. While 3- and 4-year-old children are great imaginers, they struggle with comprehending counterfactual utterances. For example, they perform better answering hypothetical future questions (e.g., "If I draw on this piece of paper, which box will it go into?") than counterfactual questions (e.g., "If I had not drawn on the piece of paper, which box would it be in?") even though the utterances are roughly comparable in their overall structure, length and the presence of non-actual content (Riggs et al., 1998; Robinson & Beck, 2000). When questions like these were asked in the context of a sorting game, where blank sheets of paper go into one box and papers that had drawings on them in another, children often provided realist responses to counterfactual utterances (e.g., pointing to the box with drawn on papers). This type of actual response has also been observed in other studies with 2-4 year-olds (e.g., Kuczaj & Daly, 1979; Reilly, 1982; Rouvoli et al., 2019). The "realist" bias observed in counterfactual reasoning tasks is much akin to children's tendency to engage in premature closure on modal reasoning tasks (Robinson et al., 2006), and thought to arise because children struggle to hold multiple possibilities in mind while considering a false possibility temporarily true (Beck, Riggs, et al., 2011; Byrne, 2007). Especially inhibitory control has been suggested to be at the core of difficulty in counterfactual processing, as repressing the actual world would require high levels of inhibition (Beck et al., 2009; Weisberg & Gopnik, 2016).

Like we saw for modal comprehension, task effects similarly seem to influence children's performance on counterfactual reasoning tasks. Some have argued that prior studies overestimated children's ability to reason counterfactually by providing confounded tasks. The standard tasks used to test counterfactual reasoning can be passed by using 'basic conditional reasoning', which relies on children's general knowledge of causal regularities, social norms and experience rather than counterfactual knowledge (Rafetseder et al., 2010, 2013). For example, in the paper sorting game described above, a child could also pass the task by remembering the rule that drawn-on pieces of paper go into the drawn box and simply link (not drawn → clean paper box) for the question "If I had not drawn on the piece of paper, which box would it be in?", rather than parsing or comprehending the structure of the counterfactual utterance. When controlling for basic conditional reasoning, Rafetseder et al. (2010;2013) found that children were not able to reason counterfactually until age 12. However, others have questioned these results (Grosu & Cournane, forthcoming; McCormack et al., 2018; Nyhout et al., 2017; Nyhout & Ganea, 2019) suggesting that children's performance was underestimated in Rafesteder et al.'s tasks, since the design allowed for unwarranted inferences that could lead children astray. Using an alternative task, avoiding such confounds and providing clear causal structures in the physical domain, even 4- and 5-year-olds displayed mature counterfactual reasoning again (Nyhout & Ganea, 2019). Children were correctly able to answer questions such as "If she had not put the green block on the box, would the light still have switched on?". Three-year-olds performed at chance level on this task, but the authors speculated that perhaps this was due to the grammar of counterfactual questions being too complex.

2.4.2.3. Children's Production of Counterfactual Utterances

Mirroring findings on children's comprehension (Riggs et al., 1998; Robinson & Beck, 2000), production studies have reported that future hypothetical conditionals (conditional constructions about a future possibility such as "If it rains tomorrow, we will play inside") are acquired before counterfactual conditionals (Bowerman, 1986; Reilly, 1982). Reilly (1982) found that most children produce hypothetical conditionals by age 3 and produce their first spontaneous counterfactual conditionals at age 4. Kuczaj & Daly (1979) similarly report that future hypothetical conditionals seem to be acquired before (past) counterfactual conditionals (at the end of age 3). The age at which children start producing counterfactual conditionals thus seems to align with when they are found to start understanding these constructions, around age 4. However, in a corpus study of three children, Bowerman (1986) notes some surprising instances of counterfactual (present) conditionals at age 2 (1a,b), and also notes that children already use counterfactual *wish* at this age as well (2).

(1) a. <Just having crossed a narrow street when a car goes by> (Bowerman, 1986, 43)
Christy (2;4): That car will/would hit me if I was in a street

b. <Child is tired during long wait in doctor's office> (Bowerman, 1986, 44)
Eve (2;11): If we (didn't?) have to wait for so long
we would have be gone a long time

(2) Christy (2;1): I *wish* Christy have a car. (Bowerman, 1986, 10)
I wish me have a airplane

While prior corpus studies mostly focused on the acquisition of past counterfactual conditionals, simpler counterfactual constructions such as the present counterfactual conditional (lacking the past perfect) or counterfactual *wish*-construction (lacking a causal dependency) might be available to children at an earlier age. This would be in correspondence with the finding about spontaneous modal productions, where the linguistically less complex modal adverbs were found to be acquired before modal auxiliaries (Cournane, 2021). Besides being linguistically more complex, the past counterfactual construction is also rare in spontaneous conversation (much less frequent than other conditionals), which may contribute to their relatively late acquisition (Crutchley, 2013).

All in all, research on the acquisition of counterfactual language and reasoning suggest age 4 is an important turning point in children's development. At this age they are generally found to start understanding and producing counterfactual conditionals. However, some authors have suggested that the grammatical complexity of the past counterfactual constructions used in comprehension experiments might prevent younger children from demonstrating their true counterfactual capacities. Similarly, there are a handful of instances found in child corpora where children use linguistically less complex counterfactual structures (such as the present conditional and *wish*-construction) at age 2. So perhaps, children acquire the ability to reason counterfactually earlier than is currently assumed, which would put its onset closer to the age children start being able to reason about pretend scenarios.

3. THIS DISSERTATION

3.1. Synthesis

This dissertation investigates the neural bases and development of displacement. As reviewed above, displacement is one of the linguistic properties that distinguishes human language from other animal communication systems, allowing us to communicate about situations outside the here-and-now (Hockett, 1959; Tamura & Hashimoto, 2012). We achieve this by temporarily shifting our perspective from the current reality to a representation of the described situation (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004). This seemingly effortless ability to go back and forth between reality and the imagined situation is already found in children as young as 3 years-old (Fecica & O'Neill, 2010; Rall & Harris, 2000), and recruits the default mode network in the brain (Ferstl et al., 2008; Morales et al., 2022; Yeshurun et al., 2021).

Additionally, we can also displace from any actuality by communicating possibilities compatible or incompatible with the actual situation (modal displacement). Using specialized linguistic constructions, such as modal (*maybe*, *must*, *can*) or counterfactual ("If I were...', 'I wish I was…") expressions, we can talk about all kinds of hypothetical scenarios (Iatridou, 2000; Kratzer, 1981, 2012; Lewis, 1973). While modal and counterfactual utterances both involve the postulation of possibilities, they crucially differ in what they express about the actual world. Modal utterances (e.g., "Maybe Eurydice isn't following") express that the status of the actual world is unknown (either Eurydice is there or not), while counterfactuals (like "If only Orpheus hadn't turned around…") indicate that the expressed possibility is contra to the state of the actual world (Orpheus did turn around). Modal displacement allows us talk about any imaginable possibility, facilitating our ability to share knowledge and novel ideas. However, it's unclear what the neural bases and developmental trajectory of modal displacement are.

From a neural perspective, there is a lot of research focusing on the psychology and neurobiology of counterfactual reasoning and language processing, while we know very little about the neural bases of modal processing. Research on counterfactuality shows that there is significant overlap between the brain regions involved with building and maintaining situation models and counterfactual processing, mostly within the default mode network (Kulakova et al., 2013; Nieuwland, 2012; Urrutia, Gennari, et al., 2012; Van Hoeck et al., 2015). It is therefore tempting to hypothesize that the processing of modal utterances also engages the default network. However, since counterfactuals (e.g., "If only Orpheus hadn't turned around…") have a dual actual and non-actual meaning, it is not clear whether brain activity evoked by counterfactual processing reflects postulating a non-actual possibility (Orpheus not turning around) or updating a situation model with the acquired actual information (Orpheus turned around). Research on the neural mechanisms underlying the processing of modality is therefore needed to truly isolate activity related to modal displacement.

Developmentally, we know quite a lot about the first language acquisition of modal expressions. Children spontaneously start producing simple modal utterances in appropriate non-actual contexts around age 2 or 3 (Cournane, 2015, 2021), and display a prelinguistic ability to represent possibilities (Cesana-Arlotti et al., 2018; Kovács et al., 2014; Téglás et al., 2011). Yet it has been suggested that children younger than 6 have difficulty holding one or more possibilities in mind (Acredolo & Horobin, 1987; Ozturk & Papafragou, 2015; Robinson et al., 2006). This discrepancy has been explained through assuming that young children initially represent modal utterances as a minimal representation that only picks out one possible simulation, assuming that it is actual. Around age 4, they would start to display the beginnings of an adult-like modal representation that includes the consideration of different possibilities (Leahy & Carey, 2019).

At first impression, this claim ties in well with findings about children's development of counterfactual understanding. Children are reported to start producing and comprehending counterfactual utterances around age 4 as well (Guajardo et al., 2009; Nyhout & Ganea, 2019; Reilly, 1982; Rouvoli et al., 2019). Perhaps, the ability to represent multiple possibilities (modal displacement) develops around then? However, some authors have suggested that the grammatical complexity of the past counterfactual constructions used in comprehension experiments might prevent younger children from demonstrating their true capacities (Nyhout & Ganea, 2019; Rouvoli et al., 2019). Past counterfactual conditionals contain the uncommon past perfect construction, a causal dependency between *if*- and *then*-clause, and is underrepresented in the linguistic input compared to other conditionals (Crutchley, 2013). On top of this, there are reports of children using linguistically less complex counterfactual structures such as the present counterfactual conditional and *wish*-construction already at age 2 (Bowerman, 1986). So perhaps, children acquire the ability to reason counterfactually earlier than currently assumed, begging the question how children perform on less complex counterfactual constructions.

3.2. Outline

In this dissertation, I take an interdisciplinary approach to non-actual language processing and combine insights from cognitive neuroscience and first language acquisition to investigate what underlies our ability for linguistic displacement. The broad research questions addressed are: 1)

What are the neural mechanisms underlying modal displacement, and 2) how does this ability mature across development? To answer these questions and to gain a deeper understanding of the mechanisms involved in representing discourse containing factual and non-actual information I employ neuroimaging, corpus research, and behavioral tasks with adults and children.

In the first part of this dissertation (Chapter 2), I pose the question of what the neural mechanisms underlying our ability of modal displacement and discourse updating are. In particular, I investigate the differences between factual and modal language comprehension using magnetoencephalography (MEG), a non-invasive measurement of brain activity to localize brain activity with a high temporal resolution. Through two carefully controlled designs, I contrasted utterances that contain the factual verb *do* with utterances that contained modal expressions such as *may* and *must*. The combination of the results from these two experiments suggests that the brain is sensitive to the contrast between fact and possibility rapidly after its presented, and that discourse situation updating only takes place for factual information.

The second part of this dissertation focuses on first language acquisition and investigates children's developing ability to process counterfactual language. While prior research mostly investigated children's counterfactual development using past counterfactual conditionals (e.g., "If I **had** drawn on this piece of paper, where would it have gone"), these structures are linguistically more complex and less frequently used in the child's input than other counterfactual constructions such as the present conditional (lacking past perfect: "If I **had** a piece of paper, I would draw on it") and wish-construction (lacking causal dependency: "I wish I **had** a piece of paper"). Adding to the linguistic complexity of the counterfactual construction, is the presence of the so called "fake" past tense (in bold) (Iatridou, 2000), which refers to a temporal mismatch between the morphological tense and intended temporal orientation. In order to acquire counterfactual constructions, children have to detect this “fakeness” and learn to map this past to counterfactuality instead. Besides lacking a causal dependency, wishes in English are also linguistically less complex than counterfactual conditionals because of their transparency. The temporal mismatch between morphological tense and temporal orientation is more transparent in wishes than in conditionals. *Wish* is dedicated to embedding counterfactual conditions and cannot co-occur with the present tense (*I wish I have a piece of paper), which might facilitate the task of mapping counterfactual meaning to the “fake” past tense form. In my research, I considered the form-to-meaning mapping challenge that children encounter when linking counterfactual meaning to counterfactual constructions. Specifically, I aimed to de-confound cognitive complexity and linguistic complexity to get a more accurate estimate of young children’s development.

In Chapter 3, I addressed two questions about children’s early counterfactual productions: 1) Do children ever make spontaneous tense errors in counterfactual constructions, using present tense rather than past, which would suggest they initially struggle with the temporal mismatch present in counterfactuals? 2) Do children start to produce counterfactual wishes before the more complex counterfactual conditionals? A large-scale corpus study on children’s transcribed speech shows that wishes (first uses around age 2 or 3) are indeed produced before counterfactual conditionals (first uses around age 3 or 4), although children’s initial usage is not distinguishable from expressing a regular desire. Similarly, the study finds that children make the productive error of producing counterfactuals with present tense marking instead of past. These errors are consistent with a stage where children have yet to figure out that counterfactual past tense signals a present non-actuality, rather than a past event on the timeline. However, usage data alone cannot confirm whether children also understand wishes before counterfactual conditionals, and whether they are initially misled by the counterfactual's "fake" past.

In Chapter 4, I examined 4- and 5-year-old children's comprehension of counterfactual constructions. In particular, I asked: 1) Do children ever misunderstand the counterfactual's "fake" past to be real? And, 2) do children comprehend the less complex counterfactual wishes before they understand counterfactual conditionals? The results show that children's performance on the counterfactual *wish*-construction exceeds their performance on counterfactual conditionals, and that there is some evidence that the comprehension of counterfactual utterances is further complicated by the presence of a "fake" past tense marking, which sometimes gets interpreted as being real. The combined findings of Chapter 3 and 4 suggests that prior work that focused only on children's comprehension of counterfactual conditionals might have underestimated children's early counterfactual reasoning abilities.

The dissertation concludes with a general discussion combining the insights of these separate studies and the open questions for future investigation (Chapter 5). With the knowledge gained from this dissertation work, I evaluate different models for representing actual and non-actual information and discuss how the fields of cognitive neuroscience and first language acquisition can inform each other and help us build towards a broader understanding of the cognitive ability of human language displacement. A greater understanding of the neural mechanisms that underly modal displacement in adults and a clear idea of when this ability develops throughout childhood, will eventually pave the way for future research on the development of these neural mechanisms and provide insight into the development of displacement from an evolutionary perspective.

CHAPTER 2: NEURAL CORRELATES OF MODAL DISPLACEMENT AND DISCOURSE-UPDATING UNDER (UN)CERTAINTY

As published in Maxime Tulling, Ryan Law, Ailís Cournane & Liina Pylkkänen (2021, eNeuro)

1. ABSTRACT

A hallmark of human thought is the ability to think about not just the actual world, but also about alternative ways the world could be. One way to study this contrast is through language. Language has grammatical devices for expressing possibilities and necessities, such as the words *might* or *must*. With these devices, called "modal expressions," we can study the actual vs. possible contrast in a highly controlled way. While factual utterances such as "There is a monster under my bed" update the here-and-now of a discourse model, a modal version of this sentence, "There might be a monster under my bed" displaces from the here-and-now and merely postulates a possibility. We used magnetoencephalography (MEG) to test whether the processes of discourse updating and modal displacement dissociate in the brain. Factual and modal utterances were embedded in short narratives, and across two experiments, factual expressions increased the measured activity over modal expressions. However, the localization of the increase appeared to depend on perspective: signal localizing in right temporo-parietal areas increased when updating the representation of someone else's beliefs, while frontal medial areas seem sensitive to updating one's own beliefs. The presence of modal displacement did not elevate MEG signal strength in any of our analyses. In sum, this study identifies potential neural signatures of the process by which facts get added to our mental representation of the world.

2. INTRODUCTION

Speculating about possibilities employs our unique human capacity to displace from the *here-and-now* (Hockett, 1959; Bickerton, 2008; Suddendorf et al., 2009). We can express possibility using 'modal expressions' like "There might be a monster", shifting our perspective from the immediate present to a hypothetical scenario. Other cognitive abilities that shift into alternative perspectives, like thinking about the past or future and conceiving the viewpoints of others, seem to share a brain network consisting of hippocampal and parietal lobe regions (Buckner & Carroll, 2007; Mullally & Maguire, 2014). However, we know surprisingly little about the neural mechanisms involved in modal displacement. While factual statements like "There is a monster" update our beliefs about a situation, modal utterances indicate uncertainty instead. Are the mental operations of discourse updating and modal displacement dissociable in the brain? Here, we investigated the neural correlates of integrating factual and modal utterances into an existing discourse representation.

2.1. Cognitive Processes Involved with Comprehending Discourse

When comprehending discourse, we represent the perspective, place and time of the discussed situation (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998), and distinguish between facts and possibilities compatible with the here-and-now of this alternative reality. Consider this scene from Ovid's tale about the ill-fated lovers Pyramus and Thisbe.

> *When a lioness, bloody from hunting, approaches, Thisbe flees into a cave, losing her shawl in the process. As Pyramus encounters the lioness hovering over Thisbe's bloodstained shawl with his lover nowhere in sight, he quickly concludes <u>she **must** have been devoured by the beast</u>.*

All but the underlined sentence are factual claims made about the actual state of affairs (R. Stalnaker, 1996). We use these utterances to build a mental situation model, which is dynamically updated as new information becomes available (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004). Maintaining these discourse models elicits activation in the medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC) and temporo-parietal areas (Speer et al., 2007; Whitney et al., 2009; Xu et al., 2005; Yarkoni et al., 2008). To interpret the narrative above, we also engage in higher order cognitive processes such as modal displacement and Theory of Mind (ToM) reasoning (Premack & Woodruff, 1978). ToM is the ability to represent someone else's belief state separately from our own, allowing us to understand how Pyramus induced that Thisbe died, even though we know she is still alive. Pyramus based his conclusion on indirect evidence (the bloody shawl), signaling with the modal verb *must* that the devouring is not actual or known. Modals like *must* or *may* allow reasoning about open possibilities compatible with a situation (Kratzer, 1981, 2012; Phillips & Knobe, 2018; von Fintel, 2006).

Since ToM and modal displacement both require a representation that is different from the actual situation (Phillips & Norby, 2019), they may recruit overlapping brain areas. While there has been no systematic study of the neural bases of modal processing, ToM tasks are consistently reported to activate the dorsal/posterior inferior parietal lobule (IPL), temporoparietal junction (TPJ), medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC) and rostral anterior cingulate cortex (rACC) (e.g., Koster-Hale et al., 2017; Mahy et al., 2014; Schurz & Perner, 2015). In particular the right TPJ seems involved in representing other's mental state (Saxe & Powell, 2006; Saxe & Wexler, 2005; Vistoli et al., 2011) though some suggest this activity may be attributable to more domain general cognitive processes such as reorienting attention (Corbetta et al., 2008; Decety & Lamm, 2007; Mitchell, 2008; Rothmayr et al., 2011). Definitions of the key concepts used throughout this paper are provided in Figure 2.1.

Key Concepts	Definitions
Counterfactuality	Language category used to discuss alternative ways the world could be or could have been, e.g. *if the monster were big, it wouldn't fit under the bed*
Discourse Updating	Updating an existing situation model when the situation's *here-and-now* changes, e.g. change in protagonist, goal, location, event or time
Factuality	Language category concerning what is known to be true or false in a situation
Hypothetical Scenario	Situation that is temporarily stipulated to be true that may or may not conflict with what is accepted as true about the world
Modal Base	The grounds on which the likelihood of a hypothetical scenario is determined, i.e. based on what you know (KNOWLEDGE-BASED) or on what the circumstances are, e.g. rules and norms (RULE-BASED)
Modal Displacement	An operation that shifts our perspective from the immediate present (*here-and-now*) to a hypothetical scenario
Modal Force	The degree of certainty for a hypothetical scenario *x* to be true, i.e. whether it is POSSIBLE (imagining all reasonable possibilities there is at least one in which *x* is the case) or NECESSARY (*x* is the case for each reasonable possibility imaginable)
Modality	Language category used to discuss hypothetical possibilities, e.g. *may*, *must*, *might*
Presupposed Content	Information that is taken for granted within the discourse, e.g. '*since the monster is big'* presupposes prior knowledge of the existence of a big monster
Situation Model	Mental representation of a situation, tracking events, actions and persons related to the *here-and-now* of that being discussed
Theory of Mind (ToM)	The ability to reason about mental states and represent the belief state of others separate from our own

Figure 2.1 Table containing key concepts and definitions as used throughout this paper.

2.2. This Study

How do our brains distinguish between information that states facts versus information that only conveys possibilities? We investigated the differences between factual and modal language comprehension in two experiments (Figure 2.2). We used magnetoencephalography (MEG), providing us with high temporal resolution and relatively good spatial localization of brain activity during sentence comprehension. Experiment 1 investigated the neural bases of discourse updating and modal displacement by contrasting sentences that contain modal verbs against sentences containing the factual verb 'do' embedded in short narratives. In experiment 2, we further investigated under which conditions discourse updating takes place by manipulating the certainty of the sentential context in which the target verbs (factual vs. modal) were embedded: factual (certain), conditional (uncertain) or presupposed (already known). Discourse updating should take place under actual situational changes (e.g., when new factual information is added to a factual context), but not when novel information is hypothetical (modal conditions) or when the entire context is hypothetical (conditional context). Modal displacement should occur whenever utterances postulate hypothetical possibilities.

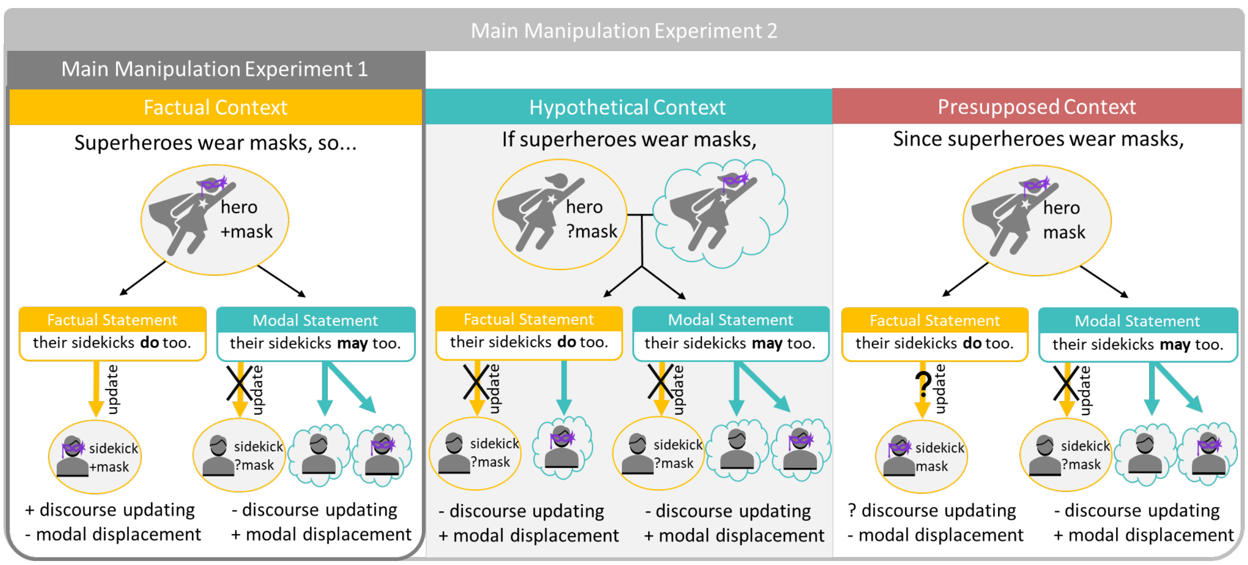

Figure 2.2. Simplified illustration of main manipulations Experiment 1 and 2. Model of operations assumed to be present during the processing of factual (yellow) and modal (teal) statements (simplified from actual stimuli). Experiment 1 contrasts factual and modal statements in a factual discourse context, while Experiment 2 varies whether the discourse context is factual, hypothetical, or presupposed. Updating of the discourse situation model (round) is expected to take place under certainty (in factual contexts with a factual update). Both modal (*may*) and conditional expressions (*if superheroes wear masks*) evoke hypothetical situations (cloud) involving modal displacement. Since the presupposed context marks information already known, we are not sure whether updating would take place.

3. METHODS

3.1. Experiment 1

3.1.1. Participants

26 right-handed, native English speakers participated in the experiment (4 male) taking place at the New York University (NY) campus. One participant was excluded from further analysis for having an accuracy lower than 70% on the behavioral task. The age range of the remaining 25 participants was 19-52 years old (M= 25.7, SD = 7.46). All participants had normal or corrected to normal vision, no history of neurological impairment and provided informed written consent.

3.1.2. Stimuli

We developed an experimental paradigm where we contrasted the modal verbs *may* and *must* against the factual auxiliary verb *do*. In order to have *do* naturally appear in the same position as *may* and *must*, our sentences contained verb phrase (VP) ellipsis, e.g., “Normally only knights sit at the round table, but the king says that the squires *may*/*must*/*do* <~~sit at the round table~~> too.” While the verb *do* indicates factuality, modals indicate hypothetical scenarios that are compatible with the actual world given someone’s knowledge or the set of circumstances. We specifically chose to use the modal expressions *may* and *must* because they vary among two dimensions: ‘modal force’ and ‘modal base’. Modal force refers to the likelihood of a hypothetical situation, i.e., whether it is deemed a possibility (*may*) or a necessity (*must*). The modal base denotes what we base this likelihood assessment on: our knowledge or the circumstances, e.g., rules/norms. The modals *may* and *must* are ambiguous in allowing for both a knowledge-based (e.g., “Given what I know, there may/must be a monster under my bed”) and a rule-based reading (e.g., “Given what the rules are, you may/must eat your dinner now"). Using such ambiguous modals, we could compare the effect of modal base without varying the form of the target item.

We constructed 40 sets of short English narratives. Each story consisted of three sentences, starting with a context sentence designed to either bias towards a knowledge-based (epistemic) scenario, or a rule-based (deontic) scenario. The context sentence was followed by a target sentence and each story ended with a final task sentence that was either congruent or incongruent with the previous two sentences (Figure 2.3A). The target sentences contained the target modal verb (the possibility verb *may* or the necessity verb *must*) and were compared against the factual condition containing the verb *do*. In the context sentence a property or habit was introduced that applied to one group (e.g., "knights sit at the round table"), and the target sentence indicated was also (possibly) the case for another group (e.g., "their squires do/may/must too"). Each stimulus set therefore consisted of 6 sentences (2x3, BASE: [knowledge, rules] x FORCE: [possibility, necessity, factual]) adding up to a total of 240 sentences for all 40 stimuli sets (Figure 2.3B). The third sentence of the story was a task sentence either congruent (50%) or incongruent (50%) with the prior two sentences. One third of the task sentences were specifically tapping into the congruency of the modal base (Figure 2.3C). Across conditions, how often task items were congruent or incongruent with the preceding sentences was controlled for, as was how often questions tapped into information obtained from the context or target sentence.

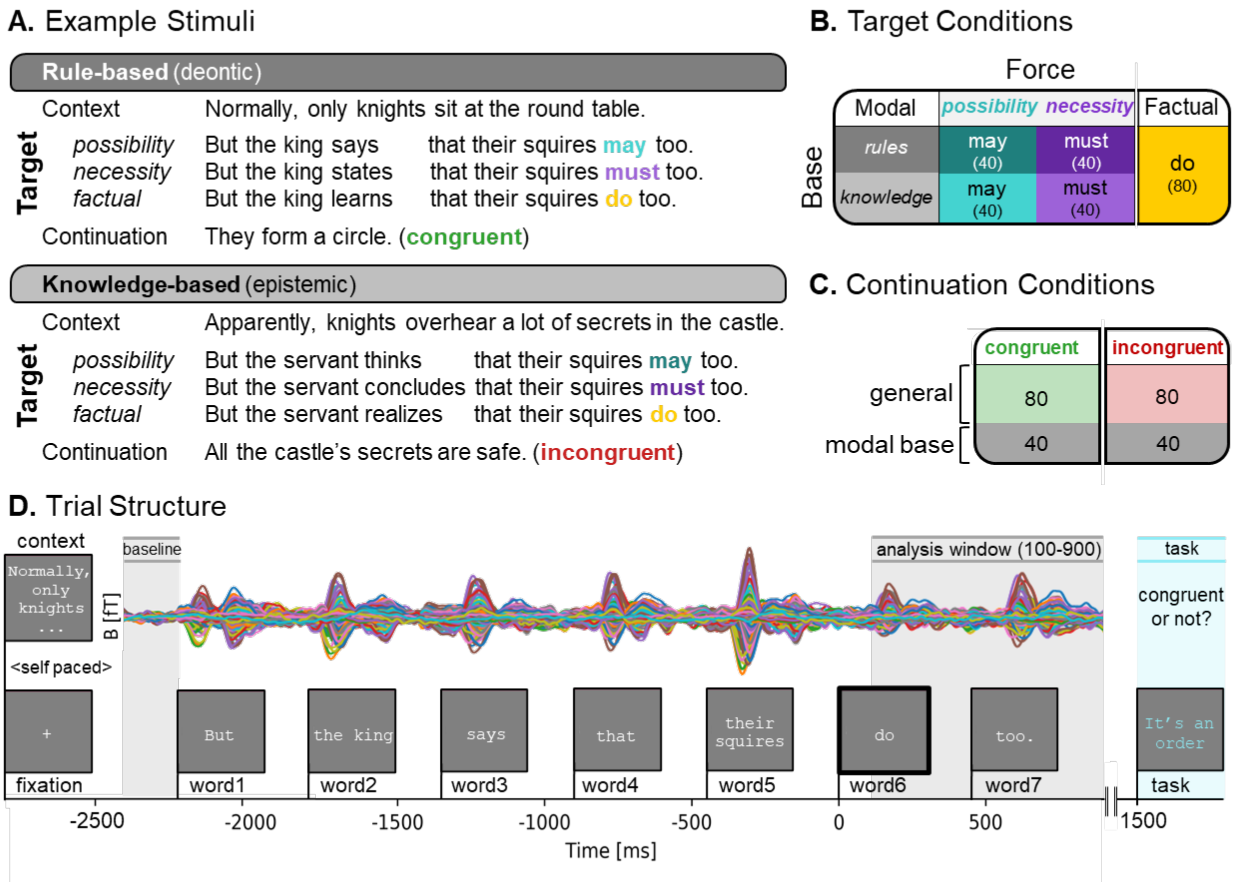

Figure 2.3. Design and procedure Experiment 1. **A**: Example stimuli set. Short narratives consisted of three parts. A context sentence biasing towards a rule-based or knowledge-based modal interpretation, followed by the target sentence containing one of the target verbs varying in force (possibility, necessity or factual). The third continuation sentence was either congruent or incongruent with prior sentences. Details on controlled between-stimuli variation can be found in Appendix S2.1 **B**: Experimental design with number of items per condition in brackets (total = 240). The stimuli vary along two dimensions: MODAL BASE [rules, knowledge] and FORCE [possibility, necessity, factual]. **C**: Continuation Conditions. Half of the continuations are incongruent with the previous sentences. One third tap into modality and are congruent or incongruent with the modal base of the previous sentences. **D**: Trial structure with evoked MEG responses from one participant. A context sentence was displayed until participants pressed a button. After a fixation cross (300ms) the target sentence was displayed word-by-word for 300ms each followed by a 150 ms blank screen. The continuation sentence was displayed with a 600ms delay, and participants indicated by button press whether this was congruent or incongruent with the prior story. Time windows for baseline correction (-2450 to -2250ms) and statistiacal analysis (100-900ms) are relative to the target verb (word6) onset.

All target sentences had the same sentence structure: CONNECTIVE (but/and/so)| the | NOUN.SG | VERB1 | that | DETERMINER | NOUN.PL | TARGET (may/must/do) | <ELIDED VP> too. The embedded clause of the sentence (introduced by *that*) was kept consistent across all conditions. We controlled for between-item variation in the other parts of the stimuli along the following dimensions: the count of different CONNECTIVES and DETERMINERS among the modal base conditions, the average length, frequency, number of syllables and morphemes of NOUN.SG among different modal base conditions, and the average length (in words and letters), stativity, transitivity and structural complexity of the <ELIDED VP> material in the target sentence across different base conditions (see Appendix S2.1). The information on lexical frequency and morpheme length was obtained from the English Lexicon Project (Balota et al., 2007). Within the modal base dimension, the target sentences only varied in the embedding verb (VERB1) to support biasing the reading of the target modal verb. Embedding verbs were divided into three categories occurring with knowledge-based, rule-based or factual targets. Each verb category contained 12 different verbs, which were repeated maximally 7 times across the entire experiment. Between the two base conditions, the knowledge-based and rule-based sentences also differed in their preceding context sentence and subject, to help bias the interpretation of the ambiguous modals *may* and *must*. In order to encourage the rule-based reading, the context introduced an event that was compatible with both permission or obligation (e.g., sitting at the royal table), and the target sentence introduced a third person subject that was in an authority position over the sentence object (e.g., a king over squires). In order to encourage the knowledge-based reading, the context introduced an event that was very unlikely to be permitted or obliged (e.g., overhearing secrets) and the target sentence introduced a subject that was in a bystander position to the event (e.g., a servant). By embedding the target utterance into the perspective of a third person subject, the assessment of the modal force (whether something was possibly, necessarily or factually true) was linked to the perspective of this character.

The effectiveness of the biasing conditions was tested with a survey on Amazon Mechanical Turk made with the help of Turktools (Erlewine & Kotek, 2016). For this norming, the target sentences containing modal verbs (160 items in total) were adjusted so that unambiguous adjectives replaced the ambiguous target modal verbs. Knowledge-based *may* was replaced with *are likely to*, knowledge-based *must* with *are certain to*, rule-based *may* with *are allowed* to and rule-based *must* with *are obliged* to. E.g., the target sentence "But the king says that the squires may too" became "But the king says that the squires are allowed to as well". These unambiguous target sentences were then displayed with their preceding context sentence and a gap substituting the adjective. Participants (n=320) were asked to choose which of 4 options (*obliged, allowed, likely* and *certain*) would fit the gap best. Each target sentence was judged 32 times across all participants. The experiment took about 2-4 minutes and participants were paid $0.20 for completing the experiment. Each participant completed 25 sentences, comprised of 20 test items and 5 filler items that served as an attention control, in random order counterbalancing for condition. Results were excluded from participants that indicated to not have English as a native language (n=17) and from participants that made more than 1 mistake on the filler items (n=6). For the responses of the remaining 297 participants we noted whether the modal base of their response (*allowed* and *obliged* = rule-based, *likely* and *certain* = knowledge-based) matched the intended modal base of the target items or not. For each item, we calculated the average percentage of matches with the intended modal base (bias score), and only approved an item for the experiment if its bias score was 70% or higher. This norming happened in two parts. In the first round, all 160 items were tested, and 137 items were accepted. The remaining 23 items had a bias score below the 70% threshold and were altered to improve their bias. In the second round, these 23 items were re-tested (now mixed with a random selection of the previously approved items) and judged with the same criteria. This time 18 items were accepted, and 5 scored below the 70% threshold. The 5 items that did not pass the norming experiment were altered again with the help and approval of several native speakers, and then included into the experiment.

The lexical frequency of knowledge-based (epistemic) and rule-based (deontic) readings of *may* and *must* are not evenly distributed in written American English: the verb *may* is knowledge-based about 83% of the time (Collins, 2007), while *must* is knowledge-based 16% of the time (Hacquard & Wellwood, 2012), in all other cases the verb has a circumstantial base that includes rule-based meanings. While these lexical frequency differences may have an effect on the processing of the individual items, we expect that grouping the different levels of the force (grouping *knowledge-based* and *rule-based* responses together) or modal base manipulation (grouping *possibility* and *necessity* responses together) should wash out any effects of this imbalance.

3.1.3. Procedure

Before recording, the head shape of each participant was digitized using a FastSCAN laser scanner (Polhemus, VT, USA). Additionally, we recorded the location of three fiducial locations (the nasion, and left and right preauricular points) and five reference points for purposes of co-registration. Before participants entered the MEG-room they received verbal instructions and did a short practice block (of eight trials). Data collection took place in a magnetically shielded room using a whole-head MEG system (157 axial gradiometer sensors, 3 reference magnetometers;

Kanazawa Institute of Technology, Nonoichi, Japan). Before the experiment, we taped five marker coils on the location of the digitized reference points that help establish the position of the subject's head before and after the experiment. During the experiment, the participant comfortably lay down in the MEG machine, reading from a screen located approximately 50 cm away with dimmed lights. Text was displayed in a fixed-width Courier New font on a light grey background.

In the experiment, participants were asked to silently read and comprehend short stories consisting of three sentences presented with PsychoPy (Peirce, 2009). The first sentence (context) was displayed as a whole. Participants read this sentence at their own pace and pressed a button to continue. Then a fixation cross (300ms) followed and after a 300ms blank screen the target sentence was presented using Rapid Serial Visual Presentation. Participants were presented with English sentences of 9 words, mostly one word at the time, with the exception of determiner-noun pairs, which were presented together so that the sentence was divided into 7 parts (called 'words' from now on). The display time for all words was 300ms. Every word was preceded by a blank screen of 150ms. This was followed by a short third sentence in blue that was either congruent with the previous sentence or incongruent (50%). The continuations were designed such that they targeted the comprehension of different parts of the story (encouraging participants to read the entire narrative with care). One third of the continuations tapped into the modality of the target sentence, in which the continuation is congruent with the modal base (e.g., a sentence about obligation followed by "their mother told them to") or incongruent with the modal base (e.g., a sentence about obligation followed by "she's probably right"). We included this manipulation to be sure that participants are paying attention to the fine meaning of the modal target verb. The participant's task was to press one button with their middle finger for continuations that 'made sense' and another button with their index finger if the continuations 'did not make sense', after which the next trial started. The participants were instructed to move and blink as little as possible during the task. The trial structure is displayed graphically in Figure 2.3D.

The experiment consisted of 240 trials in total. The trials were divided into 6 separate blocks (containing 1 item per stimuli set) by a balanced Latin square design and randomized within blocks. Each block consisted of 40 sentences and was presented into two parts during the experiment, resulting into 12 blocks which took about 3-7 minutes each. In between blocks, participants were informed about their overall accuracy. Participants were free to rest in between blocks and were paid $15 (NY) per hour.

3.1.4. Data acquisition

MEG data were sampled at 1000 Hz with an online 200 Hz low-pass filter. The signal was offline noise reduced in the software MEG160 (Yokogawa Electric Corporation and Eagle Technology Corporation, Tokyo, Japan) using the signal from the three orthogonally-oriented reference magnetometers (located within the machine, but away from the brain) and the Continuously Adjusted Least-Squares Method (Adachi et al., 2001). Further pre-processing and analysis was performed making use of MNE-Python (Gramfort et al., 2013, 2014) and Eelbrain (Brodbeck, 2017). First, MEG channels that were unresponsive or clearly malfunctioning (separating from all other channels) during the session were interpolated using surrounding channels (6% of the channels in total underwent interpolation, 7-19 channels per participant). We extracted epochs from -2450 to 900 ms relative to the onset of the target verb, which included the entire sentence. The epochs were corrected for the delay between presentation software timing and stimulus presentation, by taking into account the average delay as measured with a photodiode. The data were filtered offline with a band-pass filter between 1 and 40 Hz. Eye blinks and heartbeat artefacts were removed by the use of Independent Component Analysis (ICA) via the "fastICA" option implemented in MNE python (Gramfort et al., 2014). Additionally, we removed a known artefact pattern ('the iron cross') that was present at that time across all NY recordings due to an electromagnetic noise source from nearby cables. Any epoch that had a sensor value that was higher than 3pT or lower than -3pT were automatically rejected. Additionally, trials were rejected after visual inspection if multiple channels were affected by obvious noise patterns that exceeded the boundaries of the epoch's window. In total, this resulted in a trial-rejection rate of 4.6% across the experiment. Baseline correction was performed using data from the 200 ms before the first word of the sentence.

The location of sources was estimated by co-registration of the digitized head shape with the FreeSurfer average brain (Fischl, 2012). A source space containing 2562 sources per hemisphere was constructed for each subject, and a forward solution was created with the Boundary Element Model method. The inverse operator was calculated based on the covariance matrix from the 200 ms pre-stimulus baseline period of the cleaned trials. This inverse operator was applied to the average evoked responses to obtain a time course of minimum norm estimates at each source for each condition (SNR = 3). The direction of the current estimates was freely oriented with respect to the cortical surface, and thus all magnitudes were non-negative. The source estimates were then noise-normalized at each source (Dale et al., 2000), generating dynamic statistical parameter maps (dSPM) that were used in statistical analyses.

3.1.5. Statistical Analyses

3.1.5.1. Behavioral data:

Responses and reaction times to the 6000 (25x240) congruency decisions were collected and overall accuracy was determined based on the responses to all items. The overall accuracy was used to exclude participants if they scored below 70%. We also examined the accuracy of the 2000 modal task items.

3.1.5.2. MEG data:

MEG data were analyzed both with an ROI analysis and with a full-brain analysis, given the explorative nature of our question.

<u>3.1.5.2.1. ROI Analysis:</u>

Since there is no prior neuroimaging work on the processing of modals, our ROIs were defined based on previous literature looking at the neural bases of Theory of Mind (Koster-Hale et al., 2017; Mahy et al., 2014; Schurz & Perner, 2015), and included the Inferior Parietal Sulcus (IPS), Temporo-Parietal Junction (rTPJ), Superior Temporal Sulcus (STS), Posterior Cingulate Cortex (PCC), rostral Anterior Cingulate Cortex (rACC) and medial Prefrontal Cortex (mPFC) bilaterally. These functional regions were translated into labels for (bilateral) areas mapped onto the FreeSurfer aparc (Desikan et al., 2006) parcellation (Table 2.1). Each source current estimate was mapped onto a parcellation, and then averaged over all the sources in each ROI.

Table 2.1. Overview of regions of interest (ROIs) based on the aparc parcellation, with approximately corresponding Brodmann Areas (BA) and number of sources.

Label	Aparc	BA	N. of Sources
Inferior Parietal Sulcus (IPS)	superiorparietal	7	162
Temporoparietal Junction (TPJ)	supramarginal + inferiorparietal	39 + 40	278
Superior Temporal Sulcus (STS)	superiortemporal	22	108
Posterior Cingulate Cortex (PCC)	posteriorcingulate	23+31	49
rostral Anterior Cingulate Cortex (rACC)	rostralanterior-cingulate	24+32	15
ventromedial Prefrontal Cortex (vmPFC)	medialorbitofrontal	25+10+11	44

The effect of the experimental manipulations on our ROIs was assessed with a cluster-based permutation test (Maris & Oostenveld, 2007), aimed to identify temporal clusters that were affected by our experimental paradigm, corrected for multiple comparisons. We performed a temporal cluster-based permutation mass univariate 2 X 3 repeated-measures analysis of variance (ANOVA) with factors MODAL BASE and FORCE. Since we had no clear predictions about the possible timing of an effect, we used the generous time window of 100-900 milliseconds after the target verb's onset. Since several trials got rejected during data pre-processing, to ensure comparable SNR across conditions we equalized trial count across conditions (M=36 trials/condition, range=31-39trials/condition)

Our temporal permutation clustering test was performed in Eelbrain 0.27.5 (Brodbeck, 2017) with a standard procedure. An uncorrected ANOVA was fitted at each time point in the analysis time window (100-900 ms). Temporal clusters were formed and chosen for further analysis when *F*-statistics corresponded to significance exceeded the critical alpha-level of .05 (uncorrected) for contiguous time points of at least 25 milliseconds. A test statistic corresponding to the cluster magnitude was then determined by summing over all the F-values contained within them and selecting the largest of the cluster-level statistics. Conditions were re-labeled, and test statistics were calculated for each subject for 10,000 times to form a null distribution of the test statistics. The observed clusters were compared to this null distribution and were assigned corrected *p*-values reflecting the proportion of which random partitions resulted in an *F*-statistic greater than the observed *F*-statistic. Since in this method, the time point clusters initially chosen for further analysis are uncorrected, the borders of the clusters should be interpreted as having an approximate nature, not making claims about the *exact* latency or duration of any effects (see Sassenhagen and Draschkow, 2019). Finally, in order to also correct for comparisons across multiple ROIs, we applied a False Discovery Rate correction for multiple comparisons (Benjamini and Hochberg, 1995).

3.1.5.2.2. Whole Brain Analysis:

To complement our ROI analysis, we conducted a full brain analysis, which both described the full spatial extent of any effects observed in the ROI analysis and provided us with information about any effects not captured by the ROI analysis. We performed a spatiotemporal clustering test almost identical to the temporal cluster test described above, only now without averaging sources within an ROI. Instead, an *F*-statistic was calculated for each time point in each source, and spatiotemporal clusters were identified where significance exceeded a *p* value of .05 for at least 10 spatially contiguous sources and for at least 25 milliseconds. Again, following Sassenhagen and Draschkow (2019), the temporal and spatial properties of the identified significant spatio-temporal clusters should be interpreted as an approximate description.

3.2. Experiment 2

3.2.1. Participants

Human subjects were recruited on New York University's New York (NY) and Abu Dhabi (AD) campuses. 24 right-handed, native English speakers participated in the experiment (8 male, 12 in AD). Four participants were excluded (1 for not finishing the experiment due to a technical complication, 1 for excessive channel loss and 2 for extreme noise during recording, rendering the data unusable). The age range of the remaining 20 participants was 19-42 years old (M= 26, SD = 6.46). All participants had normal or corrected to normal vision, no history of neurological impairment and provided informed written consent. To mitigate our participant loss, we did not exclude participants based on behavioral accuracy. Participants were pseudo-randomly assigned one of three experimental lists, such that participants were equally divided over each experimental condition.

3.2.2. Stimuli

We developed a similar experimental paradigm as Experiment 1, now manipulating the information value of the sentential context rather than manipulating properties of the modal items (modal base and force). We constructed 40 sets of bi-clausal English sentences, containing a causal relationship between the two parts. We contrasted the factual auxiliary verb *do* against the possibility modal verbs *may* and *might*, keeping modal force consistent across items.

Sentences differed in their informative content and came in three types: FACTUAL e.g., “Knights carry large swords, so the squires do too”, which introduced novel information with certainty, CONDITIONAL e.g., “If knights carry large swords, the squires do too”, which introduced novel information with uncertainty (indicated by *if*), and PRESUPPOSED, e.g., "Since knights carry large swords, the squires do too", which introduced presumed to be known information (indicated by *since*) with certainty. The main manipulation (FACTUAL VS CONDITIONAL) was added to test whether a possible effect of belief updating (expected to be present when encountering the factual target verb in the factual condition) disappeared if the information update built on uncertainty (conditional condition). For processing modal displacement, we did not expect a possible effect to be influenced by sentential certainty. We included the PRESUPPOSED condition for exploratory purposes. Each sentence was preceded by a context word, indicating the theme of the upcoming sentence, e.g., "CASTLE", to stay consistent with Experiment 1, where utterances were preceded by a context sentence. Since Experiment 2 did not vary modal base, we differentiated from Experiment 1 by no longer embedding the target utterance into the perspective of a third person subject (used to bias towards modal base readings in Experiment 1), in order to reduce sentence length. The complete stimulus design and predictions are displayed in Figure 2.4A. Each stimulus set consisted of 9 sentences (3x3, TYPE: [factual, conditional, presupposed] x VERB: [*may*, *might, do*]) adding to a total of 360 sentences for all 40 stimuli sets (Figure 2.4B).

All utterances were equal in length. Since we pursued a within-participants design and the different sentence conditions within a stimulus set differed minimally, we introduced controlled variance in the first clause of the utterance to make the paradigm seem less repetitive. We constructed three semantically related variants of the subject (e.g., *knights*, *noblemen* and *commanders*) and main event (e.g., *carrying heavy armor*, *owning many weapons* and *using large swords*) that were matched across conditions in a stimulus set so that each subject and action occurred in each of the 9 conditions once. We made three different versions of the experiment such that across versions each condition occurred with all the subject and event variants. Sentential subjects denoted generic groups (e.g., *knights* or *loyal supporters*) and personal/company names (such as *Lisa* or *Facebook*).

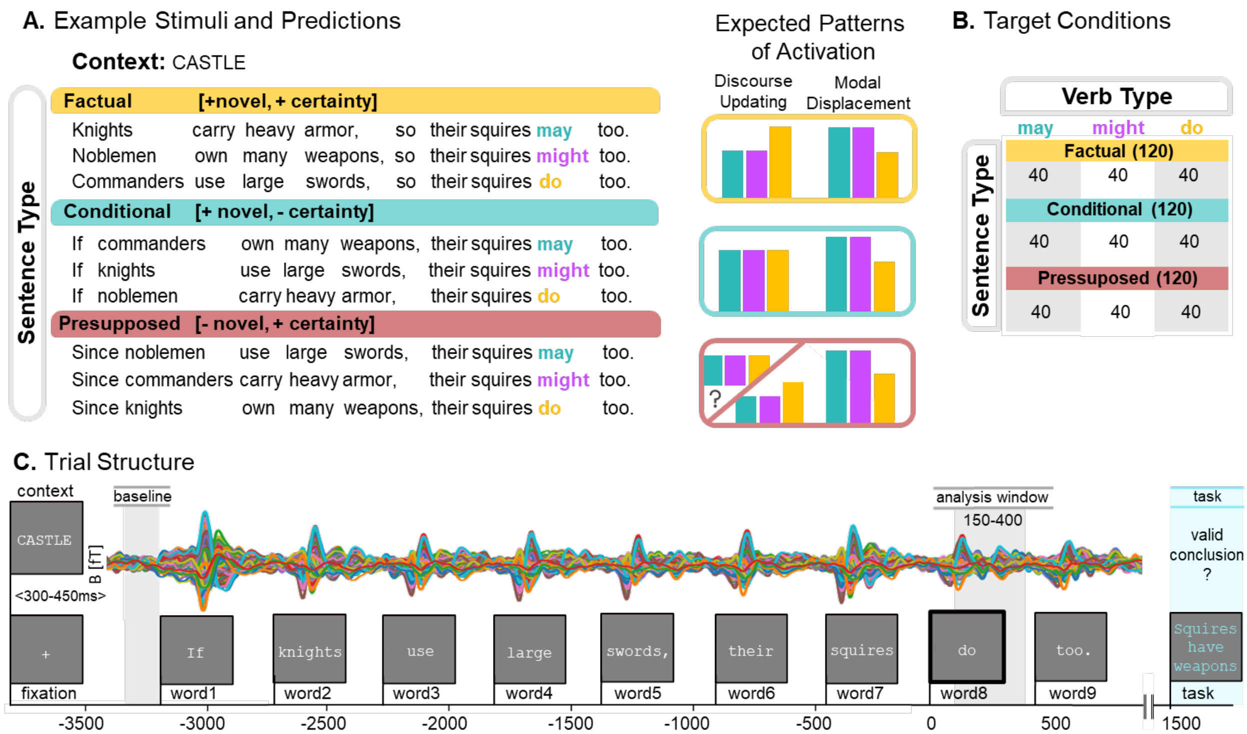

Figure 2.4. Experimental design and procedure Experiment 2. **A**: Example stimuli set and Predictions. All stimuli were bi-clausal sentences of three different types: factual (*p* so *q*), conditional (if $p \rightarrow q$) and presupposed (since $p \rightarrow q$). These sentence types differed in whether they express information that is novel and certain (factual), novel and uncertain (conditional) or known and certain (presupposed). Each sentence contained either the factual verb *do* or the modal verbs *may* or *might*. Included are expected activation patterns for each verb per sentence type under processes of belief updating and modal displacement. We expect belief updating to take place in factual contexts but not in conditional contexts. For presupposed contexts we had no clear predictions. Activity related to modal displacement is not expected to change across different sentential environments. **B**: Experimental design with number of items per condition displayed between brackets (total = 360). The stimuli vary among two dimensions: SENTENCE TYPE [factual, conditional and presupposed] and VERB [may, might, do]. **C**: Trial structure with evoked MEG responses from one participant. Procedure similar to Experiment 1. Time windows for baseline correction (-3350 to -3200ms) and statistical analysis (150-400ms) are relative to the target verb (word8) onset.

3.2.3. Procedure

Before recording, we digitized the head shape of each participant with either a FastSCAN laser scanner or a FASTRAK 3D digitizer (Polhemus, VT, USA), following the same procedure as laid out in for Experiment 1. Before participants entered the MEG-room they received verbal instructions and did a short practice block of seven trials. Data collection took place in a magnetically shielded room using whole-head MEG system with 157 (NY) or 208 (AD) channels (Kanazawa Institute of Technology, Kanazawa, Japan). Stimuli were projected onto a screen located above the participant. We made sure to keep the visual angle across both systems consistent, at approximately 0.5° vertically.

In the experiment, participants were asked to silently read and comprehend causally linked sentences presented with PsychoPy (Peirce, 2009), font and background settings identical to Experiment 1. First, a context word was displayed for 600ms followed by a blank screen which display time varied between 300-450 ms. This jitter in display time was included to approximate the temporal variety in Experiment 1 induced by self-paced reading of the context sentence. Then, a fixation cross (300ms) followed and after a 300ms blank screen the target sentence was presented using Rapid Serial Visual Presentation. Participants were presented with English sentences of 9 words, one word at the time (300ms on and 150ms off). This was followed by a conclusion (displayed in blue) that was either a valid conclusion based on prior information (50%) or not. This task was designed such that participants had to pay close attention to the fine details of the target utterances. Forty percent of the questions specifically tapped into the certainty of the prior statement (e.g., the sentence “If knights own many weapons, their squires do too” followed by the valid conclusion “Potentially, the squires own many weapons” or invalid conclusion “The squires own many weapons"). Half of these certainty-based conclusions targeted the first clause of the sentence, while the other half targeted the second half. The other conclusions (60%) were more general e.g., "Knights have (no) squires". The participant's task was to press one button with their middle finger for conclusions that were valid and another button with their index finger if the conclusions were invalid, after which the next trial started. The participants were instructed to move and blink as little as possible during the task. The trial structure is displayed in Figure 2.4C.

The experiment consisted of 360 trials in total. The trials were divided into 9 separate blocks (containing 1 item per stimuli set) using a balanced Latin square design and randomized within blocks. Each block consisted of 40 sentences and was presented in two parts during the experiment, resulting in 18 blocks which took about 3-5 minutes each. In between blocks, participants were informed about their overall accuracy. Participants were free to rest in between blocks and were paid $15 (NY) or 60 AED (AD) per hour.

3.2.4. Data acquisition

The same acquisition profile was maintained across both NY and AD systems, with settings as described for Experiment 1. Preprocessing used the same software and pipeline as described for Experiment 1. In total, 7% of the channels were interpolated due to being unresponsive or clearly malfunctioning (NY: 7-14 per participant; AD: 0-18 per participant). We extracted epochs from -3500 to 1200 ms relative to the onset of the target verb, which included the entire sentence, and rejected epochs containing signal amplitudes that exceeded a threshold of 3 pT (NY) or 2 pT (AD). The NY threshold is higher since that city and system has higher levels of overall ambient magnetic noise. In total, this resulted in a trial-rejection rate of 3.9% across all participants (NY: 5.0%; AD: 2.0%). Baseline correction was performed using data from -3350 to -3200ms relative to the onset of the target verb, before the first word of the sentence. Source estimation followed the exact procedure as described for Experiment 1. The inverse operator was calculated based on the covariance matrix from the 150 ms pre-stimulus baseline period of the cleaned trials.

3.2.5. Statistical Analyses

3.5.5.1. Behavioral data:

Overall accuracy per participant was based on responses to all 360 items. We also calculated the accuracy of the subset of task items (40%) probing the certainty of the target utterances.

<u>3.5.5.1.1. MEG data:</u>

In order to compare our results from Experiment 1 and 2, we conducted two analyses: an ROI analysis using the regions of interest as defined for Experiment 1 and a conceptual replication analysis searching for spatiotemporal clusters within a predefined region and time window based on the putative discourse updating effect of Experiment 1.

<u>3.5.5.1.2. ROI Analysis:</u>

We used the same ROIs as used for the analysis of Experiment 1, again assessing the effect of our experimental manipulations with a cluster-based permutation test (Maris & Oostenveld, 2007). We performed a temporal cluster-based permutation mass univariate 3 X 3 ANOVA with factors SENTENCE TYPE and VERB. We based our analysis time window on the results of Experiment 1, using a 150-400 ms time window after the target verb's onset to replicate the effect found in the first experiment. Again, we equalized trial count across conditions. The number of trials per condition that were analyzed was on average 36 out of 40 for NY data (ranging from 31-38 per participant) and 38 out of 40 for the AD data (ranging from 34-40 per participant).

Our temporal permutation clustering test was performed with the same procedure as laid out for Experiment 1 and corrected for comparisons across multiple ROIs (Benjamini & Hochberg, 1995).

3.5.5.1.3. Conceptual Replication Analysis:

With the expectation of replicating the results from Experiment 1, we limited our analysis to the factual sentence type condition. Then, we performed a spatiotemporal clustering analysis using the same procedure and settings as Experiment 1. Informed by the results of Experiment 1, instead of searching through the whole brain, the spatiotemporal analysis was now constrained to a predefined parcellation that combined regions in which we detected the effects of modal force in Experiment 1. This region of interest combined the right banks of superior temporal sulcus and right superior parietal, supramarginal, superior temporal, inferior parietal and middle temporal gyri from the Freesurfer aparc parcellation. Like the ROI analysis, the time window of interest was 150-400 ms after the verb's onset.

4. RESULTS

4.1. Experiment 1

4.1.1 Behavioral Results

The mean overall accuracy for the story congruency task was 83.1% (SD = .05), ranging from 71.6%-92.5% across participants. The accuracy of the one third of the congruency task items that tapped into modality was 73.3% (SD = .08) ranging from 60.0 - 88.8% across participants, and was substantially lower than the accuracy of the other general items, which was 87.9% (SD = .05) ranging from 74.4 - 94.4% across participants.

4.1.2. ROI Results

We ran a 2 (MODAL BASE: knowledge-based, rule-based) by 3 (MODAL FORCE: possibility, necessity, factual) within-subjects temporal ANOVA for the ROIs specified for Experiment 1. Since *may* and *must* differ in their lexical frequency across modal bases (*may* is high frequency as knowledge-based modal and low frequency as rule-based modal, *must* low frequency as knowledge-based modal and high frequency as rule-based modal, see 'Stimuli') we only report results that show consistent results across the *force* manipulation (knowledge-based and rule-based *may* or *must* patterning together) or the *modal base* manipulation (*may* and *must* patterning together).

The ANOVA revealed a significant effect of modal force in the right Inferior Parietal Sulcus (rIPS) within our test window of 100-900 ms after the target verb's onset (p = .046), where the factual condition (*do*) elicited more activation than the modal (*may* and *must*) conditions. This temporal cluster extended from approximately 280-340 ms. We observed a similar effect in a temporal cluster in the right Temporo-parietal Junction (rTPJ) around 240-275 ms, although this effect only survived multiple comparisons correction across time, not across multiple regions of interest (uncorrected p = .054, p = .13). Additionally, we found a trending effect of modal force in the right rostral Anterior Cingulate Cortex (rrACC), with increased activation for the necessity modal *must* over the other conditions (uncorrected p = .008, p = .099). We did not observe any other clusters in the remaining ROIs of the right hemisphere and did not observe any clusters in the left hemisphere. We summarized the ROI results in Figure 2.5 by depicting the activation patterns of the detected reliable clusters. The measured activity for each of the ROIs over our time window of interest are displayed in Figure 2.6.

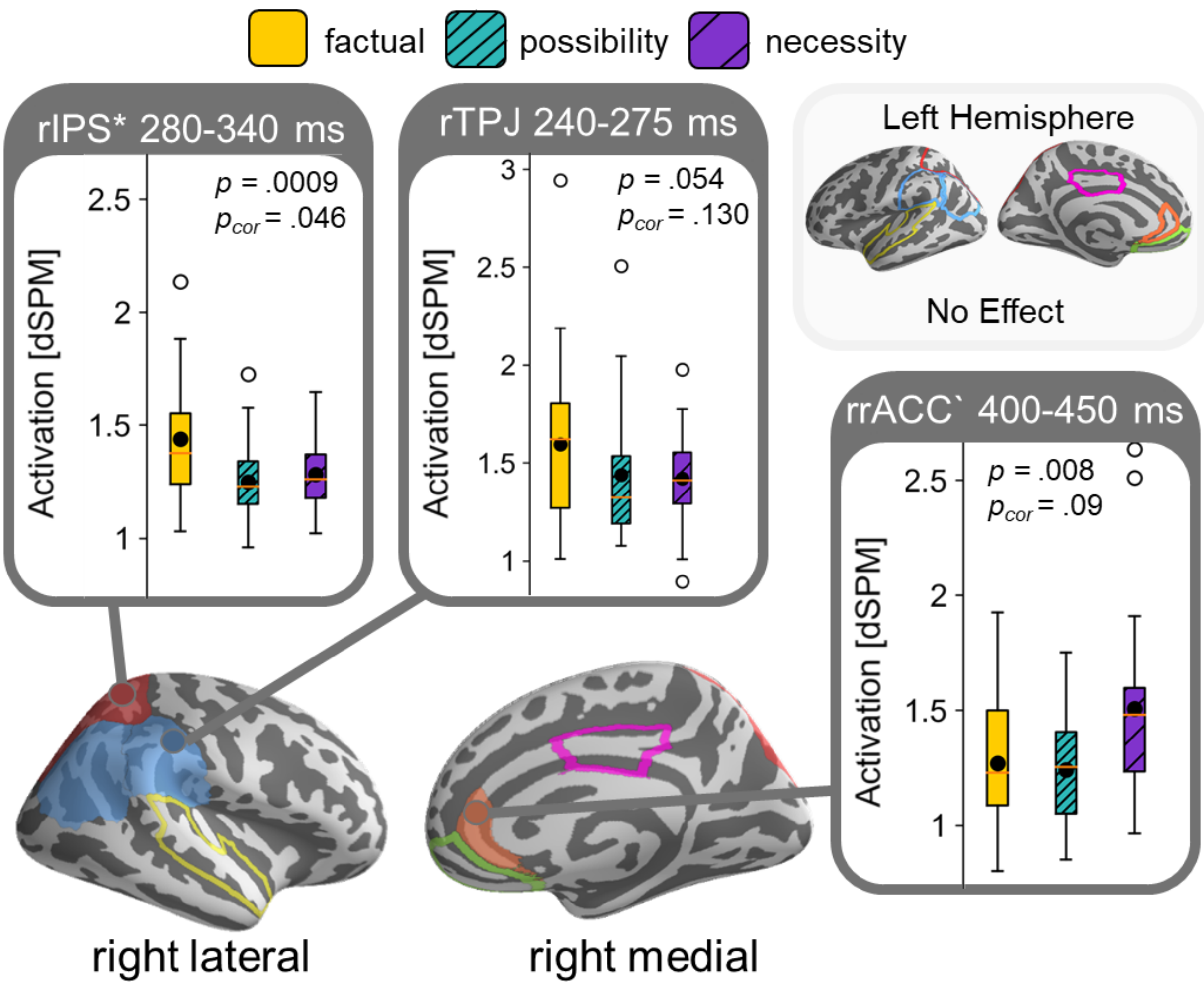

Figure 2.5. Summary Region of Interest (ROI) Results Experiment 1 showing a main effect for factual over modal conditions in right IPS and TPJ, and an increase in activation for necessity in the rrACC. Results are collapsed for MODAL BASE (knowledge-based and rule-based modals grouped together). Boxplots display estimated brain activity within the time window of the identified temporal clusters, black dots indicate mean activity. Regions of interest are outlined on brain and shaded when containing identified clusters. Clusters significant after correction comparison across multiple ROIs indicated with asterisk and with grave accent when trending.

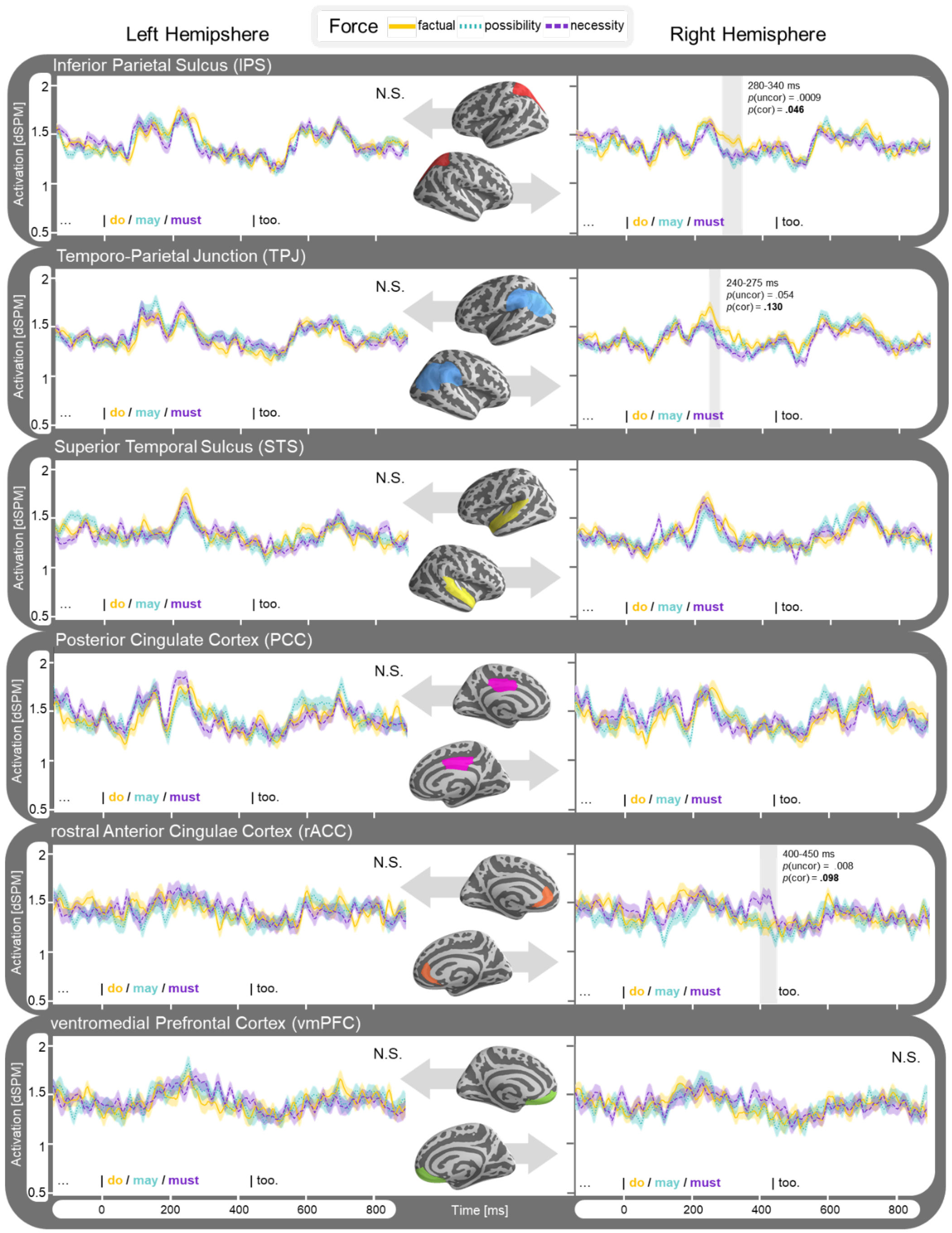

Figure 2.6. Time course of estimated average activity [dSPM] per ROI of Experiment 1. Left hemisphere ROIs displayed on the left side, and right hemisphere on the right. Results collapsed for MODAL BASE (knowledge-based and rule-based modals grouped together). Detected clusters within time window 100-900ms are highlighted and significance is indicated for the effect within the cluster (p_{uncor}) and when corrected for comparison across multiple regions (p_{cor}).

4.1.3. Spatiotemporal Results (Whole Brain)

A full-brain analysis revealed a significant effect for modal force, eliciting stronger activity for our factual condition over our modal conditions (p= .033) in our 100-900 ms time window. We detected a cluster between approximately 210-350 ms centering around the right Temporoparietal Junction (rTPJ) extending posteriorly over to the right Intraparietal Sulcus (rIPS) to the medial cortex, covering the cuneus, parts of the precuneus, and ending in the posterior cingulate cortex (Figure 2.7). The activation in this cluster reflects the activity we found for the effect of modal force in the rIPS and rTPJ of our ROI-analysis. No other significant clusters were found.

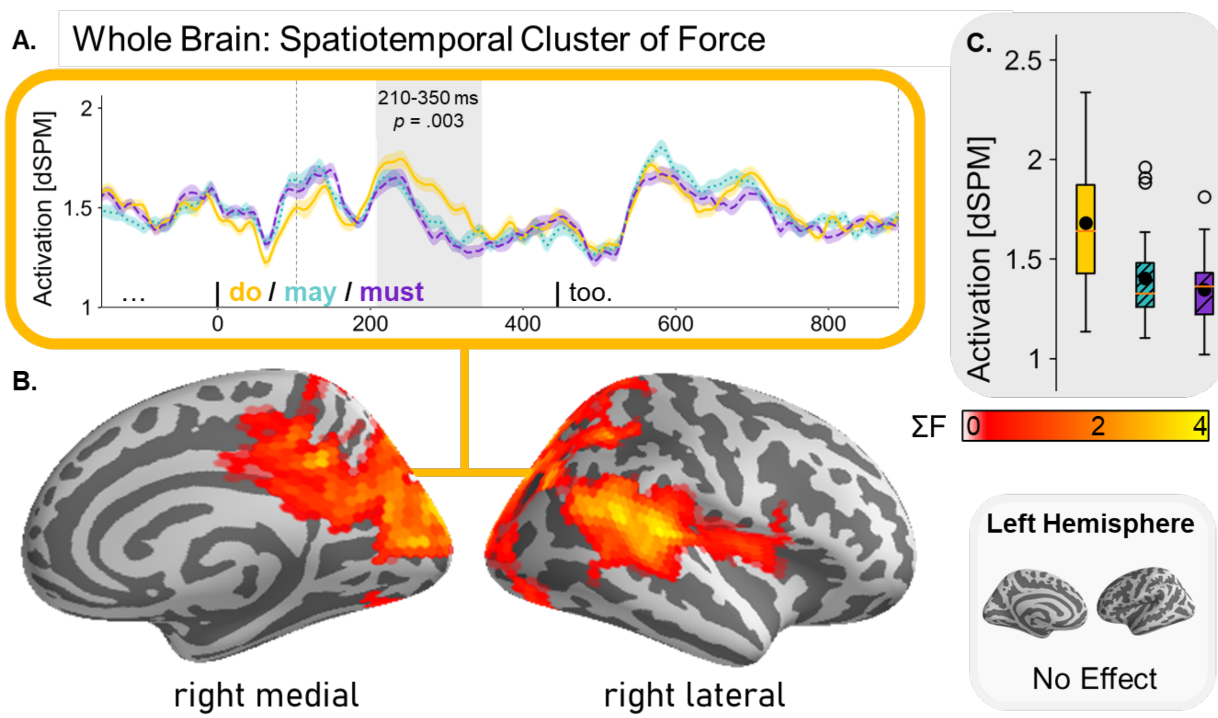

Figure 2.7. Identified spatiotemporal cluster of whole-brain analysis Experiment 1. **A.** Time course estimated brain activity [dSPM] and identified cluster (in grey). Boundaries of analysis window (100-900 ms) are indicated by dashed lines. **B.** FreeSurfer average brain shows spatial distribution of cluster, color shading indicating the sum of cluster-level F statistic (gained from cluster-based permutation test). **C.** Boxplots display estimated brain activity (factual > modal) within the identified time window of the spatiotemporal cluster, black dots indicate mean activity.

4.2. Experiment 2

4.2.1. Behavioral Results

The mean overall accuracy for the conclusion validation task was 85.6% (SD=.09), ranging from 64.7%-96.9% across participants. The accuracy of the subset of the validation task items that tapped into certainty was 83.7% (SD = .10) ranging from 57.6 - 95.2% across participants.

4.2.2. ROI Results

We ran a 3 (SENTENCE TYPE: factual, conditional, presupposed) by 3 (VERB: may, might, do) within-subjects temporal ANOVA for the same ROIs specified for Experiment 1. We only observed effects that survived multiple comparisons correction across time, but not across multiple regions of interest. The ANOVA revealed an interaction effect of VERB and SENTENCE TYPE in the left rostral Anterior Cingulate Cortex (lrACC) within our test window of 150-400 ms after the target verb's onset (uncorrected $p = .034$, $p = .341$), where the factual condition (*do*) elicited more activation than the modal (*may* and *must*) conditions in factual sentences, but not in conditional or presupposed sentences. In fact, in presupposed sentences the factual condition elicited less activity than the modal conditions. The temporal cluster reflecting this activity difference extended from approximately 365-395 ms. We observed a similar effect in a temporal cluster in the right ventromedial Prefrontal Cortex (rvMPFC) around 345-370 ms (uncorrected $p = .032$, $p = .327$). No other clusters were detected in any of the other regions of interest. We summarized the ROI results in Figure 2.8 by depicting the time course of the detected reliable clusters. The effect in the lrACC was most prominent in the NY data while the effect in the rvMPFC was more prominent in the AD data (Appendix S2.2). The measured activity for each of the ROIs over our time window of interest in the factual sentential context (for comparison with Figure 2.6) is displayed in Figure 2.9.

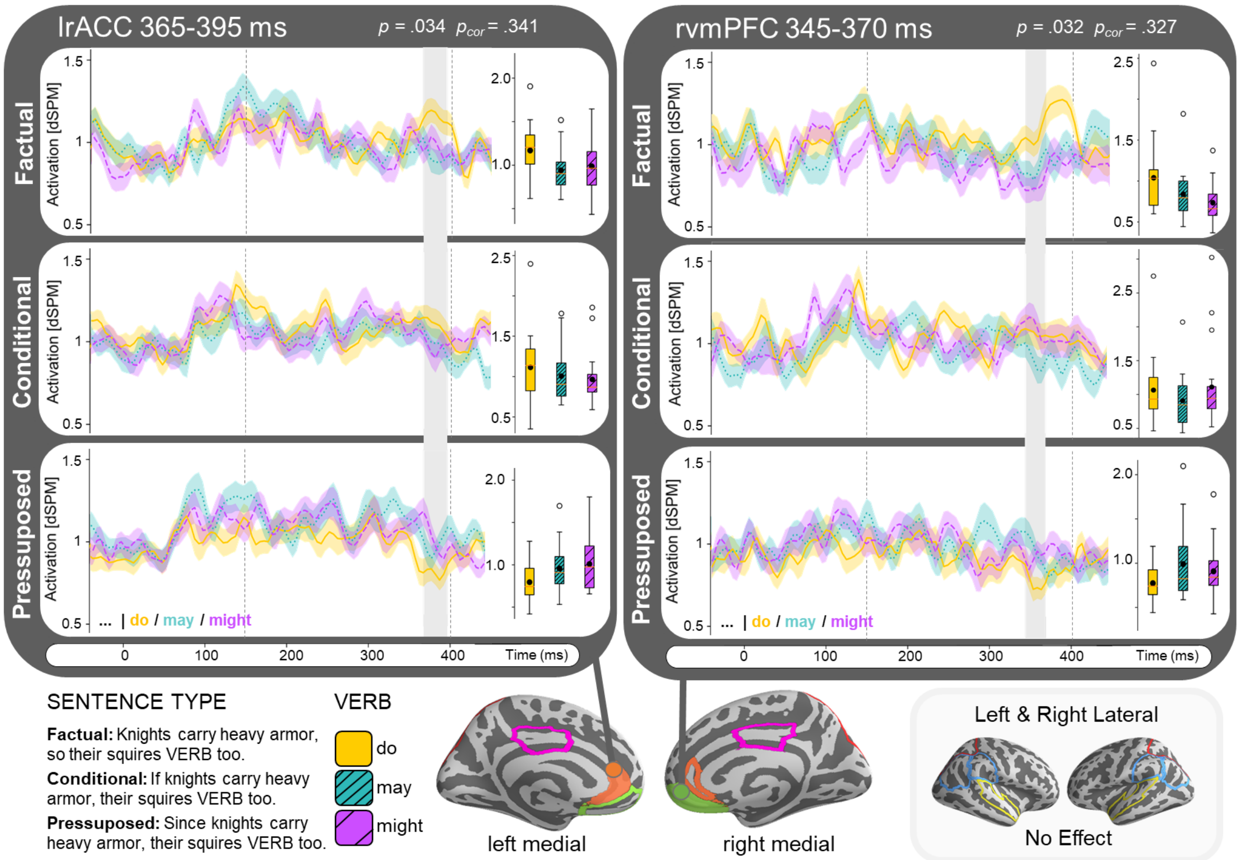

Figure 2.8. Time course estimated brain activity [dSPM] of reliable detected clusters from ROI analysis Experiment 2. Both the lrACC and rvmPFC show an interaction between sentence type (factual, conditional and presupposed) and verb (*do*, *may* or *might*) with increased activation for *do* > *may/might* when embedded in factual sentences, and decreased activation for *do* < *may/might* in presupposed sentences. Boundaries of the analysis window (150-400 ms) are indicated by dashed lines, identified clusters displayed in grey. Boxplots display estimated brain activity within the time window of the identified temporal clusters, black dots indicate mean activity. Regions of interest are outlined on brain and shaded when containing identified clusters. Cluster effects are not significant after correction comparison across multiple regions of interest. The effect in the lrACC was most prominent in the NY data while the effect in the rvMPFC was more prominent in the AD data (Appendix S2.2).

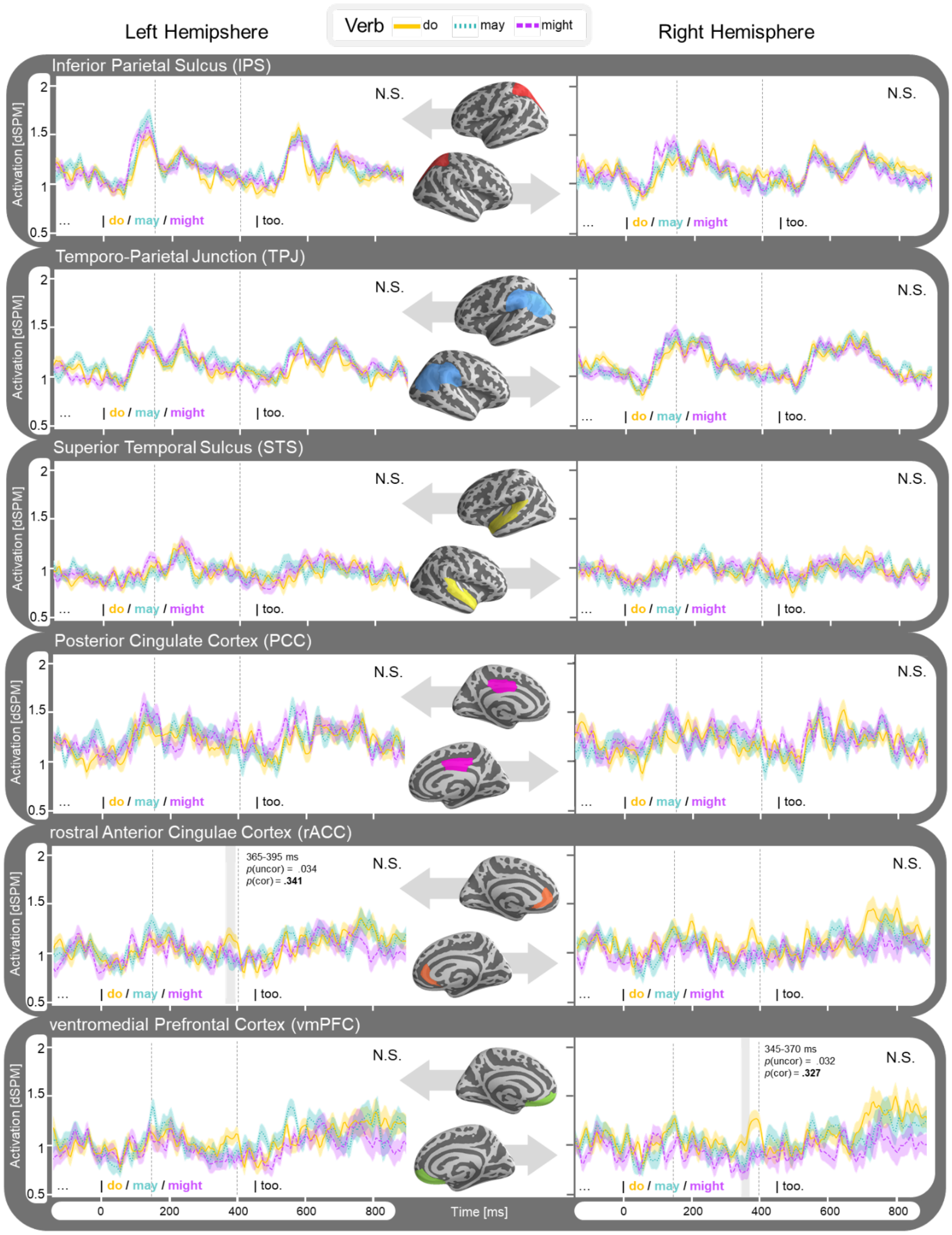

Figure 2.9. Time course of estimated average activity [dSPM] per ROI of Experiment 2 for factual sentence type (*p* so *q*). Left hemisphere ROIs displayed on the left side, and right hemisphere on the right. Results collapsed for MODAL BASE (knowledge-based and rule-based modals grouped together). Detected clusters within time window 150-400ms (indicated with dashed lines) are highlighted and significance is indicated for the effect within the cluster (p_{uncor}) and when corrected for comparison across multiple regions (p_{cor}).

4.2.3. Conceptual Replication Results

We performed a spatiotemporal clustering test in the time window 150-400 ms in a region of interest covering right lateral temporoparietal areas aiming to replicate the effect found in Experiment 1. Unlike the results of Experiment 1, a one-way ANOVA comparing activity within the VERB condition (*do*, *may* and *might*) in FACTUAL sentences detected no significant clusters in this area. This corroborates the results of the ROI analysis, in which we similarly found no difference in activity between the factual and modal verbs in the right IPS, TPJ or STS.

5. DISCUSSION

In this work, we conducted two experiments to explore the neural correlates of modal displacement and discourse model updating during language comprehension. During natural discourse comprehension, the comprehender does not only integrate incoming factual information into an evolving discourse model, but also entertains hypothetical situations denoted with modal utterances. We investigated how the brain distinguishes between factual and modal information.

Our stimuli contained short scenarios with two parts. The first part of the narrative established some property or habit that applied to one entity (e.g., “Knights carry heavy armor”), The second provided additional information about a second entity that was either factual (e.g., “the squires *do* too”) or modal (e.g., “the squires *may*/*must*/*might* too”). While the factual utterances indicated an actual change in situation, requiring the discourse representation to be updated, the modal utterances merely indicated a possible (uncertain) change. Our data showed that the factual condition elicited reliably stronger activation than the modal condition in right temporoparietal (Experiment 1) and medial frontal regions (Experiment 2). Below we discuss these increases as possible neural correlates of discourse model updating, elicited in the presence of updates that are certain (factual) but not for updates that are uncertain (modal).

5.1. Neural Correlates of Discourse Updating

Discourse updating, the operation of updating the mental representation of a situation, was modelled here as the attribution of a property to a new entity. Prior behavioral research has shown that mental representations of discourse are dynamically updated when presented with new facts (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004). Such modal updating has been associated with increased activation in the mPFC, PCC and temporo-parietal areas (Ferstl et al., 2005; Fletcher et al., 1995; Speer et al., 2007; Xu et al., 2005; Yarkoni et al., 2008). In Experiment 1, we found an increase in source-localized MEG responses for factual over modal statements. Specifically, activity increased in factual statements in the right lateral temporal and parietal hemisphere at approximately 200-350 ms after target verb onset. This effect was most pronounced in the right inferior parietal sulcus (rIPS) and less so in the right temporo-parietal junction (rTPJ). This pattern of activity is compatible with behavioral findings on discourse updating. Factual utterances signal an actual change in the discourse, and when this information is incorporated into the comprehender's mental representation this results in increased brain activity. In contrast, modal utterances only indicate a possible change of situation. Since the update is uncertain, situation model updating does not take place.

In Experiment 2, we manipulated the broader sentential context in which novel factual and modal information was presented. In contrast to Experiment 1, where the target sentence always built on a certain factual base, we now also presented the target utterance in conditionals that were hypothetical (uncertain, i.e., "If knights carry large swords…") or presupposed (presumed to be common knowledge, i.e., "Since knights carry large swords…"). We expected discourse updating to only take place when the situational change is certain, and that embedding a factual update into a hypothetical conditional should prevent discourse updating from taking place due to the entire scenario being uncertain (Figure 2.1).

While Experiment 2 was designed to replicate the results from Experiment 1 with our factual sentential context, we instead found that this time our ROI analysis (using the same regions of interest as defined for Experiment 1) revealed no differences in activity between factual and modal utterances in the right lateral hemisphere. This was confirmed by a replication analysis searching for spatiotemporal clusters targeting right lateral temporoparietal areas within the time window of 150-400 ms. Instead, we now found increased activity for factual over modal conditions in a temporal cluster in two adjacent areas: the left rostral Anterior Cingulate Cortex (lrACC) and right ventromedial Prefrontal Cortex (rvmPFC) within our test window of 150-400 ms after the target verb's onset. This effect only survived multiple comparisons correction across time, not across multiple regions of interest. The hypothesis that this activation reflects discourse updating gains weight from the fact that we only observed this pattern of activity when the sentential context was factual ("Knights carry large swords, so their squires *do*/*may*/*might* too.") but not when the sentential context was hypothetical ("If knights carry large swords, their squires *do*/*may*/*might* too."). This would be in line with the idea that discourse model updating only takes place under certain situational changes, though such a conclusion has to be drawn with caution, as the results of Experiment 2 were not that robust.

This presumed discourse updating effect resonates with prior behavioral studies on discourse updating and situation model maintenance. Discourse models representing a situation are dynamically updated as novel information indicating a change of situation comes along. As a consequence of model updating, 'old' information that is no longer relevant to the *here-and-now* of a story is backgrounded, which is measurable in longer retrieval times in probe-recognition tasks compared to information that is still relevant to the current situation (Glenberg et al., 1987; Morrow et al., 1989; Zwaan & Madden, 2004). De Vega et al. (2012; 2007) investigated whether this model updating also takes place when integrating hypothetical information, comparing accessibility after encountering factual ("As he had enough time, he went to the café to drink a beer") and counterfactual utterance ("If he had enough time, he would have gone to the café to drink a beer"). De Vega et al. (2007) found evidence for discourse updating when integrating factual information but not for counterfactual information, leading them to conclude that the hypothetical meaning of counterfactuals does not contribute to the build-up of the discourse representation. This finding was corroborated in an ERP study, where increased negativity after factual compared to counterfactual continuation utterances and reduced gamma power following counterfactuals were taken to indicate that the counterfactual's 'as if' meaning is not integrated into the discourse (de Vega & Urrutia, 2012). Our results likewise suggest that mental model updating takes place for the integration of novel factual information, but not for hypothetical information as indicated by modality (*may/must/might*) or conditionality (*if*…).

This immediate sensitivity to the factual (*do*) versus hypothetical (*may/must*) contrast is in line with ERP findings showing rapid integration of contextual information in online processing. Prior context modulates the N400 component such that it takes more effort to retrieve lexical items compatible with the actual world in counterfactual utterances (where non-actual information is expected) than in factual or hypothetical utterances (Kulakova & Nieuwland, 2016a; Nieuwland & Martin, 2012). Similarly, factive verbs like *know* presuppose complements compatible with the actual world, and when this expectation is violated it gives rise to P600 effects, taken to reflect conflict detection (Shetreet et al., 2019). While these ERP studies confirm that the brain is sensitive to the factual/hypothetical contrast during online processing, our results shed more light on when this information becomes available, possibly as soon as ~200ms after the target's verb onset.

While the results of Experiment 2 are less strong, they address some possible alternative explanations for the robust effect observed in Experiment 1, which we hypothesized to reflect discourse updating. One might wonder whether a more low-level explanation could explain the observed activity increase for *do* over *may* and *must* in the first experiment, such as an inherent difference in lexical frequency (*do* is more frequent than *may* and *must*), polysemy (*may* and *must* are polysemous while *do* is not) or type of ellipsis (*do* ellipsis syntax may differ slightly from *may*/*must*). These alternative explanations are contradicted by the results of Experiment 2, as we would have expected low-level effects like these to have been replicated in the same location and be insensitive to the experimental manipulation of our sentential context. Furthermore, we included the non-polysemous modal *might* to rule out the polysemy hypothesis. If the increase of factual over modal conditions in both experiments reflects discourse updating however, the question arises what caused the shift in location of this effect between experiments.

5.2. Updating the Representation of Someone Else's Mental State versus One's Own

In both of our experiments, we observed an increase for factual over modal expressions – henceforth "updating effect" – but the effect localized differently across the two experiments. In Experiment 1, the updating effect was found in the rIPS and the adjacent rTPJ, while in Experiment 2 we did not observe any effects in these specific areas. Instead, Experiment 2 elicited a similar pattern of activity in medial frontal areas: the lrACC and rvmPFC. Both frontal medial and temporal parietal areas have been found to be involved in constructing and maintaining discourse representations in fMRI studies (Ezzyat & Davachi, 2011; Friese et al., 2008; Speer et al., 2007; Xu et al., 2005; Yarkoni et al., 2008). For example, Xu et al., (2005) investigated natural language comprehension at the level of words, sentences and narratives. When comparing visually presented isolated sentences and narratives, they observed robust response increases in several bilateral brain regions including the precuneus, medial prefrontal and dorsal temporo-parieto-occipital cortices. In a similar manipulation, contrasting unrelated sentences with coherent narratives, Yarkoni et al. (2008) found narrative-specific activation in the mPFC and additional neural contributions of posterior parietal regions supporting situation model construction and frontotemporal regions supporting situation model maintenance.

While both temporoparietal and frontal medial areas are part of the network engaged during narrative comprehension, one may wonder why Experiment 2 did not replicate the discourse updating effect of Experiment 1 in the same regions. The reason for this may be related to a change in materials between the experiments, altering whose mental representation is updated. In Experiment 1, all target beliefs are attributed to a third person character, e.g., "But the king learns that the squires do too". This third person character was included to enhance the contrast between the knowledge-based and rule-based modal readings, varying between authority and observer figures respectively. In contrast, Experiment 2 lacked this third person character and embedding verb ("…, so the squires do too") for the target manipulation to appear in conditional structures. By making this change in stimuli, we inadvertently changed whose mental state is updated during comprehension, someone else's (Experiment 1) or the participant's own (Experiment 2). When we represent someone else's beliefs, we separate these from our own, as is evident from our ability to attribute false beliefs. For example, in the Introduction our example narrative contained the utterance "Pyramus quickly concludes she must have been devoured by the beast", which allowed us to understand Pyramus thinks that his lover has died, even though we know from the prior context that she is still alive. Theory of Mind encompasses the ability to represent someone else's mental state separate from our own (Premack & Woodruff, 1978). Theory of Mind reasoning engages a network of brain regions, but it has been argued that particularly the right TPJ is involved in representing the mental state of others (Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005; Vistoli et al., 2011) or reorienting attention (Corbetta et al., 2008; Decety & Lamm, 2007; Mitchell, 2008; Rothmayr et al., 2011). We tentatively suggest that the discourse updating effect in Experiment 1 localized around the right TPJ because it involved updating a discourse representation separate from the comprehender's own. Experiment 2 involved updating one's own global representation and elicited activation in frontal medial regions. This is in line with studies finding medial prefrontal activity for tasks that require people to reflect on or introspect about their own mental states (Gusnard et al., 2001; Mitchell et al., 2005; Zhu et al., 2007). And is also compatible with Ezzyat and Davachi (2011), who found that the bilateral vmPFC seemed especially engaged when integrating information within events, suggesting that this region could be sensitive to discourse updating.

Alternatively, it could be the case that the difference in results between Experiment 1 and Experiment 2 has to do with the different methods of contextualizing the target utterance. In Experiment 1, the target sentence appeared after an initial context sentence that was read at the participant's own pace. In Experiment 2, the context before the target utterance merely consisted of one word introducing the general setting of the following utterance. While one may wonder whether these differences in context complexity (sentence versus word) and processing pace (self-paced versus timed) interfered with the baseline of the trial, it seems unlikely that this would be the cause for different results between Experiment 1 and 2. Since all conditions within the experiments uses the same baseline region, one would expect that any artifacts resulting from task effects is consistent across the different conditions of the experiments. Since we only compare conditions within experiments, the presence of an effect relative to other conditions cannot be due to a baseline effect (e.g., pressing a button). A more pressing question is whether the differences between the results of Experiment 1 and 2 can be attributed to varying narrative complexity. In Experiment 1, the (self-paced) context sentence established a property for one entity, and the target utterance then indicated that this property was also (possibly) shared by a second entity. In Experiment 2, the target utterance consisted of two clauses, the first one establishing a (possible) property for one entity, while the second one stated that this property was (possibly) shared by a second entity. The entire target utterance was displayed with rapid serial visual presentation. Compared to Experiment 1, Experiment 2 thus allowed less time for participants to appreciate the initial situation (property being attributed to one entity) before updating this information (property also being attributed to second entity). An alternative explanation for our results could be that temporal parietal areas are more involved with constructing a larger discourse representation (coherence between sentences), while the medial frontal areas are more involved with initializing a discourse representation. This would be in line with Xu et al. (2005), who observed increased activity in the right hemisphere as contextual complexity increased.

An argument against this alternative hypothesis comes from recent work by Jacoby and Fedorenko (2018) investigating the neural correlates of expository discourse comprehension. While prior studies detected right temporal parietal engagement in comprehension of narratives (stories built around characters), expository texts (constituting facts about the real world) elicited no effect of discourse coherency in posterior ToM regions like the rTPJ (Jacoby & Fedorenko, 2020). This suggests that these regions only engage in coherence building for discourse in which you take someone else's perspective. However, Jacoby and Fedorenko (2018) did find that the mPFC was sensitive to discourse coherency of expository texts. Since their expository texts were as complex as a narrative, it cannot be the case that the lack of engagement of the rTPJ observed for expository texts is due to a lack of discourse complexity. At the same time, the finding that the mPFC is sensitive to the coherence of expository texts suggests it could be involved in updating one's own discourse beliefs.

5.3. Neural Correlates of Modal Displacement?

Before, we defined 'modal displacement' as an operation that shifts our perspective from the immediate present to a hypothetical scenario. Several prior studies have investigated the neural correlates of utterances that involve hypothetical situations, but, as far as we know, no study has succeeded in isolating the neural mechanisms involved with the operation of modal displacement. Dwivedi et al. (2006) observed stronger responses for modal utterances (“it might end quite abruptly”) compared to factual utterances (“it ends quite abruptly”), and speculated this activity increase reflects the cost of mentally representing and comparing multiple possibilities. However, their study was not controlled for utterance length or complexity, leaving uncertain whether their observed activity increases were really due to the experimental manipulation. Another branch of neurolinguistic studies that investigates hypothetical meaning is research on the processing of counterfactuality, which engages parts of the default mode network such as the medial frontal and temporal lobes, the posterior cingulate cortex, precuneus, and the lateral parietal and temporal lobes (De Brigard et al., 2013; Kulakova et al., 2013; Nieuwland, 2012; Urrutia et al., 2012; see Van Hoeck et al., 2015 for recent overview). Like modal constructions (e.g., “The monster might be big”), counterfactuals posit a hypothetical scenario (e.g., “If the monster were big…”). Unlike modal utterances, though, counterfactuals do not leave open any uncertainty about the actual state of affairs, rather they imply that the opposite is true (the monster is not big). On top of displacing from the here and now, the processing of counterfactual constructions involves keeping in mind two conflicting representations and inferencing the actual state of affairs. Any comparison between factual and counterfactual utterances (e.g., Urrutia et al., 2012) cannot separate these distinct processes.

Our study investigated modal displacement by minimally comparing factual and modal utterances. We found no reliable increases in neural activity when modal displacement occurred. However, the fact that we did find neural activation dissociating between the factual and modal condition suggests that participants processed the modal items as being different from the factual ones. Given that the increase in activation of factual over modal conditions takes place during the discourse integration of information indicating an actual change in situation, but not when integrating information regarding an uncertain (hypothetical) change, the most likely interpretation of our data is that this difference in activation reflects discourse updating.

However, if non-factual information does not get integrated into an existing situation model, the question remains how we *do* represent this information. The theoretical background for the current study was that modal displacement would involve the generation of multiple possibilities (von Fintel, 2006; Iatridou, 2000; Johnson-Laird, 1994; Kratzer, 2012). Intuitively, this would suggest that when presented with uncertainty, the comprehender postulates multiple mental representations of these different possibilities, the minimal one being a negated version (if squires *might* sit at round tables, this introduces the alternative possibility that maybe they do not). Considering multiple possibilities in parallel is thought to be cognitively demanding (Leahy & Carey, 2019), and we thus expected additional activity related to this operation. It is possible that this assumption was wrong, and that for example, the decreased activity for modal utterances compared to factual utterance is indicative of modal displacement rather than discourse updating. However, it is difficult to gauge why this modal displacement is dependent on the sentential context and why we would find this correlate shifting in location across experiments. Alternatively, there might not be any correlates of representing multiple possibilities in the cortex at the level we investigated in this paper. Recently, Kay et al. (2020) found that possibility generation in rats involves a constant cycling between possible future scenarios in hippocampal neuron populations. At a constant cycling of 8 Hz the cells alternated between encoding two different possible futures.

The authors suggest this finding might extend to the representation of hypothetical possibilities in human brains, possibly extending to brain regions connected to the hippocampus.

Lastly, some have proposed that the representation of modality involves marking a representation with a symbolic operator, indicating that this representation can be neither ruled out nor added into the actual model (Leahy & Carey, 2019). This theory would not require people to actively postulate alternative situations, though the question remains how this uncertain information would be maintained and linked to the prior discourse if not incorporated into the existing situation model. For now, these questions are still open to future exploration.

5.4. No Effect of Modal Base and Force

Our stimuli in Experiment 1 were carefully designed to investigate the online comprehension of modal verbs varying in modal base (knowledge-based versus rule-based) and force (possibility versus necessity). However, we found no reliable effects of these manipulations. We did find an effect in the right rostral Anterior Cingulate Cortex showing increased activation for necessity modals over the other conditions (Figure 2.6), but this effect only survived multiple comparisons correction across time, not across multiple regions of interest. The rostral ACC is, besides its involvement in ToM tasks, also argued to be involved in error processing and conflict resolution (Dreher & Grafman, 2003; Kiehl et al., 2000), suggesting that our effect may reflect some unnaturalness in our stimuli. The verb *must* requires strong evidence, but the surrounding context was made to be also compatible with weaker evidence (to allow for the appearance of *may*). Possibly, our stimuli contained too little evidence to naturally say *must*, eliciting increased activation in the rrACC when resolving this conflict.

6. CONCLUSION

This work investigated the integration of factual and modal information into short narratives. While the factual utterances indicated an actual change in situation, requiring the discourse representation to be updated, the modal utterances merely indicated a possible (uncertain) change as these utterances displaced from the narrative's here-and-now. In a controlled within-subjects design, we measured source-localized MEG responses while participants integrated modal and factual information into a short narrative. While we did not find any regions of the brain more engaged by the modal conditions over the factual conditions (which could reflect engagement with modal displacement), we did find the opposite pattern of activation where certain brain regions elicited stronger activation for the factual over the modal condition. This increase in activation may be a neural correlate of mental discourse representation updating. This activity difference seems to go away as soon as the factual update is presented in an uncertain (conditional) sentential environment, supporting the idea that discourse updating only takes place when the change in the situation is certain. To our knowledge, this was the first attempt to explore the neural bases of modal processing. While we have established possible neural correlates of fact comprehension, the question of how uncertain information is integrated into a discourse representation remains open. We hope that our work establishes a starting point for further investigations of this phenomenon.

CHAPTER 3: WISHES BEFORE IFS: MAPPING "FAKE" PAST TENSE TO COUNTERFACTUALITY IN WISHES AND CONDITIONALS

As submitted in Maxime Tulling & Ailís Cournane (under review)

1. ABSTRACT

Counterfactuals express alternatives that are contrary to the actual situation. In English, counterfactuality is conveyed through conditionals ("If pigs had wings, they could fly") and *wish*-constructions ("I wish pigs had wings"), where the past tense morpheme marks non-actuality rather than past temporal orientation. This temporal mismatch seemingly complicates the already challenging task of mapping abstract counterfactual meaning onto these linguistic expressions during first language acquisition. We conducted a corpus study on children's spontaneous productions of counterfactual constructions to investigate whether children make productive tense errors in counterfactuals, and gain insight into the relative acquisition onset of the counterfactual wish and conditional construction. We extracted *wish*-utterances from 52 English-speaking corpora available on CHILDES to compare children's wish productions with those of adults, and additionally extracted counterfactual conditional utterances for 6 children to provide a comparative longitudinal overview of counterfactual wishes and conditionals. Results show that *wish*-constructions are generally acquired before counterfactual conditionals, as most first wishes are produced around age 2 or 3 while the conditionals emerge at age 3 or 4. Additionally, we found that children make the productive error of producing counterfactuals with present tense marking instead of past. These errors are consistent with a stage where children have yet to figure out that counterfactual past tense signals a present non-actuality, rather than a past event on the timeline.

2. INTRODUCTION

Counterfactual expressions such as “If pigs had wings, they could fly”, express alternatives that are contrary to the actual state of affairs (pigs do not have wings). Prior research shows that children generally start producing and comprehending counterfactual conditionals around age 4, after they have developed the ability to refer to hypothetical future events, which already seem to be in place by age 3 (Bowerman, 1986; Guajardo et al., 2009; Nyhout & Ganea, 2019; Reilly, 1982, Rouvoli et al., 2019). This asymmetry between the acquisition onset of the hypothetical future and counterfactual has mainly been attributed to the additional cognitive load demanded by counterfactual reasoning, which depends not only on holding multiple possibilities in mind, but also requires temporarily considering a false possibility as true (Beck et al., 2009; Byrne, 2007). However, prior studies have not fully addressed the role of linguistic complexity in the acquisition of counterfactual constructions. More specifically, they have not addressed the challenges posed by mapping linguistics forms to counterfactual meaning. The present study demonstrates the importance of considering form-to-meaning mapping in the acquisition of abstract concepts such as counterfactuals and provides empirical evidence that more linguistically transparent constructions (counterfactual wishes) are used before the more complex ones (counterfactual conditionals). Studies solely focusing on the acquisition of counterfactual conditionals might thus have underestimated children’s ability to engage in counterfactual reasoning, confounding cognitive with linguistic complexity.

2.1. Background: Development of Children’s Counterfactual Productions

Mirroring findings on children’s comprehension (Riggs et al., 1998; Robinson & Beck, 2000), production studies have reported that future hypothetical conditionals (a conditional construction about a future possibility such as "If it rains tomorrow, we will play inside") are acquired before counterfactual conditionals (Bowerman, 1986; Reilly, 1982). Reilly (1982) used longitudinal recordings and diary entries about one child (Kate), and various elicitation tasks with children between age 2-8. She found that most children produce hypothetical conditionals by age 3, but do not yet fully comprehend hypothetical conditionals by this age and do not seem to understand counterfactuals. In fact, when asked counterfactual "what if" questions, many 2-year-olds and quite some 3-year-olds deny the counterfactuals or respond to them as if they are about reality, see (1a) and (b):

(1a) Adult: What if a snake had eaten your daddy? (Reilly, 1982, ex. 37 p.107)
Cate (2;8): No! / Can't eat my daddy

(1b) Adult: What if you were a snake? (Reilly, 1982, ex. 57 p.116)
Janine (3;0): I'm not a snake / I'm Janine.

At age 4, Reilly (1982) found that children demonstrated comprehension of both hypotheticals and counterfactuals. They no longer denied the possibility of a situation or gave realist replies to counterfactual utterances. They also produced clear spontaneous present counterfactual conditionals (2).

(2) 4-yo: If they put a goldfish in there and they ate it, they would die.
(Reilly, 1982, ex. 68, p.121)

Kuczaj & Daly (1979) investigated the longitudinal development of Abe and did a cross-sectional study of 14 other children. They similarly found that future hypothetical conditionals seem to be acquired before counterfactual conditionals, and report that Abe used his first past counterfactual conditional at 3;11.

2.2. Mapping Challenge: Attributing Counterfactual Meaning to the "Fake" Past Tense

Children appear to acquire counterfactual conditionals relatively late compared to hypothetical conditional constructions. What makes counterfactuals more complex? To acquire an abstract linguistic construction that involves complex cognitive processes, such as counterfactual reasoning, two criteria need to be fulfilled: 1) the child must have developed the cognitive ability to support the mental operations involved in representing the meaning of the utterance, and 2) the child must figure out which linguistic forms are used to express such meanings in their target language(s) (Clark, 2001; Reilly, 1982). As for the cognitive factors underlying counterfactual reasoning, the development of executive functions like working memory, attention switching and inhibition have been linked to the acquisition of counterfactuality (Beck et al., 2009, 2011, p. 20; Byrne, 2007; Guajardo et al., 2009; Robinson & Beck, 2000; Weisberg & Gopnik, 2016). As for form-to-meaning mapping, there are three properties of counterfactual constructions that make this mapping particularly challenging. First, it is not obvious how children learn to map meaning onto linguistic forms when the expressed meaning is not perceptually observable (Gleitman et al., 2005; Landau & Gleitman, 1985). In the case of counterfactual constructions (e.g., "If I were you"), this is particularly true, as by definition the proposition expressed by the counterfactual is not true in the actual world, and thus cannot be observed.

Second, there is no one-to-one correspondence between form and counterfactual meaning (Clark, 1987). Counterfactuality can be mapped onto different types of linguistic expressions, such as counterfactual (CF) conditionals (3) or wishes (4) and also involves attributing more than one abstract meaning to the same past morpheme. Past tense inflection usually refers to the actual past, and thus can only combine with a temporal adverb that matches this temporal orientation (5), like *yesterday*. However, in counterfactual constructions the past morpheme gives rise to a non-actual present interpretation instead (Iatridou, 2000; Ippolito, 2006; Karawani & Zeijlstra, 2013; Romero, 2014). This mismatch between morphological tense marking (past) and temporal orientation in present counterfactuals (dubbed "fake" past tense by Iatridou, 2000), becomes evident when the "fake" past is combined with the present temporal adverb *right now* (3a/4a). Counterfactuals with true past temporal orientation require double past marking (4b/4b). This pattern is cross-linguistically robust across distinct language families (Bjorkman & Halpert, 2017; Iatridou, 2000; James, 1982; von Prince, 2017, p.6 and references therein).

(3a) If I **had** a car (right now/*yesterday), I would drive. PRESENT CF CONDITIONAL

(3b) If I **had** had a car back then, I would have driven. PAST CF CONDITIONAL

(Ritter & Wiltschko, 2014, 62)

(4a) I wish I **had** a car (right now/*yesterday). PRESENT CF WISH

(4b) I wish I **had** had a car back then. PAST CF WISH

(Iatridou, 2000, 25-26)

(5) I **had** a car (*right now/yesterday).

Besides the present and past counterfactual, there is also a third type of construction that Iatridou (2000) groups with the counterfactual constructions: the future "counterfactual"[8] or 'future less vivid' (FLV). This construction has a flavor of 'unlikeliness', in the sense that the speaker considers this future possibility less likely (or less vivid) than other possibilities. In present counterfactual conditionals, the future less vivid reading arises when the main verb of the antecedent is eventive (6a), in wishes it arises when the complement of wish contains *would* ("fake" past + will) (6b). Similar to the present and past counterfactual, the future less vivid indicates the speaker believes the opposite to be true (e.g., "I wish he would leave tomorrow" can be used when someone is scheduled to leave next week instead).

(6a) If he **left** tomorrow, he would get there next week. FUTURE LESS VIVID (FLV)

(6b) I wish he **would** leave tomorrow. (Iatridou, 2000, 28)

How do children figure out that the counterfactual past tense refers to a non-actual present rather than a past situation? The distributional learning approach to the acquisition of semantic meaning posits that the syntactic contexts in which a word (e.g., an attitude verb like *think*) appears helps learners bootstrap meanings that are not directly observable (Landau et al., 2009; Hacquard & Lidz, 2018). Relatedly, the temporal orientation of a predicate utterance is thought to help with acquiring different flavors of modal meanings, e.g., distinguishing the obligation and epistemic reading of *may* and *must* (van Dooren et al., 2017, 2019). However, in the case of counterfactuals,

[8] The reason why this construction can strictly not be called a counterfactual is because it refers to the future, and therefore in principle is still realizable. However, Iatridou (2000, p.235) raises the question of whether we maybe should be considering it a real counterfactual after all, as it patterns alike with the other constructions. In wishes, future temporal orientation seems to indicate a desire to change a future that the speaker believes to be unlikely or impossible to change, e.g., because it's planned or determined.

morphosyntactic cues seem misleading at first sight. As a third mapping challenge, acquiring counterfactual constructions requires children to see through the "fakeness" of the past tense morphology and learn to map this morpheme to a semantic operation supporting counterfactuality instead.

2.3. The Semantics of Counterfactual Constructions: Modal-past or Past-past

What are the semantic operations supporting counterfactuality? Most theoretical accounts put a lot of explanatory weight on the counterfactual's "fake" past. There are two main approaches to analyzing the semantic role of the counterfactual's past tense morpheme (Bjorkman, 2015; Karawani, 2014; Romero, 2014; Schulz, 2014; von Prince, 2017). Past-as-past (or 'back-shifting') approaches argue that the counterfactual's past tense morpheme fulfills the function of shifting back in time (Arregui, 2009; Dudman, 1983; Ippolito, 2006; Ippolito & Keyser, 2013; Ogihara, 2000; Romero, 2014). They base themselves on a framework where time is represented as a branching tree, and the path to the present was just one out of many possible branches (Thomason, 1984). Past-as-past approaches postulate that the past tense morpheme in non-past counterfactual constructions maintains a true past tense component: it allows access to a counterfactual world by shifting back in time so that an alternative future path from the one that actually led to the present can be selected. Structurally, this would mean that the past tense morpheme is interpreted outside its usual position, as it scopes over a modal operator and logically embeds the entire proposition.

In contrast, past-as-modal ('remoteness-based') approaches believe the counterfactual's past is "fake" in the sense that the morpheme does not make any temporal reference. Within this approach, opinions differ in whether they take the past tense morpheme to correspond to one meaning (both inside and outside counterfactuals) (Bjorkman & Halpert, 2017; Iatridou, 2000; Karawani, 2014; Karawani & Zeijlstra, 2013; Ritter & Wiltschko, 2014) or whether the past tense morpheme is ambiguous between regular past tense meaning and the meaning it expresses in counterfactual contexts (Schulz, 2014). For example, Iatridou (2000) argues that the past tense morpheme is the realization of an 'exclusion' feature, that either scopes over time (excluding the present, resulting in a past tense reading) or over worlds (excluding worlds, resulting in a counterfactual reading). Access to the possible worlds (modal operator) that the exclusion feature ranges over is provided by either the conditional structure or by *wish* and takes scope inside the conditional antecedent, and not outside it.

Syntactic accounts of counterfactuals in various languages similarly show that structurally, the past morpheme in counterfactual constructions is located or licensed in a position higher than where regular past tense is licensed (Aygen, 2004; Bjorkman, 2015; Bjorkman & Halpert, 2017; Karawani, 2014; Karawani & Zeijlstra, 2013; Ritter & Wiltschko, 2014). One piece of evidence for this higher structural position of the "fake" past for English comes from the fact that counterfactual conditionals can undergo subject-auxiliary inversion (7a), while regular conditionals cannot invert (7b) (Iatridou & Embick, 1994; Ritter & Wiltschko, 2014).

(7a) **Had** she arrived, I would not have left. (Ritter & Wiltschko, 2014, 63)

(7b) ***Has** she really arrived, she will be here. (Ritter & Wiltschko, 2014, 64)

For this paper, we are not committed to any semantic or syntactic analysis. Our takeaway from reviewing the literature is that while there is disagreement about whether the "fake" past morpheme performs any temporal function, almost all approaches have in common that the "fake" past tense expresses an operation distinct from its regular past tense usage (e.g., scoping at a higher position or operating over worlds rather than times), and that there is cross-linguistic evidence to assume that the syntactic position of the past morpheme in counterfactual constructions is similarly distinct (higher) from its regular past counterpart. From a theoretical linguistic point of view, children seeing through the "fakeness" of the counterfactual past, thus entails (a) realizing that the past tense morpheme in counterfactuals must be interpreted differently than the past tense in a non-counterfactual construction, and (b) learning the semantic and syntactic operations supporting this counterfactual interpretation.

2.4. Aims and Hypotheses

In this paper, we use child corpus data to investigate the role of the "fake" past tense in the development of children's productions of counterfactuals in English. We investigate whether children go through a stage where they treat "fake" past tense as real. To our knowledge, no prior naturalistic study has looked specifically at the early acquisition of counterfactual wishes, nor has the "fake" past tense been recognized as an acquisition challenge. Additionally, we compare the acquisition of counterfactual wishes with conditionals. Although several studies included counterfactual conditionals in their overview of the acquisition of conditionals in general (Bowerman, 1986; Katis, 1997; Reilly, 1982), the development of counterfactual wishes compared to other counterfactual constructions has not yet been considered.

The mapping of counterfactual meaning onto counterfactual expressions is by no means trivial. In English, it requires the mapping of multiple abstract meanings (past and counterfactuality) to the same past tense inflection, where the counterfactual meaning is the less frequently occurring variant. Children thus have to figure out that the past tense morpheme sometimes indicates counterfactuality rather than past temporal orientation, and the exact environments in which this is the case. In English, there are several ways in which this mapping is less complex for counterfactual *wish*-constructions than for counterfactual conditionals. First of all, wishes lack the structural complexity of conditionals, which involve a dependency relationship that requires keeping in mind and causally relating two clauses (c.f. Reilly, 1982; Bowerman, 1986). More importantly, propositional embedding *wish* is a dedicated counterfactual marker, meaning that whenever it takes a full proposition as its complement, this proposition is interpreted counterfactually. (Note that while the verb *wish* does sometimes co-occur with Noun Phrase (NP) ("I wish you a happy birthday"), Verb Phrase (VP) ("I wish to sleep") or Prepositional Phrase (PP) complements ("I wish for more presents"), these uses where *wish* selects for a non-propositional complement are structurally distinguishable from propositional embedding *wish*). Acquiring the counterfactual *wish*-construction thus requires the child to learn that the verb *wish* differs from desire verbs like *want* in its counterfactual implication and can only be used when the desire is believed to be unfulfilled or unattainable. Because of the *wish*-construction's dedication to counterfactuality, it cannot occur with a present tense complement, even when the temporal orientation of the wish is present tense (8).

(8a) *I wish I **have** a car.

(8b) I wish I had a car (Iatridou, 2000, 25)

This is in contrast with conditionals, where the complementizer *if* can introduce both hypothetical conditionals (9a/b) and counterfactual conditionals (9c).

(9a) If he **has** time to bake cookies, he will bring some. PRES. CONDITIONAL

(9b) If he **had** time to bake cookies, he will bring some. PAST CONDITIONAL

(9b) If he **had** time to bake cookies, he would make some. PRES. COUNTERFACTUAL

The salient mismatch between the temporal orientation and morphological past marking of the *wish*-complement (8a) may cue the child into realizing its role in expressing counterfactual meaning. Conditionals that can appear with present (9a), real past (9b) and "fake" past tense (9c) in their antecedent, make it less transparent to discover that the counterfactual conditional's past tense does not indicate a past temporal orientation. We thus hypothesize that counterfactual wishes in English are easier to acquire than counterfactual conditionals and might even help children bootstrap into acquiring counterfactual conditionals. If wishes are easier to acquire, we expect the onset of production to precede those of counterfactual conditionals.

Since the *wish*-construction is dedicated to embedding counterfactual propositions, it could be the case that children initially map counterfactuality to the entire *wish*-expression, without realizing the role past tense inflection plays in encoding counterfactual meaning. If this is the case, we predict children to go through a stage where they interpret the "fake" past tense as having real past temporal orientation. In spontaneous production, we would thus expect to observe instances where children express counterfactual wishes about the present with present tense marking like (5), since they generate the counterfactual desire using their own non-adult grammar. We expect such present-for-past tense errors to cease once children have figured out the link between the semantic operations giving rise to counterfactual meaning and the expression of past tense morphology. Alternatively, it could be the case that realizing the counterfactual function of the past morpheme is a necessary prerequisite for expressing counterfactuality. If this is the case, tense errors are expected to indicate a non-adult like use of the counterfactual *wish*-construction. For example, if a child produces “I wish I have a car” this use of wish with a present-marked or bare verb complement could indicate a simple desire, in line with non-counterfactual desire verbs like *want* or *hope*. We then would expect to find that wishes containing present-for-past substitutions do not indicate clear adult-like counterfactual reasoning.

3. METHODOLOGY

3.1. Part 1: Comparing Children and Adult’s *Wish*-utterances

3.1.1. Selection Criteria & Preprocessing

We looked at natural child productions of counterfactual constructions by searching through English corpora of transcribed children’s speech available on CHILDES (MacWhinney, 2000) using the database ‘childes-db’ (Sanchez et al., 2019), accessed through the statistical software environment R (R Core Team, 2021). All operations involving corpus extraction were performed using the analysis package ‘childesr’ (db version = "2020.1"). We selected corpora that contained data from typically developing monolingual children between 2;5-6;0, yielding 57 corpora (48 from Northern America, 11 from the United Kingdom) including data from 585 children in total. In Appendix S3.1 we provide you can find an overview of all corpora used.

For these corpora, we extracted all utterances and calculated the amount of child and adult utterances. For this calculation, speakers with the speaker roles “Target Child”, “Child”, “Sister”, “Brother”, “Friend”, “Playmate”, “Girl” and “Sibling” were included in the child category, all other roles were treated as adults. We noticed that a small proportion of the data (77551 utterances, 3.5%) across 15 different corpora (partially) lacked age information for the children in the output of the ‘get_utterances( )’ function. Most missing age data (2.5%) could be recovered from a participant overview extracted with the function 'get_participants( )', and for the remaining 13 corpora that still (partially) lacked target child age information we manually recovered the information where available by retrieving it from the CHILDES Talkbank corpus description pages on https://childes.talkbank.org/access/. For two corpora (MacWhinney and Gathercole) age information was displayed incorrectly (based on the metadata available in the corpus descriptions), so this was manually corrected by extracting the info from the corpus description pages (Gathercole) or recalculating the children's ages based on the transcript file name (which was based on the age of the child 'Ross', so in order to calculate the age of his younger sibling 'Mark' we subtracted 01;10;25). We then filtered the data set to only include utterances from children that are within our age-range of interest 2;0-6;0 and proceeded to extract all child utterances containing the word *wish*. In total, 40 of the searched corpora contained child wishes. For these 40 corpora we also extracted all adult utterances (child-directed speech and speech addressed to other adults within the child's hearing), so we could compare *wish* usage between children and adults.

3.1.2. Exclusions

To get an idea of the proportion of *wishes* present in spoken child and child-directed speech, we calculated the percentage of *wish* utterances for the child and adult corpora. We extracted 478 child utterances containing *wish* (0.02% of 2,247,665 total utterances) coming from 40 different corpora, and 841 adult *wish*-utterances (0.03% of 2,934,114 total utterances). To make a fair comparison between the *wish*-productions of children and their input (child-directed or overheard adult *wish*-utterances), we only analyzed adult data from the 40 corpora we found child wishes in. For the adult utterances, we thus proceeded to exclude 70 utterances that came from corpora that did not yield any child wishes. For the child utterances, we excluded 10 child wishes for which the target child's age was unknown. For the remaining 468 child and 771 adult *wish*-utterances, we first excluded all utterances in which *wish* was used as a noun (e.g., "Do you want to make a wish?"), which resulted into 29 exclusions for child utterances and 129 for adults. Since the verb *wish* is counterfactual only if its complement is a full proposition (Iatridou, 2000, p.241), we then excluded utterances where *wish* did not embed a proposition. For children, this resulted in 58 exclusions (2 VP complements, e.g., "not wish to play"; 17 NP complements, e.g., "I wish you a happy birthday"; 5 PP complements, e.g., "I wish for daddy to come home" and 34 instances where there was no complement, e.g., "yeah I wish"). For adults we excluded 142 non-propositional complements (11 VP, 69 NP, 13 PP and 49 missing embeddings). Lastly, we excluded an additional 32 child wishes and 15 adult wishes for being a repetition of either themselves or someone else. This means that in total 349 child wishes and 485 adult wishes remained for further analysis.

3.1.3. Coding Conventions

All *wish*-utterances were manually coded for various structural and semantic linguistic variables. Structural linguistic variables included: person of the main subject, i.e., 'the wisher' (*I* and *we* = 1st person; *you* = 2nd person; *Mommy*, *he* and *the cat* etc. = 3rd person; no subject = omitted; inaudible subjects = unclear), person of the subject of the *wish*-embedding (same coding convention as main subject) and subjunctivity of singular 1st and 3rd person inflections of *to be*: (*was* = not subjunctive; *were* = subjunctive). We also coded for morphological tense-marking errors, i.e., tense inflections that diverge from the grammatical form used by adults in this structural context. Errors were separated into those that lack past-tense marking in the *wish*-complement, i.e., 'present-for-past' (e.g., "I wish you can't **do** that") or 'other' tense errors (e.g., "I wish we have **gotted** some mail" or "I wish I **be** a sheep"). For all present-for-past errors, we coded whether they were compatible with a 'bare verb usage' which could signal children having dropped *would/could* (e.g., "I wish I <could> do that"). If a child used an auxiliary ("I wish we **can** eat") or other inflected form ("I wish I'**m** already at home") we marked the error as incompatible with bare verb usage. As a first semantic variable, we coded for the temporal orientation of the embedded clause (e.g., "I wish I had a train" = present; "I wish I had gone to the train" = past; "I wish I would have a train" = future; "I wish want a train" = unclear). Unlike adults, who use *would* in future wishes (e.g., "I wish you **would**n't do that"), children's utterances sometimes lack *would* in wishes with a future temporal orientation (e.g., "I wish you stop bug me"). Since lexical aspect contributes to the temporal orientation (Iatridou, 2000), wishes without *would* were coded as present when containing stative verbs (i.e., *had*, *was, knew*) and as future when containing eventive verbs (e.g., *go*, *stop*, *got*). The tests used to determine stative or eventive lexical aspect came from (Dowty, 1986).

When children use *wish*-constructions, it is not assured that they understand that the *wish* statement is a counterfactual utterance, and thus indicates unobtainable desires. For this reason, we coded for the evidence we have available as coders to determine whether the wish is used counterfactually or not. We inspected the discourse and situational context as available in CHILDES transcripts, to determine whether the wish demonstrated 'clear' counterfactual reasoning. Counterfactual wishes were considered to contain clear counterfactual reasoning when lexical material within the utterance itself contrasted the actual world with a counterfactual one (e.g.,: "I wish I asked for toast **instead**" = lexical contrast, "I wish you did**n't** do that" = contrast induced by negation, "I wish I **had gone** to the station" = contrast induced by undoing past event), when the wish desired some sort of existential change, i.e., was counteridentical (e.g.,: “I wish I was a monkey”), or when the utterance was in clear contrast with prior context (e.g.,: “I wish I had green eyes.” = contextual contrast when used in a context where it is clear the speaker does not have green eyes). Wishes that were undistinguishable from a regular desire usage (e.g., “I wish I had that horse” or “I wish you’d stop”) were marked as having no evidence for counterfactuality, and wishes were transcribed without context were coded as “inconclusive”. Different than for children, we did code adult *wish*-utterances expressing desires such as “I wish I had a kitty” or “I wish I could talk to her” as contextual counterfactuals (without investigating the context it was uttered in) assuming adults always use *wish* counterfactually.

All data was coded by the first author (a fluent non-native speaker). A random subset of 100 child wishes were double-coded, by a native speaker of English (both coders were trained in semantics). An inter-rater reliability analysis was performed to determine consistency among raters in coding for the described variables, using overall accuracy, Gwet’s AC1 coefficient (Gwet, 2008) and Cohen’s kappa statistic (Cohen, 1968) to describe agreement confidence. While Cohen’s kappa statistic is often used as the default method to determine intercoder reliability, it can underestimate reliability in cases where there is high agreement in unbalanced distributions (Gwet, 2008). Since multiple of our coding variables are unbalanced (e.g., temporal orientation is overwhelmingly present), AC1 is likely a more stable measurement. The exact values for all three different statistics for our coding are displayed in Appendix S3.2. The AC1 values for all variables exceeded 0.85 (very good agreement) except for the coding indicating the available evidence for counterfactuality (percent agreement = 61%, AC1 = 0.52, $\kappa = 0.49$), which corresponds to moderate agreement (Landis & Koch, 1977). Since coding involves assessments of grammatical and situational contexts, coders discussed all disagreements and came to a consensus for items where either coder missed contextual or grammatical cues in their original rating. The first coder (who coded the entire dataset) was more accurate and conservative than coder two (who only coded the subset). 17 items were judged in favor of coder 1, and 7 items in favor of coder 2. Of the 7 items judged in favor of coder 2, only 1 item was changed from formerly being judged counterfactual to no evidence for counterfactuality. A subset of 13 disagreements remained where coders diverged and contextual cues could be interpreted in different ways. Again, coder 1 tended to code more conservatively, as 11 of these items were categorized as having no or unclear evidence for true counterfactuality, while coder 2 was willing to consider these utterances as true counterfactuals. The intercoder reliability values for evidence of counterfactuality post-discussion corresponded to very good agreement (percent agreement = 87%, AC1 = 0.84, κ = 0.83). Altogether, this suggests that the coding of our dataset might error on the side of not categorizing potentially counterfactual wishes as counterfactual, rather than overestimating the instances of wishes displaying counterfactual reasoning.

3.1.4. Data Analysis

For each coded syntactic and semantic variable, we calculated the total count and percentage of occurrences per condition for children and adults separately. We converted the error data into a binary variable coding for the presence or absence of a present-for-past substitution, and modeled the probability of making present-for-past tense errors with a generalized linear mixed-effect model (GLMM, Baayen et al., 2008). We used the glmer-function from the ‘lme4’ package available on R to perform our analysis (Bates et al., 2015; R Core Team, 2021). We ran two separate models, one over the complete dataset with the fixed effect of age group (child versus adult) to investigate whether children produced more tense errors than adults, and one over the child data with age in months as a fixed effect, to investigate whether children's age predicts their error rate. For both models, we included speaker identity as a random effect to include the variation found among speakers in the model estimates. Inclusion of a random slope or the addition of corpus identity as a random effect did not improve the fit of our models. The model fit (logit link) was estimated by maximum likelihood using the default setting of LaPlace approximation. To test the contribution of our fixed effects we performed a likelihood ratio test comparing our model and a nested model leaving out the variable of interest. We used the 'DHARMa' package (Hartig, 2022) to test the dispersion of our models, and found no indication of overdispersion, which means that the residual variance of our data was not larger than our fitted models assume.

3.2. Part 2: Individual Development of Counterfactual Utterances

3.2.1. Selection Criteria & Preprocessing

To gain more insight into the individual longitudinal development of children, we selected children that produced more than 15 wishes. From the complete dataset, six children fit this criterion: Abe - Kuczaj corpus (Kuczaj, 1977), Adam - Brown corpus (Brown, 1973), Laura - Braunwald corpus (Braunwald, 1971), Mark & Ross - MacWhinney corpus (MacWhinney, 1991) and Thomas - Thomas corpus (Lieven et al., 2009). For these 6 children, we searched for counterfactual conditionals by extracting utterances containing *if* in combination with *would*, *should* and *could*. We proceeded to compare the emergence and development of their first spontaneous counterfactual conditionals against the development of their *wish*-utterances.

3.2.2. Exclusions

The 6 children with longitudinal data were responsible for 175 of the wishes. For those 6 children, we also extracted 341 conditionals with *would*, *should* or *could*. We excluded 63 utterances where *if* was used like *whether* and not as an *if-then*-conditional (e.g., "see if you could throw two dinosaurs in"), and 93 utterances that did not contain past tense inflection in the *if*-clause. We did this to exclude (non-counterfactual) hypothetical conditionals such as "Maybe you shouldn't be there, if you scare Ellen" or "What would the toilet be like if you flush it?". A total of 185 conditionals remained. Because we were interested in the relative onset difference between counterfactual wishes and conditionals, we decided to be conservative in our inclusion criteria of what consists as a counterfactual. For this reason, we excluded all wishes and conditionals that have future temporal orientation, since their status as counterfactual is debated (strictly speaking, the future cannot be counter-to-fact, as it has not yet occurred). We excluded 26 wishes such as "I wish that you stop talking" and 80 conditionals like "Mom what would happen if I taked this balloon". We were left with 104 counterfactual conditionals and 149 wishes with present or past temporal orientation.

3.2.3. Coding Conventions

For the conditionals, we coded for the same semantic variables as we did for the wishes. For temporal orientation this included the categories 'present' (e.g., "they could fly if they had wings"), 'past' (e.g., "what would have happened if they didn't invent houses") and 'future' (e.g., "if I saw a real wolf I would kick the real wolf"). For evidence for counterfactuality this again included clear lexical counterfactuals (e.g., lexical contrast: "only if Super Man was **real** he could do it", negated contrast: "but if I was**n't** a chair I wouldn't be a chair", or past contrast: "yeah it could have lived if I would **have gotten** enough food for all of them"), counteridenticals (e.g., "if I were you I would eat food") or contextual counterfactuals (e.g.,: "if there were four one would hafta wait his turn", when used in a context where there are less than four). Conditionals that were undistinguishable from a regular hypothetical (e.g., "if I could get my boots on I could go inside") or uttered out of context were marked as "inconclusive". Since we excluded all conditional utterances that had present tense marking in the *if*-clause, we could not code for possible present-for-past substitutions.

3.2.4. Control Comparison

We hypothesized that present-for-past substitutions in the *wish*-complement could indicate children have not yet figured out that counterfactual utterances require the "fake" past morphology. Alternatively, it could be the case that some children have yet to develop the ability to use the past tense in appropriate contexts, and generally avoid using the past tense in any environment, including (but not limited to) counterfactual utterances. To investigate this possibility, we determined for each child the period in which they made present-for-past tense errors and extracted all utterances containing the word *yesterday* during this period, as well as all utterances containing a past tense morpheme. This yielded 29 utterances with *yesterday*, and 7033 utterances with past tense. We looked for signs of productive tense marking by indicating whether children correctly inflected the main verb of utterances containing the temporal adverb *yesterday* with past, and whether their other past utterances included any instances of overregularization (e.g., "I telled daddy something").

3.4.5. Data Analysis

For the coded semantic variables, we calculated the total count and percentage of occurrences per condition for all six children. We created a new variable for evidence of counterfactuality that grouped evidence into binary bins as either "clear" (lexical, counteridentical or contextual evidence) or "unclear" (inconclusive or no evidence). We then compared per child the onset of wishes and conditionals per category, and calculated the difference between the two. We then averaged over children to get an idea of the average difference between the onset of wishes and conditionals. Since we only had data for six children, we discuss these results descriptively and conducted no further statistical analysis.

4. RESULTS

4.1. Children's Wishes

In total we found 349 *wish*-constructions (*wish* + proposition) in children between the ages of 2 and 6. The first instance of the *wish*-construction we found at 25 months (10a). Like most early wishes, this wish expresses a desire about something mentioned or in direct proximity, e.g., wishing for a horse when looking at horses (10b).

Early Wishes (Like Desires)

(10a) Laura (2;1): I wish I had sandals. (Braunwald, 1971)

(10b) Becky (2;7): I wish I had a horsie. (Manchester: Theakston et al., 2001)

From these early uses, it is not clear whether children know that *wish* can only be used counterfactually, i.e., the desire is either established to be impossible or unlikely to be fulfilled. So it could be that children initially use *wish* like the regular desire verb *want*. Consistent with this possibility, we sometimes encounter clear non-counterfactual wishes, where parents comment on the incongruency (11a/b).

Non-Counterfactual Wishes

(11a) Emily (2;1): but I wish that my cold is better. (Nelson, 1989)

Father: yeah you had no cold at all everything's fine.

(11b) Laura (3;2): I wish you were my mommy. (Braunwald, 1971)

Mother: I am your mommy.

For this reason, we coded for the evidence we have available as researchers to believe that a child's wish is produced with a counterfactual meaning in mind. We separated the wishes into 5 categories: wishes that seem clearly counterfactual based on lexical information inside the utterance (12-14), i.e., contrasting the actual world against the postulated one through undoing the past, negation or a lexical contrast (n=43, 12% of total wishes); wishes that indicate an existential change (15), i.e., counteridenticals (n=27, 7.8%); wishes that are in clear contrast with reality as deduced from the discourse context (16) (n=96, 27.5%); wishes that provide no evidence for counterfactuality (n=69, 19.8%) and wishes that are not interpretable without more context and therefore provide inconclusive evidence (n=114, 32.7%).

Clear Evidence for Counterfactuality

Lexical Evidence: Undoing Past

(12) [hearing train in distance] (Thomas: Lieven et al., 2009)

Thomas (3;1): I wish **gone** Burnage Station watch that train.

<later in recording Thomas comments "I'm missing all the trains">

Lexical Evidence: Negation

(13) [mother about to braid child's hair] (Hall: Hall & Tirre, 1979)

Mia (4;9): I wish you did**n't** hafta braid it.

Lexical Evidence: Lexical Contrast

(14) [child pretends it's his birthday] (Thomas: Lieven et al., 2009)

Thomas (4;2): Oh I wish it was my birthday today **really**.

Counteridentical (Change of Identity)

(15) Ross (4;2) I wish humans were **not** humans. (MacWhinney, 1991)

Contextual Evidence

(16) Father: You don't see bumblebees in the dark at all.

Mark (5;10) I wish that the lights were on. (MacWhinney, 1991)

Most wishes uttered by two-year-olds lack clear evidence for counterfactuality. The first *wish*-constructions that we coded as having clear evidence for a counterfactual intended meaning start around 35 months, this is true for all three categories (lexical, counteridentical and contextual). This finding is visually displayed in Figure 3.1 below.

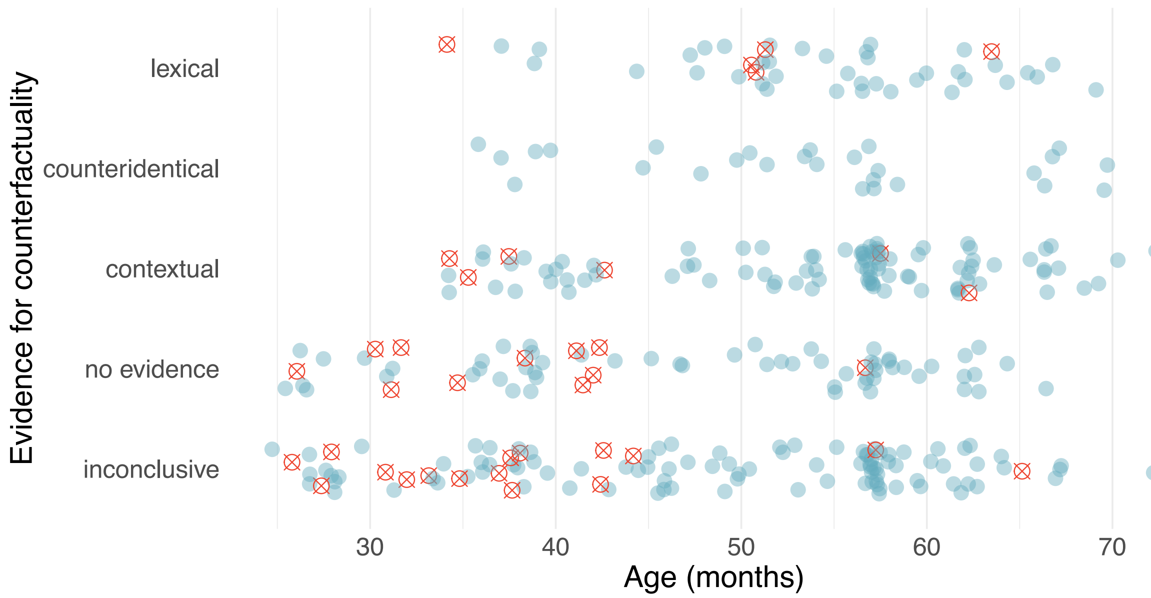

Figure 3.1. Breakdown of children's wishes. Plotted are all children's wish-productions (N=349) per evidence category for indicating counterfactuality (y-axis). Evidence that is lexical, counteridentical or contextual is considered to indicate clear counterfactuality, while no or inconclusive evidence indicates that it's unclear whether the utterance is used counterfactually. Red struck-through instances indicate the wish contained a present-for-past substitution (e.g., "I wish I have a horse"). The x-axis indicates the speaker's age in months.

4.2. Tense Errors

The tense expression in the complement of children's produced wishes diverged from the adult-form in several ways. The most frequently occurring error (38 instances, 10.9% of total), was that of using present tense in the *wish*-complement rather than past tense. For adults, we only documented 4 instances where present tense was used inside the *wish*-complement (0.8% of the total amount of 465 adult wishes). Children are thus not matching their input when making these productive tense substitutions. We modeled the presence or absence of present-for-past errors with a generalized linear mixed-effects model (GLMM) including speaker identity as a random factor to investigate whether age group (child or adult) was a predictor of error rate. A likelihood ratio test comparing our model against a nested model without fixed effects, found that age group was a significant predictor of error rate ($\chi^2(2) = 4.75$, $p = .029$). The odds of making a present-for-past substitution increased for children compared to adults ($\beta = 17.5$, $z = 3.67$, CI = 3.79 - 80.7). Children's present-for-past errors are marked on Figure 3.1. For 15 of these errors, it is not entirely clear whether they are marking present tense or are the consequence of dropping 'would', since the present tense is indistinguishable from bare verb usage in these cases (17). For the remaining 26 errors it was clear that they indicated present tense, i.e., due to inflection (18a) or from the choice of auxiliary (18b).

Present-for-Past Errors

(17)	Adam (5;2):	I wish I **have** a banjo like dis [this].	(Brown, 1973)
(18a)	Sarah (3;6):	I wish it'**s** valentine.	(Brown, 1973)
(18b)	Martin (3;6):	I wish I **can** be on the tellie.	(Wells, 1981)

Present-for-past mistakes are more common among younger children, especially those between age 2 and 3. With a second GLMM analysis considering speaker identity as a random effect, we confirmed that age in months is a predictor for children's error rate ($\chi^2(2) = 22.26$, $p < .001$). The odds of making a present-for-past mistake decreased with every month ($\beta = .911$, $z = -4.27$, CI = .088 - .951). When we group the present-for-past tense mistake counts by age group (per year) we observe indeed that most present-for-past substitutions occur before age four, and then drop off steeply. This decrease in error rate is displayed in Table 3.1.

Table 3.1. Count and percentage of present-for-past tense errors per age window

Age Group	# children	# wishes	# errors	% of total
2-3	18	47	15	31.9
3-4	21	84	14	16.7
4-5	41	148	6	4.05
5-6	19	70	3	4.29
Total	99	349	38	10.9

Besides making present-for-past errors, we also found that children sometimes express wishes about the past without using the past perfect (19a/b). A similar omission of the *had* auxiliary in the past perfect was observed in example (12). Interestingly, we observed the same for adults (20a/b).

(19a) Abe (4;4): Are we having pork chops for dinner? (Kuczaj, 1977)

Mother: Yes, that's what you asked for.

Abe (4;4): I wish I **asked** for toast instead.

(19b) [child did not have a nice time at his grandma's] (Thomas: Lieven et al., 2009)

Thomas (3;2): because I wish Mum **come** there.

Investigator: ah, did you miss your mum?

Thomas (3;2): yeah

(20a) Mother: oh don't we wish we **had** that three weeks ago

(20b) Mother: don't you wish you **had** them when you were little

(Dickinson & Tabors, 2001)

4.3. Comparing Children and Adult's *Wish*-utterances

Next, we compared the syntactic and semantic properties of the 485 adult and 349 child wishes. The proportion of child wishes (0.02% of all utterances) was overall comparable to the proportion of adult wishes across all corpora (0.03%), and we found that children and adults used wishes in a comparable way (Figure 3.2). The lion's share of wishes are produced from a 1^{st} person perspective, and children use 1^{st} person main clause subjects (83.7%) even more than adults (76.8%) (Figure 3.2A). This is compatible with the intuition that young children mostly talk about themselves. Similarly, their wishes are mostly about themselves as well, i.e., the embedded subject is first person (49.3%). In contrast, the embedded subject of adult wishes is balanced for person: 1^{st} (36.3%), 2^{nd} (31.0%) or 3^{rd} (32.3%) person (Figure 3.2B). As for temporal orientation, we see that both children and adults mostly wish about the present (children: 76.2%, adults 62.6%), followed by the future (children: 11.7%, adults: 24.9%) or the past (children: 4.0%, adults: 12.3%) (Figure 3.2C). However, it is possible that the counts for children's past and future wishes are somewhat underestimated, as they sometimes left out the past perfect *had* and future *would* auxiliary (discussed in prior section), making them hard to distinguish from the present (e.g., "I wish I come"). Below you find examples of wishes with present (21), past (22) and future (23) temporal orientation produced by children and adults. Counterfactual wishes with a future orientation often indicated a desire to change a habit or a future event that that has already been planned or whose outcome is determined (23a). The counterfactuality in these cases is the implication that this desire is unattainable. For adults, most of the future-oriented wishes express indirect requests (23b).

Wishes with Present Temporal Orientation

(21a) Ross (5;7): I wish you were a little kid then you would understand. (MW, 1991)

(21b) Mother: I wish it was real money. (Thomas: Lieven et al., 2009)

Wishes with Past Temporal Orientation

(22a) Abe (4;3): I wish we haven't come here. (Kuczaj, 1977)

(22b) Father: Boy, I wish Dallas had won the football game. (Kuczaj, 1977)

Wishes with Future Temporal Orientation

(23a) Matthew(4;7): I wish they'd give ya a fork instead of a spoon. (Gathercole, 1980)

(23b) Father: I wish you'd stop hitting. (MacWhinney, 1991)

When we break down the available evidence for counterfactuality, we see that children and adults also pattern alike here. Most wishes were judged to be clearly counterfactual based on contextual evidence (children: 27.5%, adults: 47.1%), followed by lexical evidence (children: 12.3%, adults: 19.8%) and counteridenticality (children: 7.7%, adults: 2.6%) (Figure 3.2D). The fact that we observe less contextual wishes for children than for adults could be a consequence of the fact that we conservatively coded for desire-like wishes in children (e.g., "I wish I had a horse" without clear supporting contextual evidence for counterfactuality was coded as having "no evidence") while we assumed that adults use these wishes as true counterfactuals. Last, we compared the counts of subjunctive usages, by looking at 1st or 3rd person singular conjugations of *to be* in both children (n=54) and adults (n= 67) and coded for whether these were marked with subjunctive (*were*) or not (*was*). We found that adults somewhat rarely used the subjunctive form (19.4%), and for children we observed only 3 instances (5.6%) (Figure 3.2E). For children, all subjunctive wishes came from the Northern-American corpora. For adults, we found only 2 subjunctive wishes (2.9%) in the United Kingdom corpora. This difference could be due to the fact that our sample from the Northern-American collection was bigger and skews historically older than our UK-sample. Examples of wishes with and without subjunctive mood are provided below for children (24a/b) and adults (25a/b).

Child Wishes with and without Subjunctive

(24a) David (4;9): I wish I **were** in a car. (Hall: Hall & Tirre, 1979)

(24b) Joey (4;9): Yes, I wish I **was** a spoon. (Hall: Hall & Tirre, 1979)

Adult Wishes with and without Subjunctive

(25a) Father: I wish it **were** but it's not. (Clark, 1979)

(25b) Adult: I'll tell you I wish it **was**. (Hall: Hall & Tirre, 1979)

4.4. Individual Development of Counterfactual Utterances

To understand the developmental trajectory of individual children, we investigated the emergence of counterfactual wishes and conditionals in the output of six children we had enough longitudinal observations for. We investigated both the clarity of the counterfactual (whether there is evidence that indicates the expression is used counterfactually) and whether the child made any present-for-past tense mistakes. The individual development of each child is displayed in Figure 3.3.

4.4.1. Comparing the Onset of Counterfactual Wishes and Conditionals

The age at which the 6 children started to use the *wish*-construction varied from 2;01 (25 months) to 4;00 (48 months). The age of the first clear counterfactual wish usages fall within a later range between 2;10 (34 months) and 4;11 (59 months). For (both clear and unclear) counterfactual conditionals the onset range was 2;8 (32 months) – 4;4 (52 months). Examples of children's first counterfactual conditional constructions are provided in (26a/b). The onset of the first wish/conditional was often followed with subsequent usages of the constructions within as short period of time. Repeated uses of a new construction within a short period of time is considered to be a signal of productivity (Snyder, 2007; Stromswold, 1990).

First Counterfactual Conditionals

(26a) Laura (2;8): If a really hole was in here, (Braunwald, 1971)

then I would cry for new pants.

(26b) Mark (3;7): We could fly if we had wings (MacWhinney, 1991)

well, we don't so we can't, but I know one way how you can fly

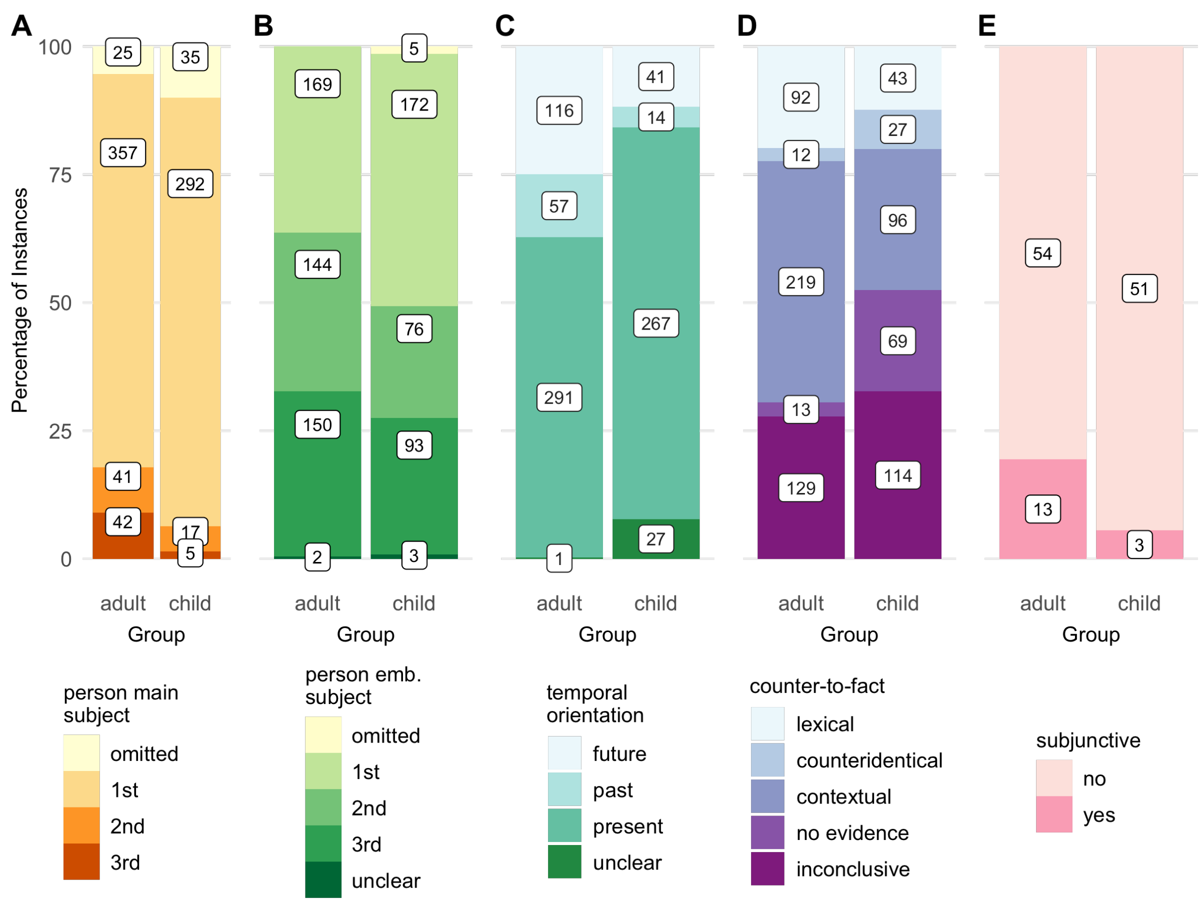

Figure 3.2. Overview of syntactic and semantic properties of child and adult wish-constructions. Count (total A-D = 465 for adults and 349 for children, E = 67 for adults and 63 for children) and Percentage (y-axis) of instances.

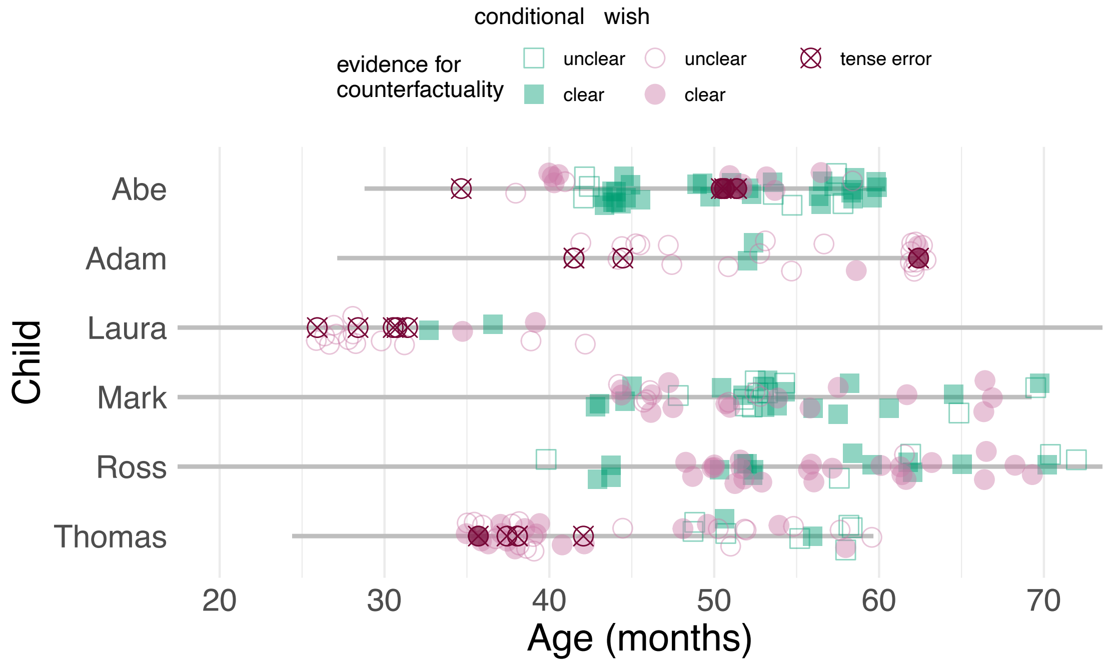

Figure 3.3. Counterfactual conditionals (green squares) and wishes (pink circles) for each child (y-axis) with age indicated in months on the x-axis. Filled shapes indicate that the evidence for counterfactuality is clear, empty shapes indicates the evidence is unclear. Struck-through wishes indicate they contained a tense error in the form of a present-for-past substitution. Grey line indicates recording span.

To quantify the average difference between the onset of wishes and conditionals for each child, we compared the onset per evidence category (unclear and clear) and calculated the average values. This numerical comparison is displayed in Table 3.2. On average, children started producing counterfactual wishes before conditionals, though the difference is more prominent if we consider unclear counterfactuals (4.7 months earlier) than if we compare the average onset of clearly counterfactual constructions (0.6 months earlier). However, there is a lot of individual variation in the presence and size of the gap between the onset of the two constructions. 4/6 children start using (unclear) counterfactual *wish*-constructions before they use conditional constructions (difference ranging from 6.6 – 13.6 months), Mark started using both constructions around the same time, and Ross was the only child who used counterfactual conditional constructions before wishes. Comparing clear counterfactual wishes and conditionals, we find that only 2 children (Abe and Thomas) start using wishes before conditionals (difference 3.6 and 15.6 months). For Mark and Laura they emerge around the same time, and for the last 2 children it seems that clear counterfactual conditionals precede the onset of clear counterfactual wishes (for Adam by 6.4 months, and for Ross by 5.2 months).

Table 3.2. Overview of children's age (in months) at time of first (clear) counterfactual wishes and conditionals (cond.)

Child	Age 1st wish	Age 1st cond.	Age 1st cond - wish	Age 1st clear wish	Age 1st clear cond.	Age 1st clear cond. - wish
Abe	34.7	42.1	7.4	39.9	43.5	3.6
Adam	41.5	52.4	10.9	58.8	52.4	-6.4
Laura	25.8	32.4	6.6	34.6	32.4	-2.2
Mark	44.6	42.8	-1.8	44.6	42.8	-1.8
Ross	48.3	39.9	-8.4	48.3	43.1	-5.2
Thomas	35.5	49.1	13.6	35.5	51.1	15.6
Average	38.4	43.1	4.7	43.6	44.2	0.6

4.4.2. Present-for-Past Errors

We observed that most present-for-past tense errors occur in the early stages of the emerging *wish*-construction, regardless of the age the child started using the construction. It should be noted that we found present-for-past mistakes in both unclear (n= 13) and clear (n= 5) counterfactual wishes. Two children (the siblings Mark and Ross) never made a present-for-past substitution in their wishes, and two children (Laura and Thomas) made multiple present-for-past substitutions when they started using the *wish*-construction, and then stopped making them before their first counterfactual conditionals emerged. This means that for 4/6 children, present-for-past substitutions did not occur after the onset of the counterfactual conditional. Adam and Abe complicate this picture. Adam initially stopped making tense mistakes around 45 months (about 7 months before his first counterfactual conditional), but then slipped up at age 5;2 (62 months). Since this also marked the end of his recording period, it is unclear whether he made any more present-for-past substitutions after this occurrence. Abe is unique in making present-for-past substitutions when both his counterfactual wishes and conditionals are productive (at age 4;3, 51 months).

4.4.3. Productive Tense Marking

Lastly, we examined children's overall productive past tense usage during the period where they made present-for-past errors in counterfactual constructions. We did this to investigate whether their present usage in counterfactual contexts is due to a variable or inconsistent use of past tense marking in general. For each child, we recorded the successful and unsuccessful instances of past tense marking in the context of the temporal adverb *yesterday*, and the period over which they exhibit overregularization. This is displayed in Figure 3.4. For all children, we found indications of productive past tense usage (both from overregularization and past tense usage with *yesterday*) outside counterfactual contexts during their error period. While Abe used present inflection once in a *yesterday* utterance at the onset of his error period, he later correctly started using past tense in this environment. For Laura we found multiple present tense errors with *yesterday* before 28 months. This indicates that some of Abe's and Laura's earliest errors could be due to an immature use of the past tense in general.

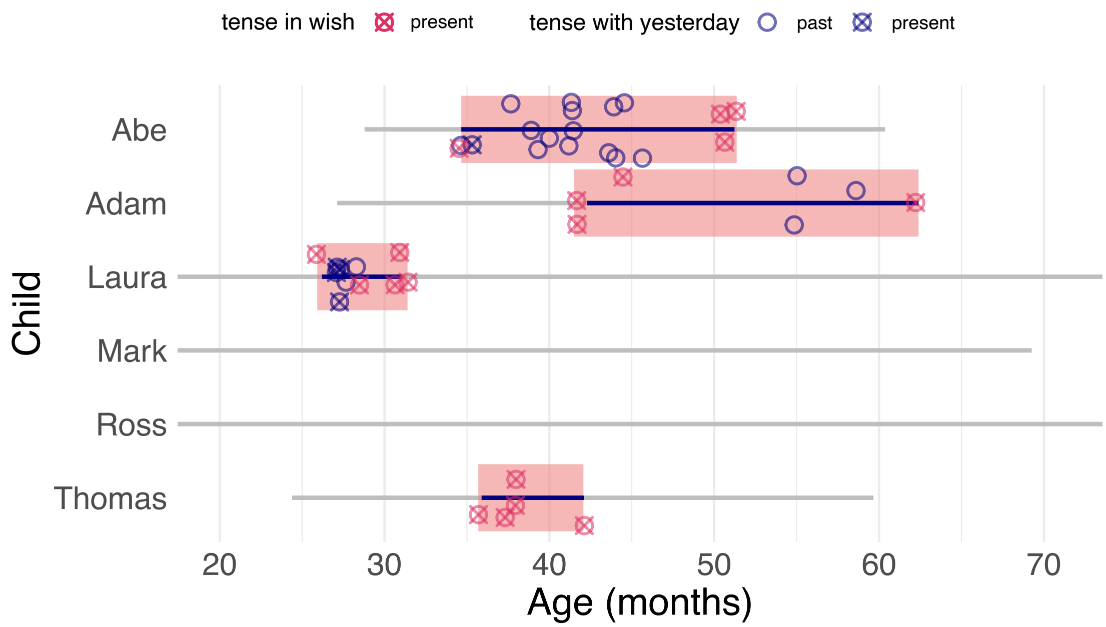

Figure 3.4. Overview of children's productivity with the past tense. Pink rectangles indicate the time span in which individual children (y-axis) produced wishes with tense errors. Each instance of a present-for-past error in wishes is displayed as a pink crossed circle. Within the error span, we plotted the tense of utterances with yesterday with blue circles (crossed means present tense was used). Blue lines within the error span indicate the time span over which we found instances of overregularization (e.g., "I putted"). Grey line indicates recording span. See Appendix S3.3 for corresponding numeric information in table format.

5. DISCUSSION

In this paper we examined the first language acquisition of counterfactual utterances, with our main focus on the development of children's wishes. We conducted corpus research that consisted of two parts. First, we extracted all child and adult utterances containing the word *wish* from eligible corpora on CHILDES and coded for various syntactic and semantic variables. We provided a detailed overview of children's *wish*-constructions and compared the properties of *wish*-utterances produced by children and adults. Second, we took a closer look at the longitudinal linguistic development of 6 children and investigated the maturation of their counterfactual language, comparing their usage of counterfactual wishes and conditionals. With this research we addressed two questions related to form-to-meaning mapping. First, we asked whether children go through a stage where they map the counterfactual's "fake" past morpheme to actual past temporal orientation, and consequently generate present tense inflected verbs in their own productions of present counterfactual constructions. Second, we asked whether linguistically more transparent counterfactual constructions (wishes) are acquired before the more complex counterfactual conditional. The combined results of our corpus work show children indeed go through a stage where they productively use present tense in the complement of counterfactual wishes, diverging from their adult input. We also found that children start using wishes around age 3;2 (onset ranging between 2;1 and 4;0), which is before the average onset of counterfactual conditionals around age 3;7 (range between age 2;8 and 4;4). Below we discuss these questions and findings in more detail, as well as some limitations to this work and suggestions for future research.

5.1. Children's Counterfactuals Contain Present-for-past Errors

The first question addressed in this study was whether children go through a phase where they make tense-marking mistakes in the complement of counterfactuals. Acquiring counterfactual utterances requires discovering that the past tense in its complement/antecedent is "fake" and marks counterfactuality instead. This mapping between counterfactuality and the past tense morpheme is thought to require complex semantic operations (Iatridou, 2000; Karawani, 2014; Ritter & Wiltschko, 2014) and a higher syntactic position (Bjorkman, 2015; Bjorkman & Halpert, 2017; Ritter & Wiltschko, 2014). Since children have to see through the "fakeness" of the past tense in order to learn this mapping, we hypothesized that children would productively form counterfactual wishes that have a present tense (rather than past tense) marking on the embedded matrix verb, as this aligns with the temporal orientation of a present wish. Indeed, we found that children make a substantial amount of past tense errors (11% of total wishes), most of them between ages 2-4 (75.6%). We observed these errors both in wishes that were judged to have clear evidence for a true counterfactual usage, and in wishes that were less clearly adult-like for counterfactuality. The fact that we observed present tense in clear counterfactual wishes, suggests children do not need the "fake past" to express counterfactual meaning. Instead, it's possible they mapped counterfactual meaning directly to the verb *wish*. The fact that you can express counterfactual meaning without relying on the "fake" past is consistent with cross-linguistic typology for counterfactual constructions: there are languages that express counterfactuality without making use of tense-marking, e.g., Mandarin Chinese (Jiang, 2019; Yong, 2016). This is also consistent with the fact that we observed some past counterfactuals productions with only one layer of past marking (18/19).

One could wonder whether the tense errors found in the complement of *wish* could be due to children not yet having acquired the past tense form in general. This seems unlikely, as children generally have productive past tense usage before age 3 (Brown, 1973; de Villiers, 2000; Kuczaj, 1977). For example, Abe acquired past tense with a 90% success rate by age 2;9, right before his first counterfactual wishes occur (Kuczaj, 1977). For three children, we showed that they display clear signs of productive tense marking during the period in which they make tense marking errors in counterfactual constructions. They use past tense in utterances with *yesterday* and overregularize the past tense morpheme to irregular verbs, showing productive usage. Only for the youngest *wish*-producer, Laura, do we find some tense marking errors outside counterfactual constructions, suggesting that her earliest errors (before 28 months) might be partially due to a general problem with applying past tense inflection. Another explanation for the tense errors could be that children actually use a bare verb construction (rather than present tense) because they treat *wish* analogously to the semantically related desire verb *want* (which selects for a non-finite complement). Or, they may be omitting the auxiliary verb *would* in future wishes, which is plausible as it is often pronounced in reduced form. However, from the 41 errors only 15 (37%) are compatible with a bare verb/dropped *would* explanation, which suggests that this cannot be the sole reason for children's past tense errors. Most tense errors in wishes are thus due to productive present tense marking, counter to the examples children receive in their input.

5.2. Children Start Producing Wishes before Conditionals

The second aim of this corpus study was to find out whether counterfactual wishes are acquired before counterfactual conditionals. Since *wish* is a dedicated marker of counterfactuality in English when it associates with propositional content, we hypothesized that counterfactual wishes would be easier to acquire than counterfactual conditional constructions. Indeed, we found that children produced the *wish*-construction either before or simultaneously with counterfactual conditionals. Counterfactual wishes mostly seem to emerge between age 2 and 4, while counterfactual conditionals emerge between age 2.5 and 4.5. However, it should be noted that there is a wide range of variation between children and the presence and size of the gap between the onset of wishes and conditionals. Some children acquire wishes before conditionals with an onset gap ranging from half a year to a year, while other children start using both constructions around the same time. We also indicated the need to be cautious not to equate using the *wish*-construction with having the ability to reason counterfactually about the world. Indeed, children's early wishes do not always seem adult-like. Especially children under age 3 seem to use the *wish*-construction to express direct desires (much like the verb *want*), and it is unclear whether they know *wish* can only be used when you believe this desire to be unobtainable or unfulfilled. We start finding clear indication of wishes with unequivocal counterfactuality (based on contextual and lexical information) between age 2.5 and 5, and for counterfactual conditionals this range is 2.5 to 4.5. While some children display a long gap between using clear counterfactual wishes and conditionals (ranging from 3-16 months), other children use clear counterfactual conditionals before wishes (difference ranging from 2 to 6 months). However, it should be noted that the distinction of "clear" versus "unclear" is difficult to make and relied on the coder's interpretation. As discussed before, the coding was done conservatively to reduce the chance of overinterpreting the counterfactuality of an utterance, which thus means we might be underestimating the counterfactuality of utterances we deemed "unclear". If we take our findings at face value however, they suggest that the *wish*-construction is generally acquired before or simultaneous with the counterfactual conditional. While it's not clear whether children always use the construction in an adult-like way, at least some children also display this pattern in the onset of clear counterfactual wishes and conditionals.

Crucially, it is unlikely that the difference we observe between the acquisition of counterfactual wishes and conditionals is solely due to the difference in causal structure (i.e., the *if..then* relationship in conditionals). While intuitively, conditionals are harder to process because they rely on linking two clauses with a causal relation, we actually find that most children start producing the non-counterfactual conditional structure (e.g., hypothetical future) before age 3 (Kuczaj & Daly, 1979; Reilly, 1982). Since most children start producing wishes after age 3, the difficulty of the conditional structure itself is not holding them back from acquiring the counterfactual conditional at that time. Another question that might arise is how accurate the ages of acquisition are that we found for the different constructions. Since corpus data is sampled and only includes a small proportion of the actual spoken input and output of the child, there is always the risk that we have missed earlier occurrences of either the wishes or conditionals. However, since the density of the used corpora was high (recording 1-5 times a month), the sample size of the observed constructions fairly similar (we observed 149 wishes and 104 conditionals) and the onset difference we observed quite large (6 to 12 months), we believe it to be unlikely that the onset differences we observed are due to unequal sampling.

5.3. Bootstrapping of the "Fake" Past Tense

When looking at the longitudinal data of six children we observed a noteworthy pattern. For 4/6 children, present-for-past substitutions did not occur after the onset of the counterfactual conditional. For half of them, this was simply because they were never observed making any present-for-past errors. This finding is compatible with a scenario where children first start to use the counterfactual *wish*-construction without have now discovered the relation between the "fake" past and the expression of counterfactual meaning. Then, once children successfully figure out this mapping, they cease using the present tense in wishes. Since they have now acquired the mapping between "fake" past and counterfactuality, they can start observing it in other environments, i.e., the counterfactual conditional, allowing them to attribute counterfactual meaning to the conditional construction as well. In other words, it is possible that the dedicated *wish*-construction in English bootstraps the acquisition of the "fake" past, which in turn facilitates learning the counterfactual conditional. However, there are children (i.e., Abe and Adam) that do not follow this pattern. Abe starts using the counterfactual conditional before the end of his present-for-past error period. Notably, Abe also participated in a longitudinal study investigating the development of hypothetical conditionals (Kuczaj & Daly, 1979), so this could have accelerated his acquisition of the counterfactual conditionals compared to other children. For Adam, the recordings ended before we could determine whether his unexpected present-for-past error at age 5 was an unremarkable slip-up or a continuation of his error period. Since we only had longitudinal data available for a small subset of children, we cannot draw any hard conclusions from this sample about this bootstrapping hypothesis. A fully analogous argument has been made for dedicated epistemic adverbs like *maybe* as potentially helping children learn the more complex variable-meaning modal verbs like *may* or *must* (i.e., auxiliaries with both epistemic and deontic (or other root modality) meanings) (Cournane, 2021).

5.4. Considerations and Future Directions

In this paper, we have investigated the acquisition of counterfactual constructions from a form-to-meaning mapping perspective and argued that the linguistic complexity of the counterfactual constructions contributes to its relatively late acquisition. The thought that complexity of linguistic structures plays a role in the emergence of such structure in children's speech is by no means original (Cournane, 2021; Reilly, 1982). For example, Reilly summarizes the relationship between cognitive and linguistic complexity as follows: "*Language and cognition are independent yet interactive systems where cognition is basically responsible for the sequence of acquisition, but it's the linguistic complexity of a structure that determines when that structure will appear in a child's grammar*." (Reilly, 1982, p.xi). We view the process of acquiring counterfactual constructions in a similar way. In order to communicate counterfactuality, children need to have reached certain developmental milestones, including the ability to hold multiple possibilities in mind (Leahy & Carey, 2019) and considering a false possibility temporarily true (Beck, McColgan, et al., 2011; Byrne, 2007). However, the onset of a linguistic construction also depends on various factors, including its linguistic complexity. Specifically, we argue that constructions that are dedicated to expressing counterfactuality (wishes in the case of English) should help children to detect these constructions in their input, and in the case of English, help discover the link between counterfactuality and the "fake" past tense.

In the future, this hypothesis can be tested by looking at other dedicated counterfactual constructions in other languages and comparing their acquisition onset with that of multi-purpose constructions. If having a dedicated counterfactual construction (such as the wish-construction) indeed facilitates the discovery of the mapping of counterfactual meaning to the "fake" past, we expect this pattern to hold for other languages as well. As mentioned before, the amount of data we extracted was relatively small, considering that we looked through all eligible corpora available on CHILDES. Since the natural occurrence of counterfactual constructions is fairly uncommon, future research directly targeting questions about "fake" past-tense usage might want to consider an elicitation task to elicit counterfactual speech, especially when working with languages that have relatively little (or no) corpus data available.

6. CONCLUSION

All in all, our findings show that counterfactual constructions are not only challenging because they require complex reasoning, but also because they involve complex form-to-meaning mapping. English children productively make tense errors predicted by semantic accounts of counterfactuality that analyze the past tense in these constructions as “fake”, and they acquire the more transparent counterfactual *wish*-construction before counterfactual conditionals. We also put forward the idea that dedicated counterfactual wishes provide more transparent cues to the temporal mismatch present in counterfactual constructions, and possibly help bootstrap the acquisition of the “fake” past tense. Future research should investigate how dedicated and undedicated counterfactual constructions are acquired in other languages as well.

CHAPTER 4: THE PAST IS "FAKE": FACILITATED PROCESSING OF WISHES COMPARED TO COUNTERFACTUAL CONDITIONALS

Maxime Tulling, Mark Bacon & Ailís Cournane (in preparation)

1. ABSTRACT

Understanding counterfactual constructions such as "If an asteroid hadn't hit the earth, dinosaurs would still be alive" involves the ability to hold multiple possibilities in mind while considering a false possibility temporarily true. While there is much literature on the development of counterfactual reasoning, there is disagreement about when exactly counterfactual thinking is fully developed, and most studies confound cognitive complexity with both linguistic and task complexity. In this study, we investigated the comprehension of counterfactual constructions in 4 and 5-year-olds (N=23) and adults (N=30) using a referent selection task, varying the linguistic complexity of the counterfactual construction. Additionally, we investigated whether children ever interpret the "fake" past tense morpheme in present counterfactual constructions as if it has an actual past temporal orientation. Our results show that children's performance on the less complex counterfactual *wish*-construction exceeded their performance on counterfactual conditionals. This suggests that children's understanding of counterfactual utterances is heavily dependent on the linguistic complexity of the structure, and that a failure to comprehend counterfactual conditionals may be caused by linguistic challenges, rather than an inability to reason counterfactually. Additionally, we found some suggestive but inconclusive evidence that the comprehension of counterfactual utterances may be further complicated by the presence of the "fake" past tense marking, which sometimes gets interpreted as being real.

2. INTRODUCTION

Counterfactual expressions such as the past counterfactual conditional, e.g., "If an asteroid hadn't hit the earth, dinosaurs would still be alive", express alternative outcomes that are contrary to the actual state (asteroids have hit the earth, and dinosaurs are not alive). Counterfactuals therefore allow us to express what could have been. Since counterfactual reasoning guides adaptive behavior and is employed when learning from the past (Byrne, 2007a; Van Hoeck et al., 2015), there has been increased interest in its development. There is disagreement, however, about when counterfactual thinking is fully developed. Some argue that this ability emerges during pre-school years (age 3-5) (Harris et al., 1996; Nyhout & Ganea, 2019; Robinson & Beck, 2000; Rouvoli et al., 2019), but others argue it is not fully developed until age 6 or 7 (McCormack et al., 2018) or even until preadolescence (age 11 or 12) (Rafetseder et al., 2010, 2013). An interesting asymmetry that has been found is that 3-to-4-year-olds seem to be better at hypothetical future questions (e.g., answering "If I draw on this piece of paper, which box will it go into?") compared to counterfactual questions (e.g., "If I had not drawn on the piece of paper, which box would it be in?") even though their structure, length and presence of an irrealis component are all comparably complex (Riggs et al., 1998; Robinson & Beck, 2000).

Counterfactual reasoning is argued to be cognitively more complex than hypothetical reasoning, since it requires us to hold multiple possibilities in mind while considering a false possibility temporarily true (Beck, Riggs, et al., 2011; Byrne, 2007a). The development of executive functions like working memory (keeping multiple possibilities in mind), inhibition, and attention switching have thus been related to the acquisition of counterfactuality (Beck et al., 2009; Beck, Riggs, et al., 2011, p. 20; Byrne, 2007a; Guajardo et al., 2009; Robinson & Beck, 2000).

Especially inhibitory control has been suggested to be at the core of difficulty in counterfactual processing, as repressing the actual world would require high levels of inhibition (Beck et al., 2009; Weisberg & Gopnik, 2016). In fact, Beck et al. (2009) found that measures of inhibitory control and receptive vocabulary were predictive of counterfactual performance.

It has thus been argued that the difficulty of comprehending counterfactual utterances is due to their cognitive complexity. However, most studies investigating counterfactual reasoning confound cognitive complexity with linguistic complexity by using past counterfactual conditionals in their tasks. These constructions are not only complex from a conceptual perspective, but also from a linguistic perspective, as the past counterfactual involves several components known to be challenging on their own, such as: perfect marking (Brown, 1973; Cromer, 1968; V. Gathercole, 1986), conditional structure (c.f. Reilly, 1982; Bowerman, 1986), and often also negation (Austin et al., 2014; Feiman et al., 2017). Using linguistically complex structures might thus have underestimated children's ability to engage in counterfactual reasoning (Nyhout & Ganea, 2019). The issue of linguistic complexity was also raised by Rouvoli et al. (2019), who investigated the acquisition of past counterfactual conditionals in relation to learning that the construction *if +past perfect* indicates counterfactuality. However, the past perfect conditional is not the only linguistic construction that can express counterfactuality, and in comparison to other counterfactual constructions it appears to be the last construction children acquire (Kuczaj & Daly, 1979; Reilly, 1982). In spoken language, past counterfactual constructions are more infrequent than other counterfactuals (Crutchley, 2013) and adolescents and even some adults display some difficulty with their use (Nippold, Nehls-Lowe, et al., 2020; Nippold, Shinham, et al., 2020). One way forward, to de-confound children's counterfactual reasoning abilities from linguistic complexity, is for researchers to use linguistically less complex counterfactual constructions.

As far as we are aware, there has been no systematic study yet breaking down counterfactual constructions into simpler parts to see how this affects children's ability to understand them. The aim of this study is to do exactly this. Here, we investigated young children's understanding of present counterfactual constructions, narrowing down the acquisition challenge to the mapping of counterfactuality onto linguistic structure, and varying linguistic complexity while keeping cognitive task demands similar.

2.1. Different Types of Counterfactual Constructions

While prior research on the acquisition of counterfactuality has mainly used past counterfactual constructions (1a), i.e., utterances that reason about an alternative past, counterfactuality can also be expressed about the present (1b). In both cases, the utterances reason about a possibility that is simultaneously asserted not to be true.

(1a) If he **had** had money back then, he would have bought a car PAST CF

→ he had no money back then (and hasn't bought a car)

(1b). If he **had** money right now, he would buy a car. PRESENT CF

→ he has no money right now (and is not buying a car)

This counterfactual interpretation is thought to arise through reinterpretation of the counterfactual's past tense morpheme. The conditional utterances in (1) include so called "fake" past tense morphemes (Iatridou, 2000). This past is considered "fake" because there is a mismatch between the morphological tense marking, and the temporal orientation of the utterance. That is, (1b) refers to an alternative state in the present (reinforced by the temporal adverb *right now*), yet the utterance contains past tense morphology (on *had*), which usually refers to a past temporal orientation. Rather than expressing regular past tense meaning, theoretical linguistic accounts propose the "fake" past contributes to expressing counterfactuality, either by going back in time to gain access to an alternative world (Arregui, 2009; Ippolito, 2006; Ogihara, 2000; Romero, 2014) or through scoping over possible worlds rather than over time (Iatridou, 2000; Karawani & Zeijlstra, 2013; Ritter & Wiltschko, 2014). In (1a) there are two levels of 'pastness', one 'real' past tense (the perfect construction) and one layer of "fake" past tense (bolded *had*) giving rise to a past counterfactual interpretation.

Conditionals consist of two clauses: the *if*-clause (antecedent) and the *then*-clause (consequent) which is logically dependent on the antecedent. In order to process conditionals one thus has to keep in mind and causally relate two clauses (c.f. Reilly, 1982; Bowerman, 1986). However, counterfactuality can also be expressed through non-causal constructions. Take for example the counterfactual *wish*-construction (2). Similar to counterfactual conditionals, counterfactual wishes contain the "fake" past tense. Counterfactual wishes express the desire for something to be the case, while simultaneously asserting that it is not the case.

(2). I wish I **had** a car right now.

→ I don't have a car.

Counterfactual wishes are thus grammatically less complex than counterfactual conditionals. This is in part because wishes lack the causal dependency relationship, but is also because the *wish*-construction is more transparent in its mapping from construction to interpretation. In the standard variety of English, *wish* is a dedicated counterfactual marker since it only embeds propositions that are counterfactual. For this reason, it cannot occur with a present tense complement, even when the temporal orientation of the wish is present tense (3). This is in contrast with counterfactual conditionals, which can appear with present (4a), real past (4b) and "fake" past tense (4c) in their antecedent.

(3). *I wish I have a car. (Iatridou, 2000, 25)

(4a). If he **has** time to bake cake, he *will* bring some.

(4b). If he **had** time to bake cake, he *will* bring some.

(4c). If he **had** time to bake cake, he *would* bring some.

This dedicated form-to-meaning mapping of the *wish*-construction to counterfactuality could facilitate first language acquisition in two ways. First, the consistent mapping between *wish* + proposition and the expression of unfulfillable desires, could cue children into understanding that the construction is linked to communicating counterfactuality. Relatedly, the obvious mismatch between the wish's temporal orientation and morphological marking could help the child detect the "fake" past and its role in expressing counterfactual meaning.

Another way in which the dedicated form-to-meaning mappings in *wish*-constructions could facilitate the comprehension of counterfactual constructions has to do with online processing. It is a well-known fact that children up to 8 years-old are often unable to revise their initial reading of temporarily ambiguous content (Weighall, 2008). This is also referred to as the "kindergarten path effect" in analogy with the temporary "garden path effect" observed in adults during the online processing of ambiguous sentence structures (Huang & Hollister, 2019;

Trueswell et al., 1999; Weighall, 2008). This phenomenon could pose a problem for understanding counterfactual conditionals, since distinguishing between a real past (4b) and a "fake" past (4c), relies on encountering *would* in the consequent, prompting the parser to revise an initial real past tense reading of *had* if they had one. Since *wish* is a dedicated marker of counterfactuality, there is no room for a garden path effect in counterfactual wishes like (3).

To sum up, counterfactual conditional constructions are linguistically complex in multiple ways: they involve a complex dependency relationship, have complex form-to-meaning mapping, and compete with other (present and past) conditionals. Due to their dedicated counterfactual meaning, *wish*-constructions lack much of this complexity, though they still require children to see through the "fakeness" of their past tense marking and reason counterfactually. However, prior literature has not considered the acquisition of counterfactual wishes nor the challenge posed by the "fakeness" of the past tense in studies investigating the acquisition of counterfactual reasoning.

2.2. Background on the Development of Counterfactuality

Previous studies on the acquisition of counterfactual utterances have focused mainly on the development of counterfactual thinking. Early studies investigating this ability in 3 and 4-year-olds used tasks in which situations are described or played out, and a critical event happens. Children were then asked to reason about what would have happened if things had gone differently. These studies report conflicting results. For example, in Harris et al. (1996) one of the scenarios involved Carol, who comes into the kitchen wearing muddy boots, leaving dirty footprints all over the floor. Children were then asked: "What if Carol had taken her shoes off, would the floor be dirty?". They reported that both 3 year-olds and 4-year-olds performed well above chance (answering ‘no’). Riggs, Peterson, Robinson, & Mitchell (1998), however, found that 3 and 4-year-olds were not able to pass their counterfactual task. In their task children heard short stories, where for example a fireman (Pete) stayed home sick in bed until an emergency call came and he goes to a fire at the post office. In this scenario the child was asked: “If there had been no fire, where would Pete be?” – to which the target response is: “in his bed/at home”. Children answered with ‘realist errors’ describing how the world currently (Pete is at the fire). Such ‘realist’ errors have also been reported in other research, labelled as an ‘actuality bias’ (Rouvoli et al., 2019).

These conflicting results led to much subsequent research investigating why such differences arose, giving rise to various attempts to improve counterfactual reasoning tasks (German & Nichols, 2003; Guajardo & Turley-Ames, 2004). Some replicated Rigg’s et al.’s finding that 3 year-olds perform poorly, and conclude that children significantly improve between age 4 and 5 (Guajardo et al., 2009; Guajardo & Turley-Ames, 2004). Others found that 3 year-olds perform above chance if you reduce the amount of information children have to maintain in mind (German & Nichols, 2003). One crucial difference between Harris et al. (1996) and Riggs et al. (1998) is the demand put on the child’s cognitive reasoning. In Harris et al. (1996) the counterfactual questions were polar questions, contrasting two alternatives directly. In Riggs et al. (1998) the counterfactual question was left open. Indeed Beck, Robinson, Carroll, & Apperly (2006) found that 3-to-4-year-olds can reason counterfactually about single events, but have trouble with counterfactuals that entertain multiple possibilities. According to them, the ability to acknowledge multiple possibilities and answer open counterfactual questions[9] develops around age 5 or 6.

The question of what exactly counts as counterfactual reasoning is a pervasive one in the literature, leading to conflicting reports on age of acquisition (Robinson & Beck, 2000), but some have also argued that previous studies overestimated children's ability to reason counterfactually by providing confounded tasks in which children could pass the task without using 'real' counterfactual reasoning (Rafetseder et al., 2010, 2013). Rafestseder et al (2010;2013) argue that children in the previously mentioned tasks could pass using 'basic conditional reasoning' rather than genuine counterfactual reasoning. In a conditional sentence like "If Carol takes her shoes off, the floor will be clean" it is the case that if the condition of the antecedent is met (Carol takes her shoes off), the consequent follows (a clean floor). With basic conditional reasoning one uses their general knowledge of causal regularities, social norms and experience (e.g., if you have clean feet → the floor will stay clean) to answer the task question ("What if Carol had taken her shoes off, would the floor be dirty"). To avoid the possibility of passing with just conditional reasoning, their study (based on Harris et al, 1996) involved a second cause of the outcome, i.e., rather than just having Carol and her muddy shoes, there is a second character (Max) who also walks in with muddy shoes. Now, when the child is asked what would have happened if Carol had taken her shoes off, only children who engage in real counterfactual reasoning should respond that the floor would still be dirty (since counterfactual reasoning involves respecting the nearest possible world[10]

---

[9] The authors acknowledged that the there is a confound with linguistic complexity, as the open counterfactual questions might also be linguistically more taxing than the closed questions (p.422).
[10] When considering a counterfactual world, you keep the alternative world almost identical to the real world, only changing the information that is causally dependent on the counterfactual utterance. E.g., when

(Lewis, 1973)). If children rely on basic conditional reasoning, they should answer the floor would be clean, ignoring the fact that in a minimally different counterfactual world Max would in fact still have been wearing dirty shoes. Rafetseder et al. (2010;2013) found that children were not able to reason this way until age 12. However, others have questioned these results (McCormack et al., 2018; Nyhout et al., 2017; Nyhout & Ganea, 2019) suggesting that children's performance was underestimated in Rafesteder et al.'s tasks, since the design allows for unwarranted inferences that could lead children astray (e.g., assuming that if Carol took her shoes off, Max would too). Using an alternative task, avoiding such confounds and providing clear causal structures in the physical domain, even 4-5-year olds are (again) argued to display mature counterfactual reasoning (Nyhout & Ganea, 2019). They were correctly able to answer questions such as "If she had not put the green block on the box, would the light still have switched on?". Three-year-olds performed at chance level, but the authors speculate that perhaps this is due to the grammar of counterfactual questions being too complex.

2.3. Questions and Predictions

As we have seen in the previous sections, prior research on the acquisition of counterfactuality suggests that children start acquiring the counterfactual conditional construction around age 3-6, with a cognitive leap around age 4 allowing children to comprehend and produce counterfactual utterances in an adult like way (Bowerman, 1986; Guajardo et al., 2009; Guajardo & Turley-Ames, 2004b; Nyhout & Ganea, 2019; Reilly, 1982). However, most tasks that tap into children's ability

---

evaluating the sentence, "If kangaroos had no tails, they would topple over", we do not usually consider worlds where kangaroos have crutches, and would be able to stand without a tail (Iatridou, 2000).

to think counterfactually involve the cognitively demanding task of actively imagining an alternative reality (based on the logic of stories or games with arbitrary rules) without any careful consideration of the linguistic content of the questions. Indeed, some researchers wonder if young children's problems with passing counterfactual reasoning tasks has to do with the linguistic complexity of the stimuli (Nyhout & Ganea, 2019). Researchers mainly use past counterfactual conditionals in their reasoning tasks (e.g., "What if he had driven the other way, where would he be?"), controlling for syntactic complexity by comparing children's performance on future hypothetical constructions such as "What if next time he drives the other way, where will he be?". However, while this comparison does control for the conditional causal structure and length of the sentence, counterfactual constructions are linguistically more complex than their hypothetical counterparts in their usage of "fake" past tense. In this study, we investigate the role of the "fake" past tense in the acquisition of counterfactual constructions. And ask the following questions:

1) Do young children display an earlier adult-like understanding of counterfactual wishes than of counterfactual conditionals?
2) Do children ever interpret the past tense morpheme in counterfactual constructions as a regular past tense?

Since counterfactual wishes in English are linguistically and cognitively less complex than conditional constructions, we expect the acquisition of wishes to be a better indicator of early counterfactual reasoning. In contrast to conditionals, wishes are dedicated counterfactual markers and lack the conditional structure which requires keeping in mind and linking two clauses. We hypothesize that these factors lead to children understanding counterfactual wishes before they understand counterfactual conditionals.

The past tense in counterfactual constructions is misleading ("fake") it does not actually express regular past tense. Figuring this out is a challenge for the acquisition of counterfactuals, and we hypothesize that children go through a stage where they misunderstand the "fake" past tense morpheme in counterfactual constructions as referring to an actual past temporal orientation. This hypothesis is compatible with findings from Reilly (1982), who observed that two- and three-year-olds deny or provide realist responses to counterfactual utterances. We predict that a past tense interpretation of the "fake" past leads to misunderstanding and/or influences children's interpretation of counterfactual constructions (e.g., gives rise to actuality biases). Some of the responses children provided to counterfactual questions suggest that they indeed interpreted the question as containing a real past tense (5).

(5a). Adult: What if you were a little girl? (Reilly, 1986, ex. 59, p.116)
3-yo: Now, I'm a big one.

(5b). Adult: What if you ate three boxes of strawberries? (Reilly, 1986, ex. 60, p.116)
3-yo: I did / I ate 'em already.

2.4. This Study

In order to investigate our research questions we chose to use a referent selection task, loosely base our study design on Rouvoli et al. (2019). More conventional counterfactual reasoning tasks require the child to answer a counterfactual question, and to do so they must actively generate alternative possibilities and reason about them (Beck, Riggs, et al., 2011; Harris et al., 1996; Nyhout et al., 2017; Rafetseder et al., 2013). With a referent selection task, children watch different characters play out a short, rule-based scenario (e.g., you get a medal for choosing healthy food and a cross for choosing candy), after which a narrator speaks about one of the characters. The children are asked to interpret a counterfactual utterance, like "If he had eaten a watermelon, he would have got the medal" and decide which character the narrator is talking about. Such a task reduces linguistic and executive function demands (Rouvoli et al., 2019, p.553), and avoids the possibility of false passes based on conditional reasoning. Unlike Rouvoli et al. (2019), which focused solely on past counterfactual constructions[11], the focus of our experiment is on present counterfactual constructions and includes both counterfactual conditionals and wishes. By focusing on present counterfactuals, we can tune into the "fake" past tense aspect of counterfactuals, while varying linguistic complexity (wish/conditional). Our task modifies the referent selection task such that it involves an option that is compatible with a real past tense reading, allowing us to distinguish between an actuality-biased response (Rouvoli et al., 2019; Kazanina et al., 2019) and a past tense interpretation.

3. METHODOLOGY

3.1. Participants

3.1.1. Adults

45 adult American English-speaking participants were recruited on the experimental platform Prolific (www.prolific.co). We excluded data from 2 participants for not passing control items and 13 for not being native speakers of English. The age range of the remaining 30 participants was 18-63 (M=33.6, SD=12.8).

---

[11] They do mention both present counterfactuals and counterfactual wishes (Rouvoli et al., 2019, footnote 1 and 2), leaving them up for future investigation.

3.1.2. Children

We recruited 28 children between the ages 4 and 6 living within the United States, to participate over the online conferencing platform Zoom (https://zoom.us/), using the secured Zoom license of New York University. All children had American English as their dominant input language and had no reported history of atypical language development. Electronically signed written consent was obtained from caregivers (present during the experiment) before the experimental session started, and verbal assent was obtained from children. Zoom sessions were recorded if the caregiver provided consent for this. We excluded data from 3 participants for failing practice or control items, and the data from 1 participant who did not complete the experiment. The age of the remaining participants 24 participants ranged from 51-74 months (M=59.2, SD=7.2).

3.2. Design and Procedure

3.2.1. Materials

In this experiment we showed children short, animated scenes taking place inside a milkshake bar. The experiment revolved around five characters: three identical 'kippies' (silent animal-like creatures), a milkshake man (male voice) and a narrator (off-screen, female voice). We chose novel identical characters to avoid any biases or predispositions that might influence children's rationale in their selection of a referent. The experiment was designed using JavaScript with the PennController (Zehr & Schwarz, 2018) and hosted online on IBEX-farm (Drummond, 2010). Audio and images were hosted on the first author's personal website. All images for the experiment were created with digital drawing software (ArtRage 5, Ambient Design), and animated with stop motion intervals of 75-100ms. The voices of the narrator and milkshake man were recorded by native speakers of American English. The recordings were selected to sound natural and engaging, and amplitudes were normalized to 65 dB in Praat (Boersma & Weenink, 2022).

3.2.2. Design

Participants completed an animated referent selection task testing the interpretation of present counterfactual conditionals and wishes, particularly tailored to detect an actual past tense reading that children might entertain for counterfactual utterances.

In our task, participants were first introduced to three identical characters "kippies" who love milkshakes. They were also introduced to a milkshake man who sells milkshakes for coins matching the flavor (chocolate, strawberry or banana) of the chosen milkshake. Participants were informed that the milkshake man has a strong preference for banana coins and were asked two comprehension questions to verify that they understood the rules of the game. The questions were phrased as hypothetical conditionals, e.g., "If this kippie wants to buy a chocolate milkshake, what does he have to pay?" to make sure participants understood (non-counterfactual) conditional constructions. Participants were asked to select (by touch or mouse click) the correct coin and were prompted to try again if they were wrong. What followed was an explanation of the main task, showing three kippies buying milkshakes from the milkshake man. The participant was informed that every time the kippies bought a milkshake, the milkshake man would say something about one of them. The participant was also told that the milkshake man sometimes talks about what is happening right now but that sometimes he imagines how things could be different instead. This information should prepare the participant for the possibility of non-actual utterances. After this explanation, two practice trials followed to make sure the participant understood the task. For the first practice trial, the participant had to select the only kippie buying a chocolate milkshake when the milkshake man said, “Hmm, that kippie will have to give me a chocolate coin”. For the second practice item, participants heard an utterance that was non-actual. In this scenario, two kippies got a strawberry milkshake and the last kippie chose a banana milkshake. Then the milkshake man said, “Hmm maybe next time that kippie will pick a strawberry milkshake instead”. The narrator repeated the utterance and also added this: “Oh, that silly milkshake man got distracted thinking about how things could be different. Who’s the milkshake man talking about?”. Choosing the correct answer (the kippie with the banana milkshake) on these trials indicates that the participant understands that they should pay attention to the linguistic structure of the utterance, and not merely pick the milkshake flavor that is mentioned. Participants got two tries to select the right character. With children, the experimenter would repeat the question or clarify the scenario if a child picked a wrong answer. When both practice questions were answered correctly, the real experiment started. If children were not able to pass any of the practice trials, the experiment was ended prematurely.

In the subsequent target trials, an animation showed each kippie buying a milkshake from the milkshake man. The structure of these trials was always the same. Kippies would start out at the bottom of the screen, with the milkshake bar at the top of the screen. Participants were informed that “some kippies already chose a milkshake, let’s see what happens next!” and then an animated scene followed where each kippie moved forward and performed an action with a milkshake. One of the kippies started out without a milkshake, and then picked one, e.g., banana flavor (actual referent). Another kippie was already drinking a milkshake (with the same flavor that the other kippie had picked) when the scene started, and finished it before paying, leaving the empty milkshake behind (past referent). The other kippie started out without a milkshake and then picked a milkshake flavor different from the other two (counterfactual referent). Once all the kippies had moved forward, they took out a coin matching the flavor of their milkshake, anticipating an upcoming transaction. However, before the kippies paid, the milkshake man produced a control or target utterance (actual, past, counterfactual) about one of the kippies (Figure 4.1A).

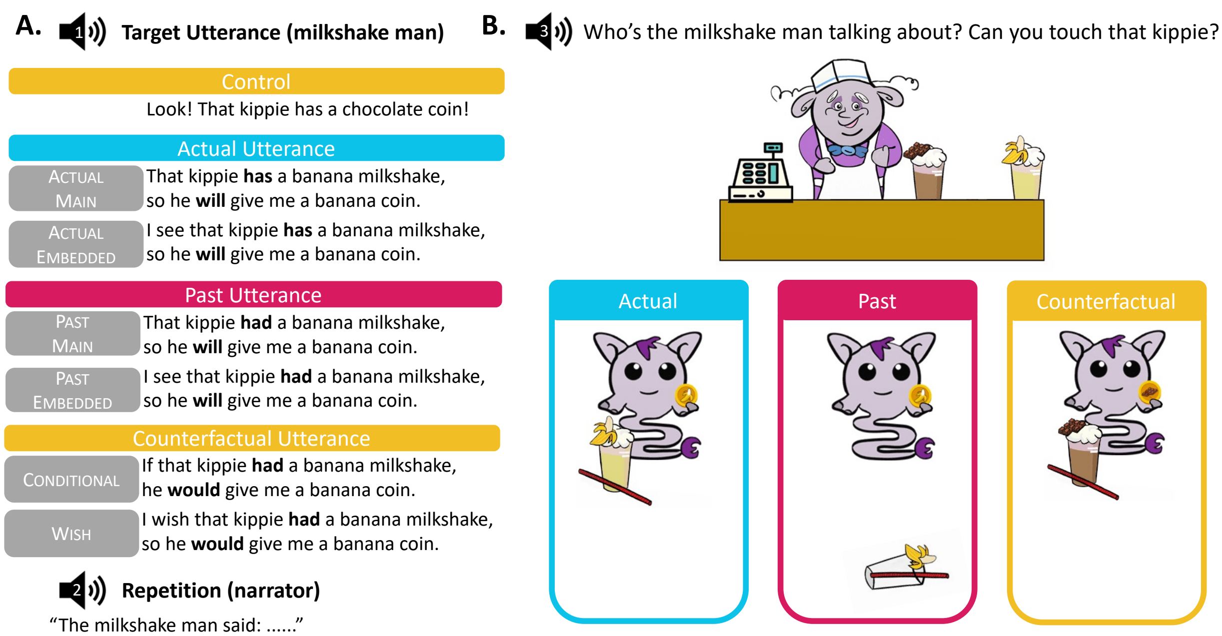

Figure 4.1. **A.** Example stimuli set and **B.** overview target responses. The milkshake man says one of the target utterances (control, actual, past or counterfactual) that refers to one of the characters (actual, past or counterfactual interpretation). Each target utterance is repeated once by the narrator.

In the ACTUAL UTTERANCE condition, the milkshake man described the current state of a kippie, e.g., "That kippie **has** a banana milkshake, so he **will** give me a banana coin". In the PAST UTTERANCE condition, the milkshake man described the past state or action that had just taken place, e.g., "That kippie **had** a banana milkshake, so he **will** give me a banana coin". The present and past condition were added to the experiment to see whether the intended past tense referent and actual referent were interpreted as such. For the COUNTERFACTUAL UTTERANCE condition, the milkshake man described how the situation could have been different e.g., “If that kippie **had** a banana milkshake, he **would** give me a banana coin”. The motivation for using such a counterfactual utterance was provided by the prior context stating that the milkshake man loved banana coins. For this reason, counterfactual trials were always about the banana milkshake. For the remaining trials, the three milkshake flavors were balanced. The order in which the different referents (actual/past/counterfactual) performed the action on screen (as well as where they appeared on screen) was balanced across the experiment.

The counterfactual utterance condition was subdivided into COUNTERFACTUAL CONDITIONALS and COUNTERFACTUAL WISHES (“I wish that kippie **had** a banana milkshake, so he **would** give me a banana coin”). We matched the overall grammatical structure of the conditional and wishes in the ACTUAL and PAST UTTERANCE conditions by including main clause sentences, i.e., “That kippie has a banana milkshake, so he will give me a banana coin” (matching the conditional clause structure) and embedded clause utterances “I see that kippie has a banana milkshake, so he will give me banana coin” (matching the wish structure). The stimuli all contained two clauses that were causally related (*if… then* or *p, so*), even though *wish* does not require this, We included the causal *so*-clause to balance our stimuli such that the comparison between COUNTERFACTUAL WISHES and COUNTERFACTUAL CONDITIONALS focused on syntactic/semantic complexity rather than utterance length or causal linking.

When hearing a counterfactual utterance like “If that kippie **had** a banana milkshake, he **would** give me a banana coin”, a child could have three possible interpretations: 1) An actual interpretation, linking the main referents (kippie → banana milkshake/coin) without paying much attention to the sentential structure; 2) a past tense interpretation (where the kippie had drunk his banana milkshake prior to the utterance); or 3) an adult-like counterfactual interpretation (the kippie does not have a banana milkshake). The three different possible referents introduced by our scenario each match one of these interpretations (Figure 4.1B).

The narrator repeated the target utterance once and then the participant was asked to select the kippie the milkshake man was talking about by touch (children) or click (adults). No feedback was provided for target items. Each trial ended with a between-trial screen where a cash register would sound and the narrator concluded "Yay! All the kippies have paid. Let's do this again!". The experimental procedure is summarized in Figure 4.2.

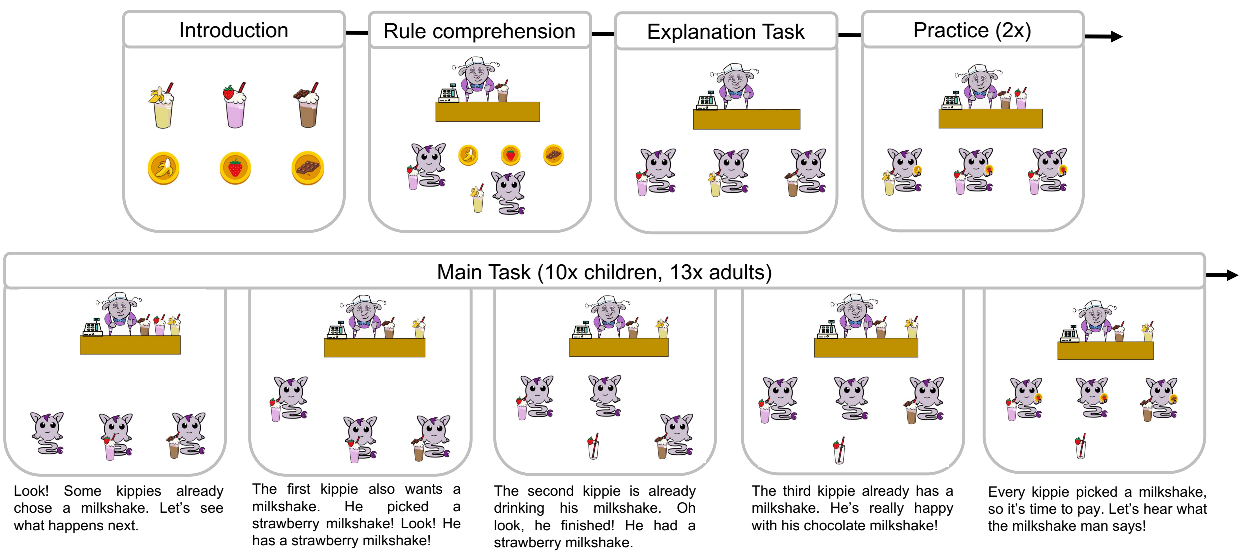

Figure 4.2. Experimental procedure with example stills from each animated scene. In the practice and experimental phase, each kippie moved forward separately performing one action (picking a milkshake from the counter, keeping the milkshake they already have or drinking the milkshake they have) before moving towards the milkshake man to pay.

3.2.3. Procedure

3.2.3.1. Adults

For adults, the experiment was hosted online on Ibex Farm, accessed through Prolific, and displayed on the participant's personal computer or android device. Depending on their device, they selected the correct response by mouse click or touch. Participants were instructed that they would be participating as adult controls in a psychology experiment created for children, and that they needed access to clearly audible sounds in order to participate. The online experiment took about 12-18 minutes. Each participant received $4.00 for their participation. The adult version of this experiment contained 13 trials: 6 counterfactuals (3 wishes and 3 conditionals), 3 controls, 2 actual trials and 2 past trials. The order in which items appeared was randomized. At the end of the experiment participants were given the opportunity to leave comments. Before the experiment took place participants filled in a short questionnaire indicating their age and language background.

3.2.3.2. Children

For children, the experiment was hosted locally on the experimenter's computer. To resemble an in-person experience, the experimenter conducted the experiment live with the child (in company of a caretaker) over video connection, sharing the experiment screen and providing remote access to the participant who zoomed in from a touchscreen tablet or pc. This way, the child could make their responses by touching their screen, and the experimenter could communicate with the child controlling the pace of the experiment and ensuring the child was engaged and understanding the task. In between trials, the experimenter would ask children to explain their responses (at least once for each condition) and tracked experiment progress via a digital trial tracker, where children got to pick an animal "sticker" to fill an empty spot on a digital sticker sheet. Children heard 10 target trials: 4 counterfactual items (2 wishes and 2 conditionals), 2 controls, 2 actual, and 2 past trials. The trials were grouped into four different blocks: Control, Counterfactual Wish + Counterfactual Conditional, Past and Actual. The items were randomized within these blocks and started with the Control block. The remaining trials were shuffled such that sequences of items were picked from these blocks in the order presented above, maximizing the distance between trials within the same block. This way children first got two trials from the Control condition, then trials from the Counterfactual condition, then a Past item, and lastly an Actual item, after which this sequence repeated. We used this order to make sure the children got some confidence with the task (starting with the 'easy' control items that require you to pick a referent that matches the actual world) and to make sure that the past condition followed the counterfactual (as to not bias their initial counterfactual response with a prior past tense trial).

The complete experiment took about 20-30 minutes. Before the experiment started, children's caregivers completed a language background questionnaire. We acquired data about the child's age, gender (male, female or non-binary), which languages children were exposed to, attestation of (a)typical development and education level of parents/caregivers. All Zoom sessions were recorded (if consented for by the caregivers), and children received a digital e-book about kippies and counterfactual constructions for their participation.

3.3. Data Analysis

3.3.1. Adults

Response and response accuracy was recorded for each individual trial, and the dataset was anonymized of any identifying markers (i.e., Prolific ID). Since accuracy is a binomial variable, we modeled the probability of making an error with a generalized linear mixed-effect model (GLMM, Baayen et al., 2008). Our maximal converging model included Main Utterance Condition (Actual, Past or Counterfactual) Construction Type (main/conditional or embedded/wish) and its interaction as fixed effects. As a random intercept we included participant identity. The model fit (logit link) was estimated by maximum likelihood using the default setting of LaPlace approximation. To test the contribution of our fixed effects, we performed a likelihood ratio test comparing our model and a nested model leaving out the variable of interest. We used the glmer-function from the 'lme4' package available on R to perform our analysis (Bates et al., 2015; R Core Team, 2021), and the 'emmeans' package to estimate the marginal means from the model (Russell, 2022). We performed a post-hoc pairwise comparison using the estimated marginal means to compare between the levels of our predictor variables. Then, we calculated the percentage of responses per utterance condition of interest (Actual, Past, Counterfactual Conditional and Counterfactual Wish), and visualized the responses per condition, to inspect what types of errors participants made.

3.3.2. Children

Response and response accuracy was recorded for each individual trial. In addition, we transcribed children's responses to follow-up questions about their reasoning for picking the chosen referent, if this data was recorded (recordings were missing for 3 of the 28 participants) and if the child chose to respond to this query (some children refused). We excluded one trial which got lost due to a bad zoom-connection. Upon inspecting children's individual response patterns, we further excluded data from one participant (z027) who had provided the same (counterfactual) response for each trial (target and controls), employing the strategy of always picking the referent that had a different milkshake from the other two kippies. For the remaining 23 participants, we modeled accuracy over different conditions with a similar generalized linear mixed-effect model as described for adults, with the inclusion of age in months (centered and z-scored) as an additional fixed effect. We similarly performed a post-hoc pairwise comparison using the estimated marginal means to compare between the levels of our predictor variables and visualized error patterns. In addition to this, we qualitatively inspected the rationale children had provided for their responses and described the non-adult like reasoning errors children sometimes seemed to make.

4. RESULTS

4.1. Adults

Adults displayed high performance on all conditions (see overview of mean accuracy in Table 4.1). We used a generalized linear mixed-effects model (GLMM), including participant as a random factor, to investigate whether accuracy was dependent on the experimental main condition (actual, past and counterfactual), construction type (main/wish or *embedded/conditional*) or interactions thereof. A likelihood ratio test comparing our model against a nested model without fixed effects, found that while the interaction of main condition and construction type was not a significant predictor of error rate ($\chi^2(2) = 2.96$, $p = .228$), main condition ($\chi^2(1) = 12.68$, $p < .001$) and construction type ($\chi^2(1) = 4.55$, $p = .033$) separately were. The model output of the maximal

GLMM is displayed in Appendix S4.1. A post-hoc pairwise comparison of the model's estimated marginal means (see Table 4.1 column 5) revealed that there was no significant difference between embedded and main construction types in the ACTUAL (OR 0.69, 95% CI 0.13-3.77, $p=0.67$) and PAST conditions (OR 0.63, 95% CI 0.17-2.40, $p=0.50$). For the COUNTERFACTUAL condition, we found that adults made significantly more errors for counterfactual conditionals than for wishes (OR 0.10, 95% CI 0.01-0.87, $p<.05$). These results are visually displayed in Figure 4.3A.

When we plot the mean accuracy for wishes against the mean accuracy for conditionals per participant (Figure 4.3B), we can observe that errors were rare, but not completely sporadic. Three participants made 2 or more errors on COUNTERFACTUAL CONDITIONALS. In contrast, we only found 1 error for the COUNTERFACTUAL WISH condition. To gain more insight into the nature of these errors, we plotted the response patterns of participants per utterance category in Figure 4.4A (collapsing main and embedded ACTUAl and PAST utterances together, as the difference between them was negligible). Notably, if we look at the type of errors participants made on COUNTERFACTUAL CONDITIONAL trials, in all cases participants chose the past referent instead of the counterfactual one. We also observe that about 10% of the time the past referent was selected in ACTUAL utterances, and that about 20% of the time the actual referent was selected for PAST utterances.

Table 4.1. Adult's Mean Accuracy Scores and Estimated Marginal Means per Condition

Condition	Mean Accuracy	SD	# Responses	Estimated marg. means	SE	CI
Actual Main	0.90	.305	30	0.93	.043	0.78-0.98
Actual Embedded	0.87	.346	30	0.91	.054	0.74-097
Past Main	0.80	.407	30	0.85	.073	0.65-0.95
Past Embedded	0.73	.450	30	0.78	.090	0.56-0.91
Counterfactual Wish	0.99	.105	90	0.99	.007	0.95-0.99
Counterfactual Conditional	0.91	.286	90	0.94	.027	0.86-0.98

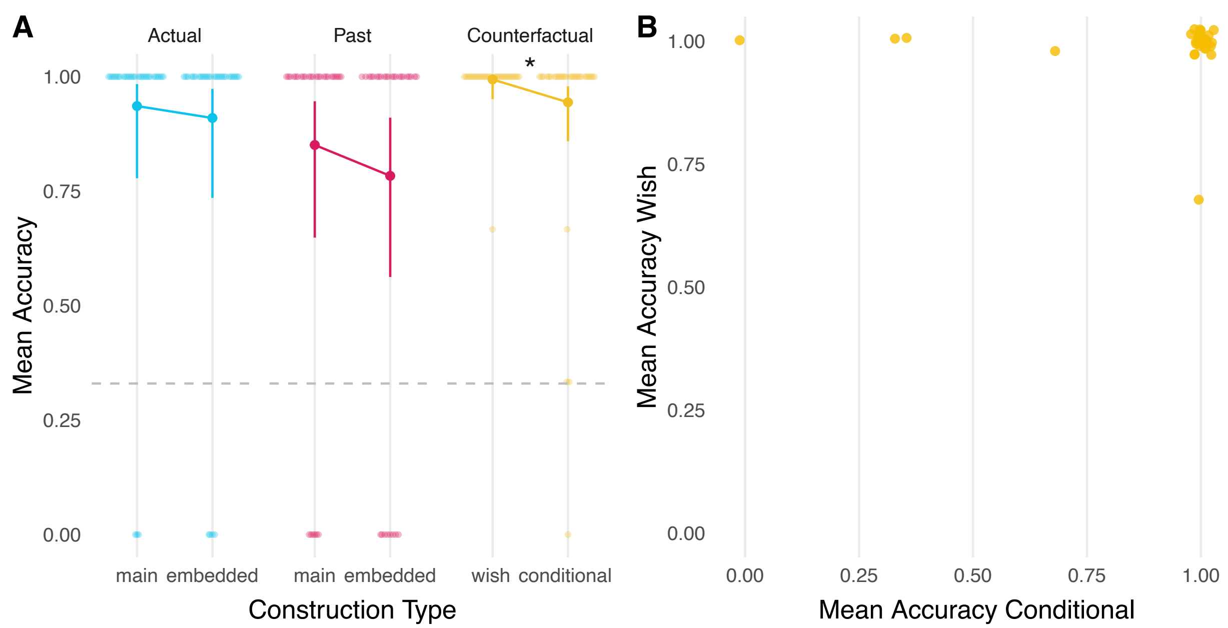

Figure 4.3. **A**. Adult's mean accuracy per condition (ACTUAL, PAST and COUNTERFACTUAL) and construction type (MAIN/WISH or EMBEDDED/CONDITIONAL). Raw mean accuracy per participant plotted as scattered dots in the background, estimated marginal means from generalized linear mixed model plotted in foreground, with bars indicating model-based standard errors. Asterisk indicates significance from post-hoc pairwise comparison. Dotted line indicates chance-level. **B.** Mean accuracy of counterfactual wishes (y-axis) plotted against the mean accuracy of counterfactual conditionals (x-axis) for each participant (scattered dots). Scattered dots are slightly jittered for the sake of legibility.

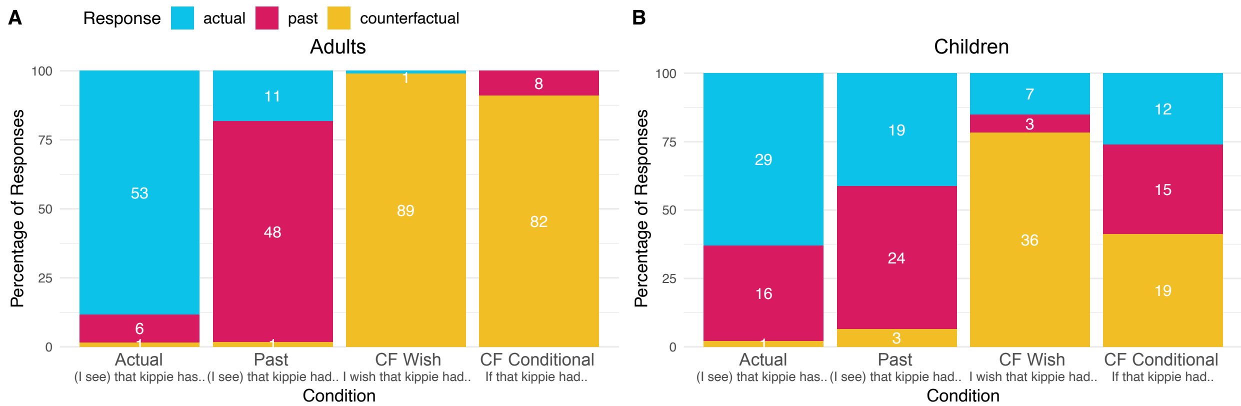

Figure 4.4. Count and percentage (y-axis) of responses (actual, past or counterfactual referent) per condition (ACTUAL, PAST or COUNTERFACTUAL WISH and COUNTERFACTUAL CONDITIONAL). **A**. Adult data (n=30) on the left. **B**. Child data of 4- and-5-year-olds (n=23) on the right. Adults perform at ceiling for all conditions (chose target referents above 75% of the time) and chose more past referents for COUNTERFACTUAL CONDITIONALS than WISHES. Children performed better on COUNTERFACTUAL wishes than CONDITIONALS, sometimes selecting the actual referent, and other times the past referent. Since children chose both past and present referents for ACTUAL and PAST utterances, their reasoning behind selecting a past/actual referent is not clear.

4.2. Children

Children's overall accuracy was much lower than that of adults, see overview in Table 4.2. We conducted a similar analysis as we did for the adult data but adjusted the generalized linear mixed-effects model to include age in months as a predictor. The model output of the maximal GLMM is displayed in Appendix S4.2. Similar to the adult data, we found that construction type ($\chi^2(1)$ = 12.96, p < .001) was a significant predictor, but main condition ($\chi^2(1)$ = 1.89, p = .388) and their interaction were not ($\chi^2(2)$ = 3.80, p = .149). Children's age was also not a significant predictor of performance ($\chi^2(1)$ = 0.06, p = .808). A post-hoc pairwise comparison of the models estimated marginal means (see Table 4.2 column 5) revealed that as for adults there was no significant difference between embedded and main construction types in the ACTUAL (OR 0.62, 95% CI 0.16-2.40, p=0.49) and PAST conditions (OR 0.53, 95% CI 0.15-1.91, p=0.33). For the

COUNTERFACTUAL condition, we found that children made significantly more errors on counterfactual conditionals than on wishes (OR 0.15, 95% CI 0.05-.041, p<.001). These results are displayed in Figure 4.5A.

When we look at the error patterns of individual participants on the counterfactual trials (Figure 4.5B), we see that children fall into several approximately equal-sized groups. Adult-like children, perform well on both COUNTERFACTUAL CONDITIONALS and COUNTERFACTUAL WISH trials (top right). Non-adult-like children fail on both types of counterfactuals (left bottom). A third group of children performs well on COUNTERFACTUAL WISHES but fails on the COUNTERFACTUAL CONDITIONALS (top left), and a fourth group of children made one mistake on the counterfactual items (top middle), this one mistake was always on COUNTERFACTUAL CONDITIONAL items. Remarkably, we never observed a child who passed COUNTERFACTUAL CONDITIONAL trials but made mistakes on COUNTERFACTUAL WISHES (middle and bottom right). When we inspect what type of errors children made on COUNTERFACTUAL items (Figure 4.4B), we see that errors are roughly split between choosing the past or actual referent instead of the counterfactual one. However, similar to but much more pronounced than in the adult data, we observe that children also select the past referent for ACTUAL utterances (35%), and the actual referent for PAST utterances (40%). We can thus not say with confidence that a past selection on a counterfactual indicates that the child entertained a past-temporal interpretation. The response data from individual children can be found in Appendix S4.3.

Table 4.2. Children's Mean Accuracy Scores and Estimated Marginal Means per Condition

Condition	Mean Accuracy	SD	# Responses	Estimated marg. means	SE	CI
Actual Main	0.74	.449	23	0.78	.241	0.24-0.58
Actual Embedded	0.65	.487	23	0.67	.654	0.65-0.91
Past Main	0.65	.487	23	0.67	.298	0.30-0.74
Past Embedded	0.52	.511	23	0.52	.111	0.44-0.85
Counterfactual Wish	0.78	.417	46	0.81	.111	0.44-0.85
Counterfactual Conditional	0.41	.498	46	0.40	.096	0.54-0.91

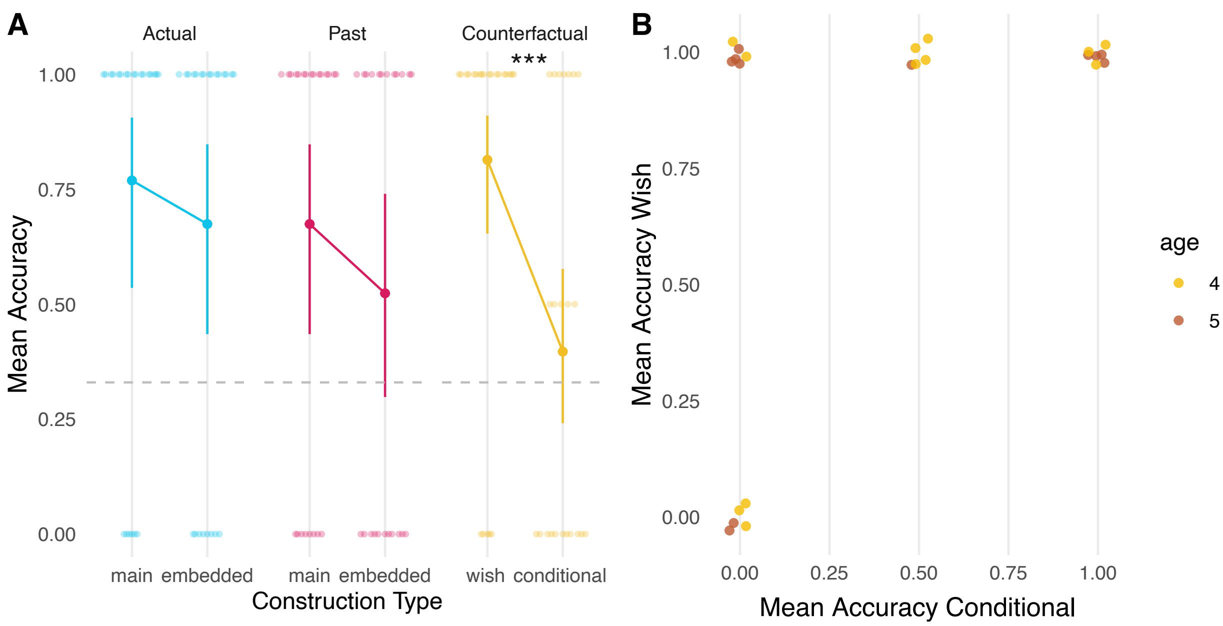

Figure 4.5. **A**. Children's mean accuracy per condition (ACTUAL, PAST and COUNTERFACTUAL) and construction type (MAIN/WISH or EMBEDDED/CONDITIONAL). Raw mean accuracy per participant plotted as scattered dots in the background, estimated marginal means from generalized linear mixed model plotted in foreground, with bars indicating model-based standard errors. Asterisks indicate significance from post-hoc pairwise comparison. Dotted line indicates chance-level. **B.** Mean accuracy of counterfactual wishes (y-axis) plotted against the mean accuracy of counterfactual conditionals (x-axis) for each participant (scattered dots). Yellow dots mark 4-year-old participants, brown dots 5-year-old participants. Scattered dots are slightly jittered for the sake of legibility.

We can gain more insight into children's reasoning by considering the prompted explanations they gave for their decisions. Children that correctly chose counterfactual referents on COUNTERFACTUAL trials, often commented on the fact that that kippie has a different milkshake from the others (7,8), pointing out the unspoken assertion that counterfactual utterances evoke.

(7). Audio: If that kippie had a banana milkshake, he would give me a banana coin.
<child selects counterfactual referent (kippie without a banana milkshake) >
Experimenter: Why that kippie?

z006 (5;4): because it's chocolate and it's different from the others

z018 (4;4): because it has a strawberry, not- um, a banana

(8). Audio: I wish that kippie had a banana milkshake,
so he would give me a banana coin.
<child selects counterfactual referent (kippie without a banana milkshake) >
Experimenter: Why that kippie?

z028 (4;4): um, because, um, it doesn't have a banana coin

z023 (4;5): 'cause it doesn't have a banana milkshake

z026 (4;4): because he said 'I wish that one had a banana'

Not all children understood the counter-to-fact meaning expressed through the COUNTERFACTUAL utterances though. Some children picked a past referent on COUNTERFACTUAL trials, and provided argumentation for this that without a doubt suggests a past interpretation of the utterance (9):

(9). Audio: If that kippie had a banana milkshake, he would give me a banana coin.
<child incorrectly selects past referent (kippie who had a banana milkshake) >
Experimenter: Why that kippie?

z009 (5;7): the one that finished that one is the milkshake man is talking about

z025 (5;1): because- because they said 'that one HAD', he already drank, that's how I knew, that's a little bit easy for me

Other children seemed to reduce the game to matching utterances to the actual referent, the kippie holding the mentioned milkshake (10):

> (10). Audio: If that kippie had a banana milkshake, he would give me a banana coin.
> <child incorrectly selects actual referent (kippie holding a banana milkshake) >
> Experimenter: Why that kippie?
>
> z032 (4;11): Uhh, because it has a banana
>
> z026 (4;4): Because that one has-had a-, uhm, the other one had its cup empty

Surprisingly, children's performance on ACTUAL and PAST utterances was somewhat inconsistent and non-adult like. More so than adults, children chose the actual referent on PAST trials, and the past referent on ACTUAL trials. Some of their responses also suggested that for them, the two could be used interchangeably (11,12):

> (11). Audio: That kippie has a banana milkshake, so he will give me a banana coin.
> <child incorrectly selects past referent (kippie who had a banana milkshake) >
> Experimenter: Why that kippie?
>
> z009 (5;7): 'cause one of the ones that are bananas you can pick and that's a banana one
>
> (12). Audio: That kippie had a strawberry milkshake,
> so he will give me a strawberry coin.
> <child incorrectly selects actual referent (kippie holding a strawberry milkshake) >
> Experimenter: Why that kippie?
>
> z028 (4;4): um, two were strawberry so I just touched them one each.

In some cases, it might be the case that children misheard the utterances (as the only difference between the ACTUAL and PAST condition lies in the morpheme *has* or *had*). However, many children were able to recall the utterances correctly or remarked on the use of past or present tense (13), even when picking the incorrect character, suggesting that mishearing is at least not the sole underlying reason for the non-adult like responses (14). It could be that these children focused on the second part of the utterances, matching the utterance to the right coin.

(13). Audio: That kippie had a strawberry milkshake,
so he will give me a strawberry coin.
<child selects past referent (kippie who had a strawberry milkshake) >
Experimenter: Why that kippie?

z007 (6;1): because the milkshake man said HAD and this one is finished

(14). Audio: I see that kippie had a banana milkshake, so he will give me a banana coin.
<child incorrectly selects actual referent (kippie who has a banana milkshake) >
Experimenter: Why that kippie?

z001 (5;8): because that one, it was having a strawberry milkshake

All in all, from this qualitative discussion it becomes clear that at least some children interpreted the past referent in counterfactual utterances as real. However, it is not clear that all past responses to COUNTERFACTUAL utterances can be interpreted as the child necessarily understanding the utterance as having a past temporal meaning. Some children chose the past referent for present actual utterances, showing that there is some ambiguity to children's use of this response.

5. DISCUSSION

In this paper we addressed two questions about children's comprehension of counterfactual utterances: 1) Do children understand counterfactual wishes before they understand counterfactual conditionals? And, 2) Do children ever interpret the "fake" past tense morpheme in counterfactual constructions as a regular past tense? To answer these questions, we designed a novel referent selection task that could detect a regular past tense interpretation for counterfactual utterances, and contrasted counterfactual wishes and conditionals. The data from 23 4-and 5-year-olds and 30 adults provides strong evidence that there is facilitated processing for counterfactual wishes over conditionals, not just for children, but also for some adults. Similarly, we have compelling evidence that misinterpretation of the counterfactual's "fake" past contributes to participant's difficulty with counterfactual constructions. We discuss these findings in more detail below.

5.1. Facilitated Processing for Counterfactual Wishes over Counterfactual Conditionals

Children performed significantly better on counterfactual wish trials than on counterfactual conditional trials, displaying adult-like reasoning in most of the cases. Notably, there was not a single individual child that showed the opposite pattern (performing well on counterfactual conditional trials while making errors on wishes). We observed better accuracy for counterfactual wishes over counterfactual conditionals for 12 children (a little over half the sample), and even found a similar pattern for 4 adults (13 percent of the sample). What accounts for this facilitated processing of wishes? In the introduction, we laid out the hypothesis that counterfactual wishes might be easier for children to understand (compared to counterfactual conditionals) because of the wish's dedicated counterfactual marking. This dedicated marking (the fact that whenever *wish* selects for a full proposition, this proposition is counterfactual) could facilitate both form-to-meaning mapping during language acquisition and online processing during comprehension.

During language acquisition, children have to learn to map linguistic meaning to input forms. Since counterfactual *wish*-constructions are dedicated in their mapping to counterfactual meaning, they could be easier to learn than counterfactual conditionals, as the conditional construction is used in multiple different ways. If this is the case, we might expect to find that children acquire counterfactual wishes before their conditional counterparts. While it has been found that children usually start to produce wishes before counterfactual conditionals in spontaneous production corpora (Tulling & Cournane, 2019), production data does not provide a conclusive answer as to whether children also comprehend counterfactual wishes before conditionals. In our current study, 6 children passed counterfactual wish trials but failed on counterfactual conditionals, while the opposite pattern (children failing on wishes while they passed counterfactual conditionals) was never observed. This suggests that these children mastered the *wish*-construction before they acquired the meaning of counterfactual conditionals, possibly because they were detected earlier in their input. One adult participant also did not interpret the counterfactuality of the conditional (providing past responses instead), perhaps suggesting that even some adults have difficulty interpreting the "fake" past as counterfactual.

However, alternatively, participants performed better on wish trials because the dedicated wish structure facilitated online processing in comparison with the counterfactual conditional. In theory, the string "If that kippie had a banana milkshake" can be both the beginning of a counterfactual conditional (like it always was in our experiment) or a past conditional. It is only later in the sentence, when you encounter the word would (in "he *would* give me a banana coin") that you can disambiguate between these readings. Possibly, the difficulty children (and some adults) have with the counterfactual conditional finds it roots in online processing. If you first entertain a non-counterfactual conditional reading of the sentence, you will have to revise your initial interpretation once you parse the word *would*. In other words, perhaps the present counterfactual conditional behaves as a garden-path sentence. It is a known finding that children have more trouble than adults revising their initial parse of an utterance (Huang & Hollister, 2019; Trueswell et al., 1999; Weighall, 2008). So perhaps children (and some adults) struggled getting out of the garden path induced by the multifaceted conditional construction's antecedent.

Supporting this possibility are the facts that: (a) some children made errors on only half of the counterfactual conditional items, (b) we observed the same pattern for adults, and (c) children's performance on counterfactual conditionals did not improve with age.

The fact that age was not a predictive factor of children's performance on our task was surprising considering that both 4- and 5-year-olds scored high in comprehending past counterfactual conditionals in a similar referent-selection task (Rouvoli et al., 2019), and children generally start producing counterfactual conditionals around age 3.5-4 (Kuczaj & Daly, 1979; Reilly, 1982; Tulling & Cournane, 2019). Perhaps the processing cost we observed for present counterfactual conditionals is greater than for past counterfactual conditionals, as the usage of the past perfect "If he had eaten a watermelon..." signals counterfactuality in a more transparent way (due to both the past morpheme and the past participle form of the verb). However, we intentionally decided to test children on present counterfactual utterances so that we could isolate the role of the "fake" past tense. Additionally, corpus research shows that present counterfactual constructions are much more common in natural spoken language than past counterfactual constructions (Crutchley, 2004, 2013) and acquired later than other hypothetical conditional constructions (Bowerman, 1986; Kuczaj & Daly, 1979; Reilly, 1982), so it would be surprising if present counterfactual conditionals are harder to understand than their past counterpart. Another way in which our study differed from that of Rouvoli et al. (2019) is in the fact that we tested 6 different types of utterances, while children in the Rouvoli experiment were repeatedly exposed to counterfactual conditionals. So perhaps varying construction types throughout the experiment increased processing costs and required higher levels of attention.

5.2. Past Temporal-Orientation of the "Fake" Past Contributes to Misinterpreting Counterfactuals

The second question addressed by our experiment pertains to the interpretation of the counterfactual's "fake" past tense. In present counterfactuals (such as "If that kippie **had** a banana coin, he would give me a banana milkshake") the past tense morpheme (in bold) indicates present counterfactuality rather than past temporal orientation. We hypothesized that children might not realize this initially and go through a phase where they interpret the "fake" past as real. A real past interpretation of the "fake" past could contribute to previously observed actuality biases in comprehending counterfactual utterances (Riggs et al., 1998; Rouvoli et al., 2019). The results of our study show that such a past temporal orientation is present in at least some children and perhaps even in some adults.

From the qualitative discussion of children's reasoning as to why they selected the past tense referent on counterfactual trials, we learn that some children (like z009 and z025) definitely entertained a past temporal orientation for the utterance. While other children were not as vocal about this, we did observe a substantial amount of past responses on counterfactual trials, suggesting that this misinterpretation of the "fake" past is a more widespread phenomenon. However, as we indicated in the result section, some caution is required in making this generalization. In our experiment, children also selected the past referent after actual utterances about the present. This prevents us from being able to draw any strong conclusions about children's past referent selections. The fact that we observed many actual responses after past and counterfactual utterances as well adds to this uncertainty. However, intriguing and unexpected support for the hypothesis that children misinterpreted the counterfactual's "fake" past comes from our adult data. Although most adults displayed the expected adult-like behavior (selecting a counterfactual referent after hearing a counterfactual utterance), a small but significant proportion of adults displayed the behavior we predicted for children. As discussed above, adults made a significant amount of errors on counterfactual conditional trials, and whenever they made such an error they always selected the past referent instead of the counterfactual one. This suggests that even adults may sometimes fail to detect the "fakeness" of the counterfactual past morpheme. Again, a question that arises is whether this is due to immature knowledge about the counterfactual construction (not knowing that the "fake" past tense does not refer to a past temporal orientation in counterfactual utterances), or whether this behavior can be attributed to misguided online processing. The latter is more likely for adults.

5.3. Open Questions and Future Directions

The garden path that might arise from the string "If that kippie had a banana milkshake" only exists if for the participant the reading of the morpheme had is ambiguous between past (contained in a past hypothetical conditional) and "fake" past (contained in a present counterfactual). A misinterpretation of the "fake" past tense is thus a prerequisite for the garden-path explanation. When participants give non-target responses to our counterfactual items, our task does not allow us to differentiate between two possible causes: 1) participants do not know whether the counterfactuals' past is "fake"; or 2) participants know the past is "fake" but (sometimes) misparse the utterance during online processing. Moreover, both could be simultaneously true at the population level. Future research could disentangle these possibilities by investigating the online processing of present counterfactual conditionals, e.g., recording eye movements with an eye tracker. If present counterfactual conditionals (temporarily) induce garden paths during sentence comprehension, we would expect participants to look at the past referent before looking at the counterfactual referent.

Another open question pertains to the acquisition order of counterfactual wishes and conditionals. Both the 4- and 5-year-olds in our experiment performed very well on counterfactual wish trials, and from their provided rationale it is clear that they understand the counter-to-fact meaning these utterances express. As we laid out at the beginning of this paper, we hypothesize that the linguistically less complex wish construction is acquired before the counterfactual conditional. While we did find facilitated processing for wishes compared to conditionals in our present experiment, our data cannot definitively answer the question as to whether the *wish*-construction is acquired before the conditional. Future research should investigate whether 3-year-olds also understand counterfactual wishes, and perhaps even test younger children, given the fact that 2-year-olds start using *wish*-constructions in their spontaneous production (Tulling & Cournane, 2019). In a pilot study for the experiment discussed in this paper, we tested some 3-year-olds, but quickly noticed the current task was too complex for them. In the future, we should thus aim to test young children on the least complex counterfactual construction with a simple task, to really pinpoint when they command the cognitive prerequisites of counterfactual reasoning.

6. CONCLUSION

All in all, the results of our study support the hypothesis that counterfactual wishes are easier to process than their conditional counterparts, in a task where we keep all other cognitive demands (sentence length, causal relations and task) identical. We also found some evidence suggesting that children sometimes misinterpret the counterfactual's "fake" past as real. These findings illustrate that linguistic complexity plays a role in children's understanding of counterfactual constructions. Tasks investigating children's counterfactual reasoning abilities that only use counterfactual conditionals (Beck et al., 2009; Beck, Riggs, et al., 2011; Byrne, 2007a; Guajardo et al., 2009; Nyhout & Ganea, 2019; Robinson & Beck, 2000; Rouvoli et al., 2019), thus might underestimate children's underlying potential by confounding cognitive complexity with linguistic complexity.

CHAPTER 5: GENERAL DISCUSSION

In this dissertation, I discussed the properties, neural bases and first language acquisition of language that displaces from the here-and-now. Specifically, I focused on two types of displacement that are thought to be unique to human language: 1) experienced displacement, that is, our ability to shift our perspective from the current reality to a representation of the described situation; and 2) modal displacement, our ability to communicate possibilities compatible or incompatible with the actual situation. Open possibilities can be communicated using modal expressions (such as *maybe* or *must*), while counterfactual expressions (such as "If I were…" or "I wish I was") are used to describe a situation that is contra to the state of the actual world. Taking an interdisciplinary approach, I first reviewed literature from linguistics, developmental psychology and cognitive neuroscience on experienced and modal displacement (Chapter 1). Then I formulated broad questions about the neural bases and development of modal displacement. In the first part of this dissertation, I reported two experiments addressing the question of what neural mechanisms underlie the processing of factual and modal utterances using the neuroimaging technique magnetoencephalography (MEG) (Chapter 2). In the second part, I focused on children's acquisition of counterfactual utterances. In Chapter 3, I conducted a corpus study on children's and adult's spontaneous counterfactual productions to investigate the influence of linguistic complexity on the acquisition of counterfactuality. In Chapter 4, I conducted a behavioral study with adults and children aged 4 and 5 to investigate the influence of linguistic complexity on the comprehension of counterfactual utterances. Below, I will first synthesize the insights gained from each chapter and discuss what they tell us about the language ability of displacement. Then, I discuss the questions that remain open and new questions that arise for future investigation.

1. SYNTHESIS OF MAIN FINDINGS

1.1. Model of the Language Capacity of Displacement

As we learned from the literature review in Chapter 1, we have a fairly good understanding of how our minds represent experienced displacement. Whenever we use language to talk about things beyond our immediate environment, we represent the described situation in a discourse representation (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). These situation models represent the properties of the described situation in relation to the here-and-now point of the described actuality, and we mentally shift ourselves into the perspective of that here-and-now point. Children as young as 3-years-old are able to do this (Fecica & O'Neill, 2010; Rall & Harris, 2000; Ziegler & Acquah, 2013), and perhaps even children younger than that (Ganea et al., 2007). Postulating a situation model is thought to play a crucial role in children's ability to imagine situations and engage in pretend play (Harris, 2001), and for adults to immerse in stories (Glenberg et al., 1987). The construction and maintenance of this discourse representation is supported by the default mode network in the brain (Ferstl et al., 2008; Morales et al., 2022; Yeshurun et al., 2021), which is also involved in semantic processing, imagination, daydreaming, and mental time travel (remembering the past and thinking about the future) (Binder et al., 2009; Buckner & Carroll, 2007; Hassabis & Maguire, 2007; Spreng et al., 2009). As it appears, all these functions underly the same mechanism of mentally generating and maintaining a complex scene, whether it is motivated by intrinsic thinking or processing external input. I therefore believe that different types of displacement (spatial, temporal and experienced) rely on exactly this principle of projecting yourself into a different here-and-now point (see Figure 5.1). A common mechanism for these abilities is supported by the fact that 3-year-olds also already have the emerging ability to think and talk about the past (Kuczaj, 1977; Quon & Atance, 2010) and future (Atance & Meltzoff, 2005; Atance & O'Neill, 2005).

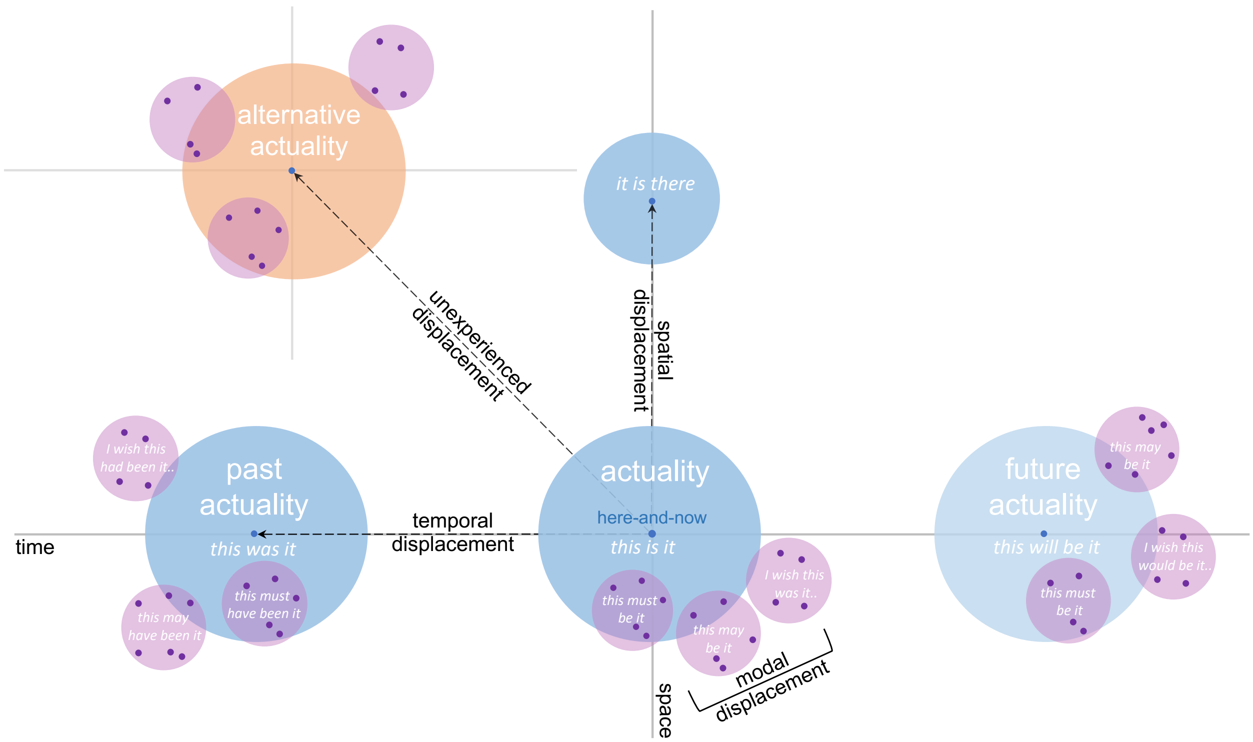

Figure 5.1. Different ways of displacing from the here-and-now of an actuality. The here and now point (blue dot) is depicted in the center, along an x-axis indicating chronological time and the y-axis indicating space. Spatial displacement takes place when moving the here-and-now into a different location, temporal displacement takes place when moving into the here-and-now into a different point in time (e.g., into the past or future). The future situation is constructed based on recombining current facts and past experiences. Actual worlds are indicated with blue circles, the future projection of the actual world is light blue circle. Unexperienced displacement is achieved by shifting into the here-and-now point of an alternative projected actuality (orange circle), that projects its own time and space dimensions. Modal displacement (purple circles) is achieved by postulating possibilities (purple dots) that are possibly (indicated by modal *may*) or necessarily (indicated with modal *must*) compatible with the actual world under consideration, or by postulating possibilities that are not compatible (indicated with counterfactual *wish*).

Modal displacement, on the other hand, is much less understood. Modal displacement differs from other types of displacement in the fact that it does not involve shifting into an alternative here-and-now perspective. Instead, it involves the postulation of possibilities whose compatibility and likelihood are evaluated from the here-and-now point of the actual world under discussion. Note, that I say ‘actual world under discussion’ and not ‘reality’, because we can engage in modal displacement from any actuality (real or imagined). This is displayed in Figure 5.1 by the postulation of a set of possible worlds (purple) that are possibly (some possible worlds are compatible with the actual world, e.g., “this may be it”), necessarily (all possible worlds are compatible with the actual world, e.g., “this must be it”) or not compatible with different actualities (none of the possible worlds are compatible with the actual world, e.g., “I wish this was it”). This idea of postulating multiple possible worlds is borrowed from theories about the underlying semantics of modal and counterfactual utterances (e.g., Kratzer, 1981, 2012; Lewis, 1973; Portner, 2009), but used for visualization purposes rather than making a statement about the underlying mental representation. As we discussed before, there are different ideas about the mental representation of non-actual information, some of which involve representing a multitude of alternatives (Phillips et al., 2019; Phillips & Knobe, 2018), while others involve the representation of as little alternatives as possible (Byrne & Johnson-Laird, 2009; Johnson-Laird, 1994; Johnson-Laird & Ragni, 2019).

In Chapter 2, I performed an experiment to investigate the neural bases of processing modal and factual utterances. Here, my initial expectations presumed that modal utterances would involve the postulation of multiple possibilities, and that this should result in distinctive neural activity, as supporting multiple possibilities would require engaging in modal displacement and holding more than one representation in mind. The idea that postulating multiple possibilities would require more cognitive effort, is a common assumption entertained in psychology and child development research (e.g., Johnson-Laird & Byrne, 2002; Johnson-Laird & Ragni, 2019). However, as discussed thoroughly in the discussion of Chapter 2, we did not observe any increased neural activity for modal utterances (containing *may, might* or *must*) over factual utterances (containing *do*). Instead, we observed the opposite effect and linked this to discourse updating, which was supported by a second experiment where we could cancel the discourse updating effect when we embedded the factual/modal contrast into an uncertain sentential context. The implications of these results are discussed in the following section, where I will lay out the different ways one can think about the representation of possibilities considering the current evidence obtained from the experiments conducted in this dissertation and reviewing prior work.

1.2. The Representation of Possibilities

The experiments described in Chapter 2 observed increased brain activity for factual utterances over modal utterances in brain regions that are part of the default mode network (the right temporoparietal junction, and the ventromedial prefrontal cortex). This activity seems to be reflective of discourse updating, suggesting that factual but not modal meaning updates a situation model. The results are compatible with prior work advocating for a separation between actual and non-actual language representations, arguing that non-actual information does not get incorporated into the discourse representation (Claus, 2008; de Vega et al., 2007; Dwivedi, 1996; Urrutia, de Vega, et al., 2012). What does this mean for the representation of non-actual possibilities?

As discussed in Chapter 1, there are different ways one can think about the representation of possibilities, and all of them involve some type of separation between factual, modal and counterfactual information. In particular, I discussed 3 accounts that have been put forward about the representation of non-actual information: 1) the approach that simultaneously represents a multitude of alternatives (Phillips et al., 2019; Phillips & Knobe, 2018); 2) Mental Model Theory (Byrne & Johnson-Laird, 2009; Johnson-Laird, 1994; Johnson-Laird & Ragni, 2019); and 3) the experiential approach (de Vega & Urrutia, 2011; Kaup et al., 2007). The different models proposed by these accounts are depicted in Figure 5.2. According to Phillips et al. (2019), modal cognition first involves partitioning a relevant possibility space, e.g., based on knowledge or circumstances. A next step involves considering specific possibilities within this domain (the more time you have, the more possibilities you consider), and a last step ordering these possibilities based on their probability and value (e.g., for some deontics their morality) (5.2A). Counterfactual reasoning undergoes the same partitioning and ordering process, except that the actual situation is excluded from consideration at the first step of partitioning. In Mental Model Theory (Johnson-Laird & Ragni, 2019), possibilities are represented in small finite mental models. Modal possibilities (e.g., "It might be a hamster") are represented as the proposed content (e.g., "It is a hamster") with a mental footnote (…) that this model left out some possibilities (i.e., the negative "It is not a hamster") and a symbolic marker indicating the utterance's modal status. Counterfactuals are represented in two mental models, representing both the factual and counterfactual meaning separately (5.2B). Last, the experiential approach has proposed an auxiliary representation system for non-actual meaning. This auxiliary system was proposed to account for the non-actual representation of negation, where a negated utterance like "there is no bird in the sky" involves the first step of representing the non-actual counterpart "there is a bird in the sky" (Kaup et al., 2007), Such an approach has also been applied to account for the counterfactual's dual meaning (de Vega & Urrutia, 2011). The conflict between the actual situation represented in the situation model, and the non-actual situation in the auxiliary model, is thought to give rise to the concept of negation or counter-to-fact meaning. (5.3C). It is not clear how modal meaning would be represented in the experiential approach. In contrast to negation and counterfactuality, modal utterances indicate the status of the actual world is open. So, representing the non-actual modal meaning in the auxiliary system should not lead to conflict with the actual world. There are 3 ways in which I can envision representing modal meaning (open possibility) in a non-actual auxiliary system (5.3D). 1) the experiential simulation of the non-actual modal meaning is less vivid or fuzzier than that of negation or counterfactuality; 2) within the auxiliary system you can represent multiple simulations in parallel; or 3) within the auxiliary system you can simulate multiple possibilities in a cyclic fashion. Such cycling between possibilities has been suggested based on works on decision making in rat brains (Kay et al., 2020).

Each theory has its own upsides and downsides. While the approach of Phillips et al. (2019) matches well with how we generally think about possibilities in semantics, representing a multitude of possibilities is thought to be too taxing or slow on the computation system (Johnson-Laird & Ragni, 2019). Mental Model Theory avoids this problem by using small finite mental models but relies heavily on abstract ideas like ‘footnotes’ and symbolic operators marking discourse type (e.g., factual, modal or counterfactual). It is not clear how a system that relies on symbolic operators to distinguish different types of information can be implemented during online discourse processing (do different types of information ever get integrated, and if yes, how), or what the neural mechanisms supporting such a system are. The experiential view is easier to incorporate with how we think our brains represent discourse situations, as the non-actual auxiliary system is directly tied to a situation model. However, as the experiential approach wants to avoid the use of symbolic operators, it is not clear how modal information is represented (5.3D). Perhaps, a hybrid system that combines situation model + auxiliary system with symbolic operators will be the best way forward. Future work should try to distinguish between these different possibilities.

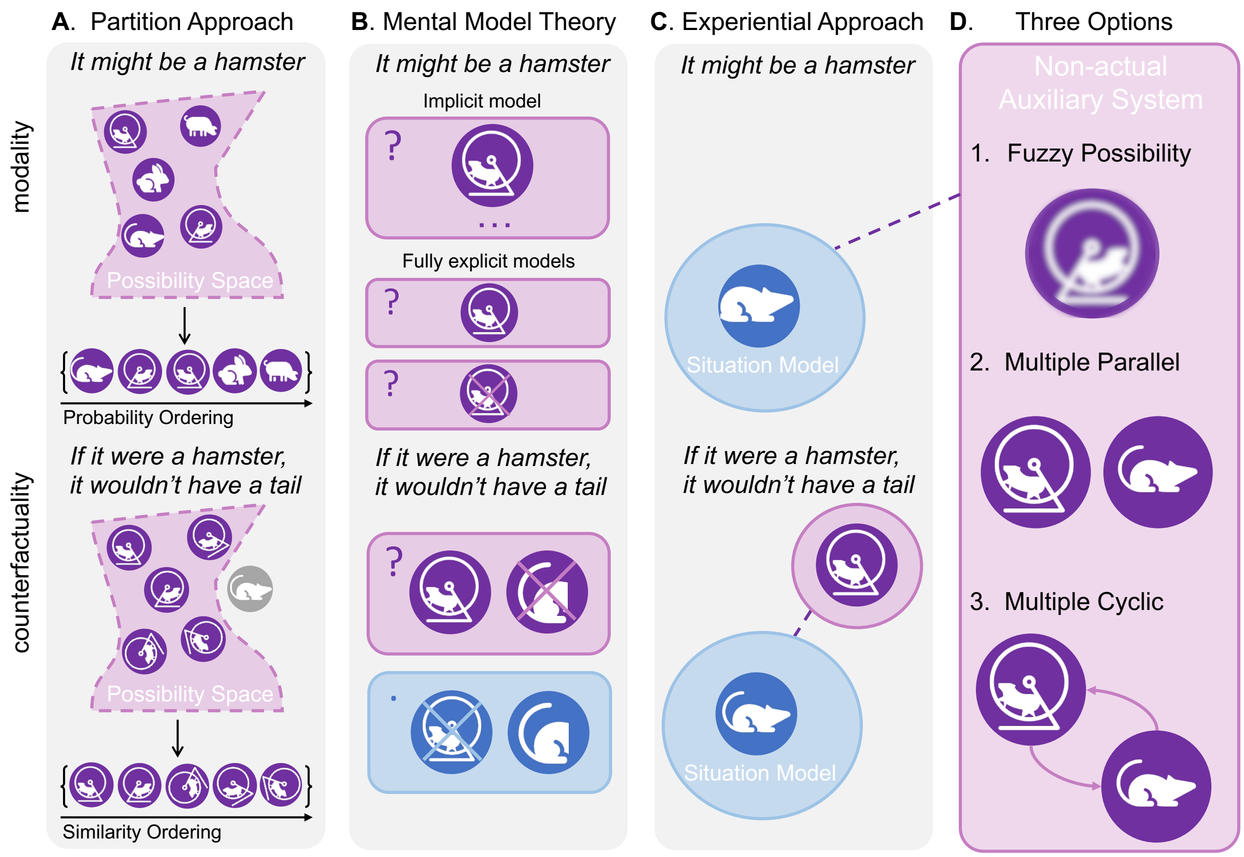

Figure 5.2. Overview of different theories on the representation of possibilities. **A**. Partition Approach: the simultaneous representation of a multitude of alternatives. Top displays the postulation of modal possibilities, bottom represents counterfactual possibilities (leaving the actual world out of the initial possibility space). **B**. Mental Model Theory. Top = mental model of modal possibility, ellipsis indicates a 'mental footnote' about the model being reduced. Question mark is symbolic marker of modality. Bottom = mental models of dual counterfactual meaning. Dot is symbolic marker of factuality. **C**. Experiential simulation view with a situation model keeping track of actual information, and an auxiliary system representing non-actual information. For Modality it is unclear what the non-actual representation looks like. **D**. Proposed possible representations for modal possibilities in an experiential simulation approach: 1) a fuzzy representation, 2) representing multiple parallel possibilities in the auxiliary system, and 3) representing multiple possibilities in (cyclic) sequence.

1.3. Development of Modal Displacement

If we are on the right track that experienced displacement and modal displacement rely on different computational operations (Figure 5.1), then it is possible that these abilities develop at different points throughout development. As discussed above, experienced displacement seems to be in place around age 3 and might even start to develop earlier. Currently, the leading assumption is that cognitive abilities supporting modal displacement likely develop at a later point, starting around age 4 (Leahy & Carey, 2019). This view is supported by the fact that children up to age 4 are not adult like on language tasks involving the comprehension of modal and counterfactual language (Acredolo & Horobin, 1987; Beck et al., 2006; Guajardo et al., 2009; Nyhout & Ganea, 2019; Ozturk & Papafragou, 2015; Reilly, 1982; Rouvoli et al., 2019). However, in this dissertation I emphasized the fact that cognitive complexity is often confounded with linguistic complexity. This is especially true for counterfactual language, as in English (and many other languages) its morphological tense (the "fake" past) does not match the utterance's temporal orientation.

In Chapter 3 and 4, I considered this form-to-meaning mapping challenge for the acquisition of counterfactual constructions. Specifically, I compared the acquisition of counterfactual conditionals with that of counterfactual wishes, as they differ in linguistic complexity. Counterfactual conditional constructions are linguistically complex in multiple ways: they involve a complex causal dependency relationship, have complex form-to-meaning mapping, and compete with other (present and past) conditionals while being less frequent in the child's input. Due to their dedicated counterfactual meaning, *wish*-constructions lack much of this complexity: they do not require causal dependency and only embed propositions that are counterfactual, therefore providing a salient cue for the presence of "fake" past tense. With a large-scale corpus study on children's transcribed speech, I showed that children start spontaneously producing counterfactual utterances in appropriate contexts well before age 4 (Chapter 3). Children's early counterfactual utterances contained a significant number of present-for-past substitutions, suggesting that initially the do not realize the "fake" past expresses counterfactual meaning. Children generally started producing wishes around age 2 or 3, and counterfactual conditionals around age 3 or 4, suggesting that linguistically less-complex structures are acquired before more complex ones. This corroborates findings by Cournane (2021) who similarly found children started producing dedicated epistemic modal adverbs that are linguistically less complex before epistemic uses of complex polysemous modal auxiliaries. Like the counterfactuals, children started producing modal utterances around age 2 or 3, before the age children have currently been found to postulate possibilities (Leahy & Carey, 2019). However, children's usage data alone cannot confirm whether they also understand what they are saying or represent that language in an adult-like way.

In Chapter 4, I tested 4- and 5-year-olds comprehension of counterfactual wishes and conditionals. I found that children performed better on wishes, and some evidence suggesting that the comprehension of counterfactual utterances is further complicated by the presence of a "fake" past marking, which sometimes gets interpreted as being real. Notably, I found no improvement with age, and many of the 4-year-olds had already mastered the counterfactual *wish*-construction, which suggests younger children might also have this construction already. While we tried to test 3-year-olds in a pilot version of this study, children this age were not able to do the task I designed.

Currently, I am working on a study intended for 2- and 3-year-olds that probes children's implicit understanding of counterfactual wishes without imposing any excessive task demands.

Together, these findings suggest that modal displacement might be available earlier than researchers currently assume, and emphasize the necessity of untangling the effects of linguistic, cognitive and task complexity in research on first language development. However, in order to clearly answer the question of when modal displacement develops, we have to have a better idea of what this ability entails and formulate explicit models about the representation of possibility.

2. FUTURE DIRECTIONS

In this dissertation, I posed two broad questions about the neural and developmental bases of modal displacement, that are far from being fully answered. In the previous section, I laid out the developments I have made in approaching these questions, and proposed models for how language displacement and the representation of possibility might work. To distinguish between different theories about the representation of possibility, and in order to increase our understanding about our ability of modal displacement there are several areas that should be explored in future research.

2.1. Discourse updating and knowledge updating

We have discussed how our brains are remarkably equipped to imagine any type of situation, shifting into the perspective of the described actuality whether the situation is real or imagined. We represent the here-and-now of a fictional (or believed to be untrue) situation while keeping this information separate from our general knowledge. However, using the phrase 'keeping it separate' grossly simplifies what is actually happening. The true power of representing situations that we have not experienced ourselves lies in our ability to learn from them. We gain a lot of knowledge from second-hand information, suggesting that at some point the information that we represent in a situation model must get integrated into our own beliefs. How does this work? And when do we accept information? A recurrent finding in first language development is that young children are ‘naïve skeptics’, and are more likely to claim that that real-life events and characters are fantasy rather than the reverse (Woolley & Cox, 2007; Woolley & Ghossainy, 2013), suggesting they are initially cautious to integrate discourse information into their beliefs. It would be interesting to see how this ability develops in relation to the ability to represent possibilities. In this dissertation I reported neural correlates of ‘discourse updating’ for factual, but not for modal information. How does this translate to ‘knowledge updating’? Do we ever incorporate non-factual statements into our set of beliefs? And when do children start to do so? Are they over-accepting or possibly skeptic in this aspect as well?

Relatedly, there are questions to be asked about how knowledge affects the way we integrate information into our discourse models. What happens if you get a factual utterance from a source that is known to be unreliable? Will you still update your representation of a situation? If our brains resist updating under uncertainty in general (as for language indicating mere possibility) we would expect that discourse updating would not take place if you know the source to be unreliable. However, it could also be that the updating effect evoked by information presented as fact is so automatic, that it does not matter what the source is. Answers to these questions could inform how we can communicate information in the most effective and responsible way.

2.2. Storing non-actual information

As discussed, theories about the representation of possibilities assume that non-actual meaning is somehow separated from factual meaning, either by representing it in a separate model, or by marking it with some type of symbolic operator of uncertainty or incompatibility. In the experiential approach, the representation of the non-actual meaning of counterfactuals and negation is thought to be temporary, until the conflict between actual and non-actual information gets resolved, approximated to last between 750-1500 ms (de Vega et al., 2007; de Vega & Urrutia, 2012; Kaup et al., 2007). For non-actual modal meaning however, it is an open question as to how long this information is retained. Imagine you are watching a murder mystery and try to solve the murder before the detective does. You could encounter a sentence such as “Maybe Marie came back to finish the job afterwards” and this information could be highly relevant to solve the case. It could also be one of the many distractors to keep the audience guessing what really happened until the detective cleverly puts together all the pieces in the end. Imagine we have our situation model tracking the here-and-now of the story and represented “Maybe Marie…” in the auxiliary system. What happens next? Do we hold on to this information? If yes, for how long – until it gets resolved? What if it is a loose end and it never gets resolved? And what happens if before the possibility gets resolved, another possibility comes up (e.g., “Or maybe the butler was the late night visitor”). Can the auxiliary system hold on to two unrelated possibilities at once, or do we overwrite the other? Does it matter whether the possibilities are compatible or incompatible? There are endless questions to ask.

On a related note, there are questions to be asked about long-term memory and the recollection of different types of information. In particular, it would be valuable to do recall tasks comparing the long-term accessibility of non-actual information and factual information. Are facts easier to recall than possibilities? And are open possibilities more accessible than disconfirmed possibilities? A recent paper on this subject suggests that memory differences across different types of information might vary across categories, i.e., people and locations were remembered similarly across past, future and counterfactual conditions, but the objects and times for future and counterfactual situations were harder to remember than for the past (De Brigard et al., 2020).

2.3. Cross-linguistic Development of Non-Actual Language

Last, cross-linguistic research on the development of non-actual language can provide valuable insights into the relevance of the form-to-meaning mapping in the first language acquisition of modal and counterfactual meaning. The studies in this paper were conducted testing children that had English as their native language. For that reason, we discussed properties of counterfactual constructions that were specific to the English language, i.e., the dedicated meaning of the *wish*-construction. There are many other languages where the language constructions used to express counterfactual desires is not dedicated, e.g., in Dutch or Greek. Similarly, there are languages that, unlike English, have dedicated counterfactual conditional constructions, such as Russian or Mandarin Chinese. If the form-to-mapping challenge is an important factor determining the onset of different language constructions during first language acquisition, then we would expect dedicated counterfactual (or modal) expressions to appear before polysemous or non-transparent ones regardless of the structural category. If this is the case, dedicated expressions should be used whenever possible to probe children's modal or counterfactual comprehension to get a more accurate idea about their underlying cognitive abilities.

CONCLUSION

In this dissertation I discussed the language property of displacement, the ability to process language outside the here-and-now. I have covered both displacing oneself from the immediate environment by shifting into an alternative here-and-now perspective, and modal displacement, which is achieved by using language expressing possibilities compatible (modal) or incompatible (counterfactual) with the current actuality. I have conducted different experiments investigating the neural bases underlying modal displacement using magnetoencephalography, and conducted corpus and behavioral work to gain insight into the development of this ability. Through these studies I uncovered the neural responses involved in updating discourse representations with factual (but not uncertain) information, and learnt that linguistic complexity seems to influence the onset of different counterfactual constructions in first language acquisition. I have discussed these results in relation to theories about the representation of situations and possibilities and formulated different directions for future investigation. Perhaps, I have raised more questions in this dissertation than I have provided answers for. However, I hope this dissertation will inspire future exploration of one of the most fascinating language abilities.

APPENDICES

<figures on next two pages>

Figure S2.1. Details on controlled between-stimuli variation Experiment 1. The target sentences were identical in structure, e.g. “But the king says that their squires may too” but varied in controlled manner in five ways: **A**: Overview of the variation in count of used connectives (*and*, *but* and *so*) across modal bases. **B**: Variation of nouns (main subject) across modal base conditions in average length (in letters), average lexical frequency, average log lexical frequency, number of syllables and number of morphemes. **C**: Variation of the determiners used to refer to the embedded subject: the, a long distance pronoun (LD) referring to a referent in the prior context sentence or a short distance pronoun (SD) referring to a referent in the target sentence. **D**: Variation of the elided VP across modal base conditions in average length (in words and letters), percentage of verb phrases that included verbs indicating a state (in contrast to an event), percentage of verbs taking two arguments (transitive) versus verbs that take one argument (intransitive), average syntactic node count (how many phrase nodes are present counting phrases containing a noun (NP), verb (VP), adjective (AP), preposition (PP) and infinitive (IP)) and average syntactic complexity (maximum amount of nodes opened at the same time), e.g. *to see dusty books at the library* includes 5 syntactic phrases [IP to [VP see [AP dusty [NP books]]]] [PP at the library] and has at most 4 nodes open at the same time. **E1**: List of different embedding verbs used with count of usage across modal bases **E2**: Variation of embedding verbs used across modal base conditions in average length (in letters), average lexical frequency, average log lexical frequency, number of syllables and number of morphemes.

Figure S2.2. Time course estimated brain activity [dSPM] of reliable detected clusters from ROI analysis Experiment 2, displayed separately for the data collected in NY and the data collected in AD. Both the lrACC and rvmPFC show an interaction between sentence type (factual, conditional and presupposed) and verb (*do*, *may* or *might*) with increased activation for *do* > *may/might* when embedded in factual sentences, and decreased activation for *do* < *may/might* in presupposed sentences. The effect in the l**r**ACC was most prominent in the NY data while the effect in the rvMPFC was more prominent in the AD data. Boundaries of the analysis window (150-400 ms) are indicated by dashed lines, identified clusters are displayed in grey. Boxplots display estimated brain activity within the time window of the identified temporal clusters, black dots indicate mean activity. Regions of interest are outlined on brain and shaded when containing identified clusters. Cluster effects are not significant after correction comparison across multiple regions of interest.

	A		B	E		C		D	
e.g.	CON	the	NOUN.SG	VERB1	that	DET	NOUN.PL | **TARGET**	<ELIDED VP>	too
	"But	the	king	says	that	their	squires | **may**	<sit around the round table>	too."

A. Variation of CONNECTIVES across bases

Base	And	But	So	Total
factual	30	28	22	80
rules	30	28	22	80
knowledge	30	28	22	80
Total	90	84	66	240

C. Variation of DETERMINERS across bases

Flavor	LD[i]	SD[ii]	the	Total
factual	22	22	36	80
rules	20	24	36	80
knowledge	24	20	36	80
Total	66	66	108	240

[i]LD indicates long distance pronouns, pronouns that refer back to a referent in the context sentence, e.g. 'Apparently knights overhear a lot of secrets in the castle. But the servant concludes that **their** squires must too'.

[ii]SD indicates a short distance pronoun, pronouns that refer back to a referent in the target sentence, e.g. 'The bride's father is known to give boring speeches at receptions. But the bride hints that **her** brothers may too'.

B. Variation of NOUNS across modal base conditions

Values	Rules	Knowledge
Average Length (letters)	6,8	7,6
Average Freq_HAL	83,9	87,20
Average Log_Freq_HAL	8984,9	8585,2
Average #Syll	2,25	2,5
Average #Morph	1,6	1,75

D. Variation of ELIDED VP across modal base conditions (Av. = average)

Base	Av. Ellipsis Length (words)	Av. Ellipsis Length (letters)	Stative Verbs[i]	Transitive Verbs[ii]	Av. Node Count[iii]	Av. Node Complexity[iv]
rules	4,975	28,15	8%	88%	2,7	2,2
knowledge	4,775	27,9	15%	88%	2,925	2,55

[I] Percentage of stative verbs (e.g. know, accept, live).

[ii] Percentage of transitive verbs (with two arguments), 12 percent are intransitive verbs (taking one argument).

[iii] Estimation of size syntactic structure by counting structural phrase nodes. Counted nodes are: noun phrase (NP), verb phrase (VP), predicate adjective phrase (AP), preposition phrase (PP), infinitival phrase (IP). NPs within APs or PPs were not counted. IPs were only counted in presence of the infinitival marker 'to'. Adverbs, particles and attributive adjectives were ignored.

[iv] Layers of embedded structure (complements). PPs are considered adjuncts not adding to complexity: e.g. [IP [VP [NP]]] = counted as having 3 structural layers, [IP [VP] [PP]] = counted as having 2 structural layers.

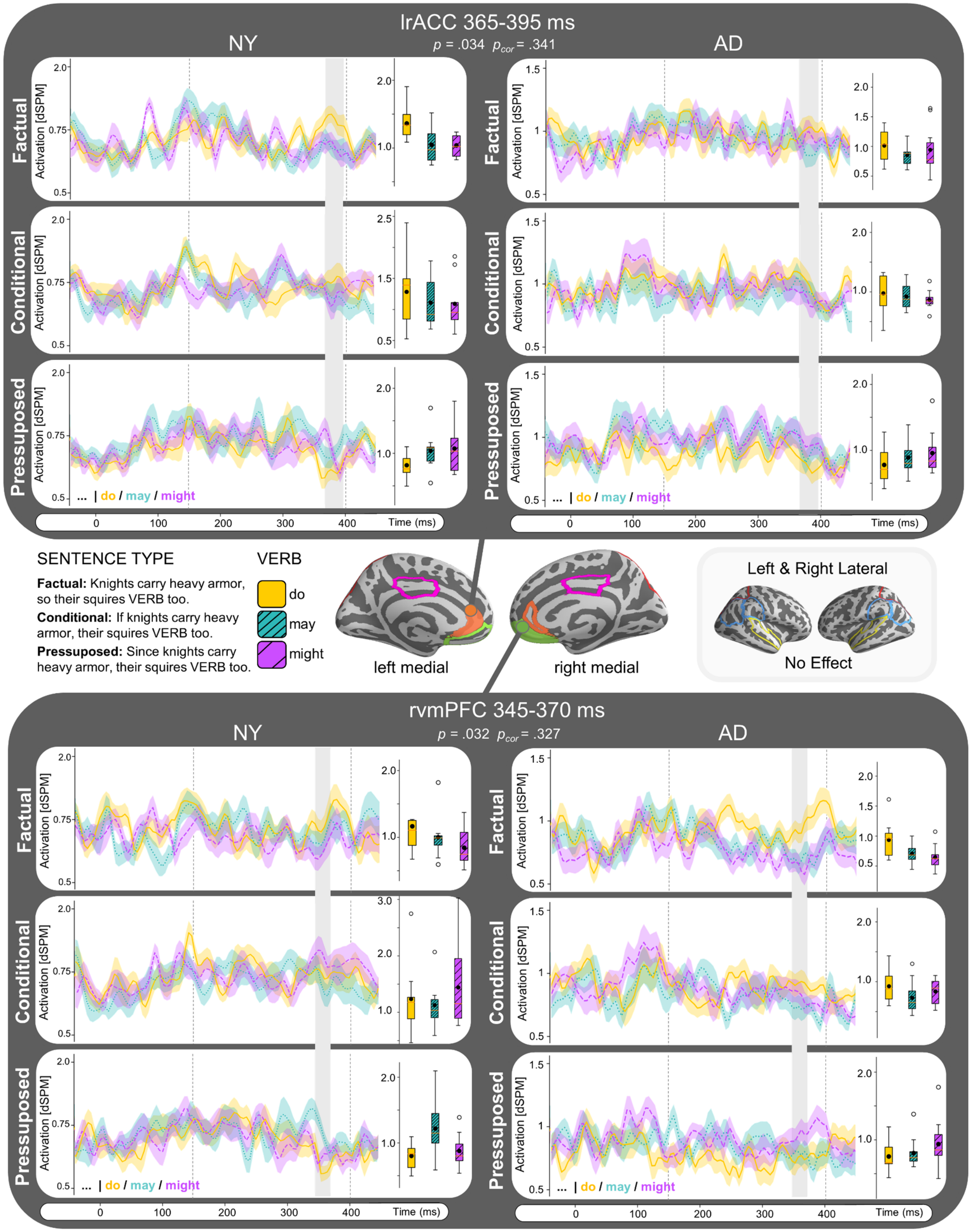
lrACC 365-395 ms
NY
AD
p = .034 p_cor = .341
Factual
Conditional
Pressuposed
Activation [dSPM]
... | do / may / might
Time (ms)
SENTENCE TYPE
Factual: Knights carry heavy armor, so their squires VERB too.
Conditional: If knights carry heavy armor, their squires VERB too.
Pressuposed: Since knights carry heavy armor, their squires VERB too.
VERB
do
may
might
left medial
right medial
Left & Right Lateral
No Effect
rvmPFC 345-370 ms
p = .032 p_cor = .327

Table S3.1. Overview of All Corpora Used: corpus name, collection, children's age range (in months), the number of children documented, the number of utterances and wishes found separated by children and adults, and corpus citation. Shaded rows indicate corpora that did not include any wish-utterances from children.

Corpus	Collection	Min Age	Max Age	N	N Child Utterances (N wishes)	N Adult Utterances (N wishes)	Citation
Belfast	Eng-UK	24.1	54.2	11	25781 (1)	80899 (28)	(Henry, 1995)
Bliss	Eng-NA	40.0	64.0	4	1302 (1)	1011 (0)	(Bliss, 1988)
Bloom	Eng-NA	19.2	37.7	2	31970 (0)	36071 (NA)	(L. Bloom et al., 1974)
Bohannon	Eng-NA	36.0	36.0	3	4057 (0)	6737 (NA)	(Bohannon & Marquis, 1977)
Braunwald	Eng-NA	15.0	84.5	1	53311 (30)	33970 (21)	(Braunwald, 1971)
Brown	Eng-NA	27.1	62.4	2	96747 (32)	86172 (32)	(Brown, 1973)
Clark	Eng-NA	26.5	38.1	1	18185 (2)	24283 (9)	(Clark, 1979)
Compton-Pater	Eng-NA	8.0	38.7	3	25169 (1)	0 (0)	(Pater, 1997)
Cruttenden	Eng-UK	17.6	46.1	2	3061 (0)	0 (NA)	(Cruttenden, 1978)
Davis	Eng-NA	6.4	36.1	6	97128 (3)	0 (0)	(B. L. Davis & MacNeilage, 1995)
Davis-CDI	Eng-NA	8.9	35.7	4	3763 (3)	0 (0)	(Davis et al., 2018)
Demetras1	Eng-NA	24.9	47.9	1	6971 (1)	8293 (0)	(Demetras, 1989)
Demetras2	Eng-NA	26.5	33.8	1	9411 (3)	11119 (5)	(Demetras et al., 1986)
EllisWeismer	Clinical-MOR	30.0	66.0	13	71074 (11)	102876 (11)	(Weismer et al., 2013)
ENNI	Clinical-MOR	48.4	119.8	1	29269 (1)	650 (0)	(Schneider et al., 2006)
Evans	Eng-NA	71.3	71.3	1	4787 (0)	10 (NA)	(Evans, 1985)
Fletcher	Eng-UK	36.0	86.4	48	22073 (2)	26251 (0)	(Johnson, 1986)

Forrester	Eng-UK	12.0	60.0	1	7536 (2)	8919 (3)	(Forrester, 2002)
Garvey	Eng-NA	34.0	67.0	62	10338 (26)	9 (0)	(Garvey & Hogan, 1973)
Gathercole	Eng-NA	33.0	78.0	14	6724 (11)	2743 (1)	(Gathercole, 1986)
Gelman	Eng-NA	18.0	84.2	2	52281 (19)	126964 (32)	(Gelman et al., 1998, 2004, 2014; Jipson et al., 2016)
Gleason	Eng-NA	26.5	62.3	22	20247 (3)	38880 (6)	(Bellinger & Gleason, 1982)
Goad	Eng-NA	17.6	42.6	2	8853 (1)	0 (0)	(Parsons, 2006)
Gopnik	Eng-NA	24.0	64.7	1	3754 (1)	6347 (0)	(M. Gopnik, 1989)
Haggerty	Eng-NA	31.6	31.6	1	1739 (0)	0 (NA)	(Haggerty, 1930)
Hall	Eng-NA	54.0	57.0	36	124924 (71)	107305	(Hall & Tirre, 1979)
Hicks	Eng-NA	61.0	95.0	21	8992 (0)	5248 (NA)	(Hicks, 1991)
Higginson	Eng-NA	22.0	35.0	1	5953 (0)	9672 (NA)	(Higginson, 1985)
HSLLD	Eng-NA	42.6	141.9	11	130124 (25)	172908 (75)	(Dickinson & Tabors, 2001)
Inkelas	Eng-NA	6.3	45.9	1	1873 (0)	0 (NA)	(Inkelas & Rose, 2007)
Kuczaj	Eng-NA	28.8	60.4	1	32172 (25)	25622 (14)	(Kuczaj, 1977)
Lara	Eng-UK	21.4	40.0	1	57639 (4)	99728 (14)	(Rowland & Fletcher, 2006)
Mac-Whinney	Eng-NA	1.0	92.1	3	57675 (69)	63605 (17)	(MacWhinney, 1991)
Manchester	Eng-UK	20.7	36.3	13	249504 (5)	374198 (39)	(Theakston et al., 2001)
Morisset	Eng-NA	30.0	39.0	100	12964 (1)	19341 (0)	(Morisset et al., 1990)
MPI-EVA-Manchester	Eng-UK	24.0	37.1	2	253910 (14)	320710 (83)	(Lieven et al., 2009)

Nelson	Eng-NA	19.6	32.8	1	4552 (4)	1624 (1)	(Nelson, 1989)
New-England	Eng-NA	13.5	33.0	24	12041 (0)	43667 (NA)	(Ninio et al., 1994)
Newman Ratner	Eng-NA	11.0	288.0	1	23268 (0)	164190 (NA)	(Newman et al., 2016)
Paido-English	Eng-NA	27.0	69.0	1	10169 (0)	0 (NA)	(Edwards & Beckman, 2008)
Penney	Eng-NA	59.9	72.1	21	1491 (0)	944 (NA)	(Judd, 2018)
Peterson-McCabe	Eng-NA	48.0	113.0	1	10361 (1)	7216 (0)	(Peterson & McCabe, 1983)
Post	Eng-NA	22.7	32.2	1	16893 (0)	18755 (NA)	(Demetras et al., 1986)
Providence	Eng-NA	11.1	48.1	6	176132 (16)	283927 (109)	(Demuth et al., 2006)
Sachs	Eng-NA	15.0	57.1	1	17236 (0)	12222 (NA)	(Sachs & Nelson, 1983)
Smith	Eng-UK	26.1	45.4	1	5308 (0)	0 (NA)	(Smith, 1973)
Snow	Eng-NA	29.6	45.1	1	13520 (2)	21033 (16)	(MacWhinney & Snow, 1990)
Sprott	Eng-NA	33.0	61.0	27	4718 (2)	1606 (0)	(Sprott, 1992)
Suppes	Eng-NA	23.5	39.7	1	33950 (1)	35172 (4)	(Suppes, 1974)
Thomas	Eng-UK	24.4	59.7	2	218984 (58)	372363 (153)	(Lieven et al., 2009)
Tommerdahl	Eng-UK	29.0	45.0	1	12027 (2)	13879 (2)	(Tommerdahl & Kilpatrick, 2014)
Valian	Eng-NA	21.7	32.8	1	15945 (1)	27715 (2)	(Valian, 1991)
VanHouten	Eng-NA	28.0	43.4	26	4455 (1)	8736 (0)	(Van Houten, 1986)
VanKleeck	Eng-NA	37.0	48.0	20	6677 (0)	8756 (NA)	(van Kleeck et al., 1985)
Warren	Eng-NA	30.0	70.0	11	3563 (0)	5847 (NA)	(Warren-Leubecker, 1982)

Weist	Eng-NA	25.0	60.2	7	47577 (8)	65165 (12)	(Weist & Zevenbergen, 2008)
Wells	Eng-UK	17.7	60.8	31	57537 (14)	40756 (11)	(Wells, 1981)
Total	NA	1.0	288.0	585	2247665 (478)	2934114 (771)	NA

Table S3.2 Intercoder Reliability Values. Results from calculating overall accuracy (%), Gwet's AC1 coefficient and Conger's kappa statistic for each coded variable.

Variable	Test	Value	CI (95%)
Main Subject	Percent Agreement	0.94	(0.893,0.987)
	AC1	0.94	(0.884,0.987)
	Conger's Kappa	0.80	(0.64,0.951)
Embedded Subject	Percent Agreement	0.96	(0.921,0.999)
	AC1	0.96	(0.913,0.999)
	Conger's Kappa	0.94	(0.881,0.998)
Subjunctivity	Percent Agreement	0.96	(0.921,0.999)
	AC1	0.95	(0.903,0.999)
	Conger's Kappa	0.89	(0.784,0.997)
Temporal Orientation	Percent Agreement	0.88	(0.815,0.945)
	AC1	0.87	(0.792,0.941)
	Conger's Kappa	0.60	(0.403,0.797)
Bare Error	Percent Agreement	0.93	(0.879,0.981)
	AC1	0.92	(0.871,0.982)
	Conger's Kappa	0.28	(-0.034,0.6)
Tense Error	Percent Agreement	0.89	(0.828,0.952)
	AC1	0.88	(0.807,0.95)
	Conger's Kappa	0.61	(0.439,0.79)
Evidence Counterfactuality	Percent Agreement	0.61	(0.513,0.707)
(before discussion)	AC1	0.52	(0.401,0.64)
	Conger's Kappa	0.49	(0.358,0.612)

Evidence Counterfactuality	Percent Agreement	0.87	(0.803,0.937)
(after discussion)	AC1	0.84	(0.757,0.922)
	Conger's Kappa	0.83	(0.743,0.918)

Table S3.3. Overview of Children's Past Tense Productivity. For each child we recorded their age range (in months), total amount of utterances, total amount of produced present-for-past errors, age range while making errors, the proportion of correct past tense marking in the context of the temporal adverb yesterday (YD), total amount of past tense overregularization (OR) and age range of during which overregularized.

Child	Abe	Adam	Laura	Mark	Ross	Thomas
Corpus	Kuczaj	Brown	Braunwald	Mac-Whinney	Mac-Whinney	Thomas
Min Age	28.8	27.1	15.0	5.5	16.4	24.4
Max Age	60.4	62.4	84.5	69.3	92.1	59.7
N Utterances	31958	46651	39750	20754	36379	218439
N Errors	4.0	4.0	5.0	NA	NA	5.0
Error Min Age	34.7	41.5	25.9	NA	NA	35.7
Error Max Age	51.4	62.4	31.4	NA	NA	42.1
N Past with YD	13/14	3/3	2/6	NA	NA	NA
YD Min Age	34.7	55.0	28.0	NA	NA	NA
N OR	218.0	22.0	8.0	NA	NA	22.0
OR Min Age	34.7	42.3	26.2	NA	NA	35.9
OR Max Age	51.2	62.4	31.0	NA	NA	42.1

Table S4.1. Model output of generalized linear mixed effect model on adult data.
Accuracy ~ Construction_Type * Main_Condition + (1 | id)

Effect	estimate (odd ratio)	std. error	z.value	CI 5%	CI 95%
Intercept	16.90	8.79	5.43	6.09	46.90
Construction Type					
wish (compared to conditional)	9.76	10.60	2.09	1.15	82.70
Main Condition					
past (compared to counterfactual)	0.21	0.13	-2.50	0.06	0.72
actual (compared to counterfactual)	0.60	0.42	-0.74	0.15	2.34
Construction Type * Main Condition					
wish * actual	0.16	0.21	-1.42	0.01	2.00
wish * past	0.15	0.21	-1.37	0.01	2.26

Nakagawa $R^2 = 0.487$

Table S4.2. Model output of generalized linear mixed effect model on child data.
Accuracy ~ Construction_Type * Main_Condition + Age_Months + (1 | id)

Effect	estimate (odd ratio)	std. error	z.value	CI 5%	CI 95%
Intercept	0.66	0.25	-1.13	0.32	1.36
Construction Type					
wish (compared to conditional)	6.65	3.42	3.68	2.43	18.2
Main Condition					
past (compared to counterfactual)	1.67	0.93	0.92	0.56	4.97
actual (compared to counterfactual)	3.15	1.81	1.99	1.02	9.73
Age in Months	1.06	0.26	0.24	0.65	1.73
Construction Type * Main Condition					
wish * actual	0.28	0.23	-1.53	0.06	1.42
wish * past	0.24	0.21	-1.66	0.05	1.30

Nakagawa $R^2 = 0.280$

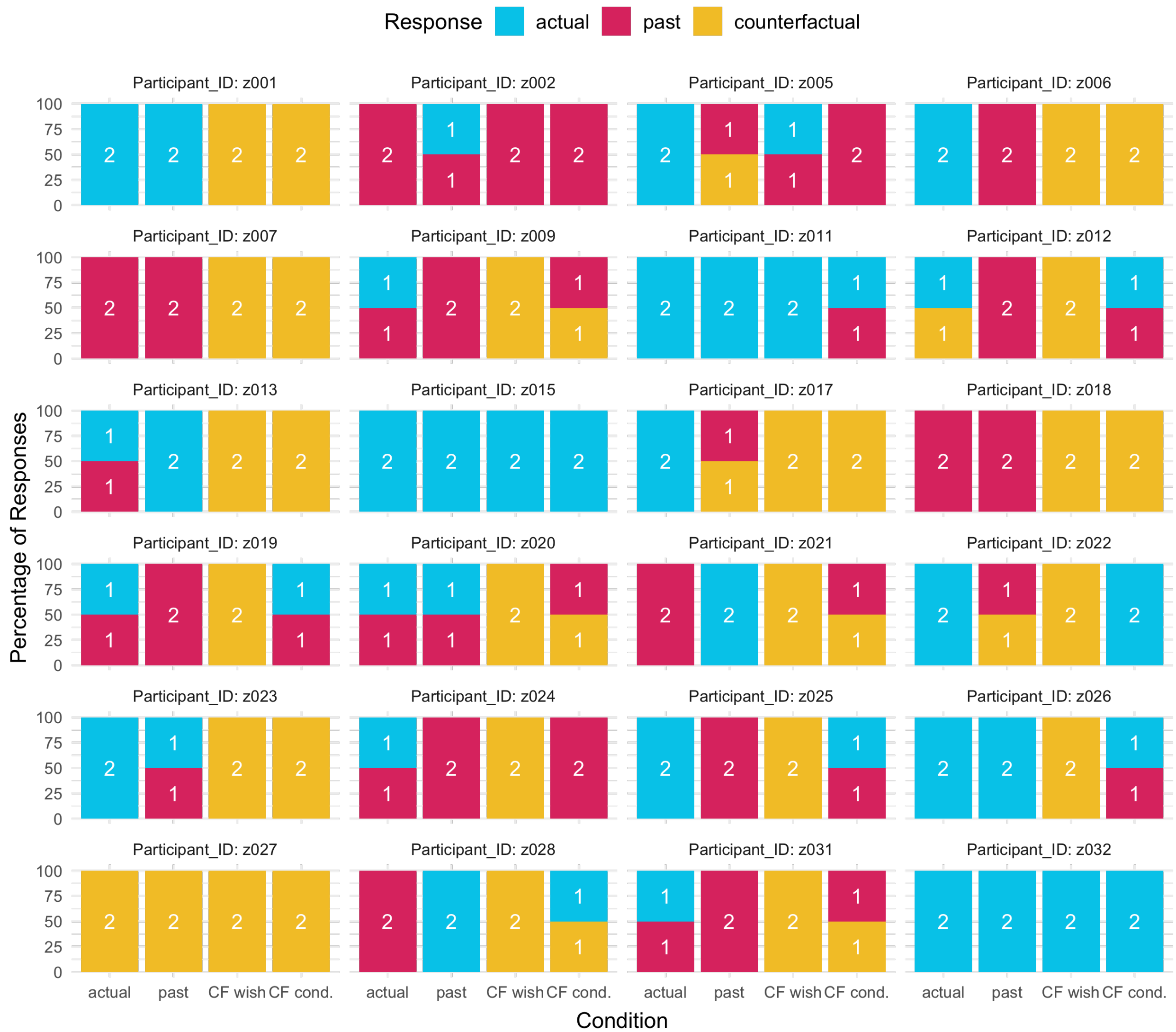

Figure S4.3. Count and Percentage (y-axis) of responses (actual, past or counterfactual referent) per condition (ACTUAL, PAST or COUNTERFACTUAL WISH and COUNTERFACTUAL CONDITIONAL) split by participant (n=24). CF = Counterfactual.

REFERENCES

Acredolo, C., & Horobin, K. (1987). Development of relational reasoning and avoidance of premature closure. *Developmental Psychology*, *23*(1), 13. https://doi.org/10.1037/0012-1649.23.1.13

Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., & Ochiai, M. (2001). Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Transactions on Applied Superconductivity*, *11*(1), 669–672. https://doi.org/10.1109/77.919433

Arregui, A. (2005). *On the accessibility of possible worlds: The role of tense and aspect*. BookSurge Publishing.

Arregui, A. (2009). On similarity in counterfactuals. *Linguistics and Philosophy*, *32*(3), 245–278. https://doi.org/10.1007/s10988-009-9060-7

Arregui, A. (2020). Counterfactuals. In *The Wiley Blackwell Companion to Semantics* (pp. 1–33). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118788516.sem007

Atance, C. M., & Meltzoff, A. N. (2005). My future self: Young children's ability to anticipate and explain future states. *Cognitive Development*, *20*, 341–361.

Atance, C. M., & O'Neill, D. K. (2005). Preschoolers' talk about future situations. *First Language*, *25*(1), 5–18. https://doi.org/10.1177/0142723705045678

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology*, *50*(8), 2061.

Aygen, G. (2004). *T-to-C and overt marking of Counterfactuals: Syntactic and Semantic Implications*. Harvard Working Papers in Linguistics 10: 1–18.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bailey, H. R., & Zacks, J. M. (2015). Situation model updating in young and older adults: Global versus incremental mechanisms. *Psychology and Aging*, *30*(2), 232–244. https://doi.org/10.1037/a0039081

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Barnes, M. A., Raghubar, K. P., Faulkner, H., & Denton, C. A. (2014). The Construction of Visual-spatial Situation Models in Children's Reading and Their Relation to Reading Comprehension. *Journal of Experimental Child Psychology*, *119*, 101–111. https://doi.org/10.1016/j.jecp.2013.10.011

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beck, S. R. (2016). Why What Is Counterfactual Really Matters: A Response to Weisberg and Gopnik (). *Cognitive Science*, *40*(1), 253–256. https://doi.org/10.1111/cogs.12235

Beck, S. R., McColgan, K. L. T., Robinson, E. J., & Rowley, M. G. (2011). Imagining what might be: Why children underestimate uncertainty. *Journal of Experimental Child Psychology*, *110*(4), 603–610. https://doi.org/10.1016/j.jecp.2011.06.010

Beck, S. R., Riggs, K. J., & Burns, P. (2011). Multiple Developments in Counterfactual Thinking. In *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology.* Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199590698.001.0001/acprof-9780199590698-chapter-6

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning*, *15*(4), 337–354. https://doi.org/10.1080/13546780903135904

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's Thinking about Counterfactuals and Future Hypotheticals as Possibilities. *Child Development*, *77*(2), 413–426. JSTOR.

Bellinger, D. C., & Gleason, J. B. (1982). Sex differences in parental directives to young children. *Sex Roles*, *8*(11), 1123–1139.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bhatt, R. (1999). *Covert modality in non-finite contexts* [Ph.D., University of Pennsylvania]. http://www.proquest.com/docview/304517430/abstract/3A58CF56F54A4731PQ/1

Bickerton, D. (2008). But how did protolanguage actually start? *Interaction Studies*, *9*(1), 169–176. https://doi.org/10.1075/is.9.1.12bic

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual Processing during the Conscious Resting State: A Functional MRI Study. *Journal of Cognitive Neuroscience*, *11*(1), 80–93. https://doi.org/10.1162/089892999563265

Bjorkman, B. M. (2015). *Only some "fake" pasts are real: Contrasting counterfactuals and sequence of tense* [Manuscript].

Bjorkman, B. M., & Halpert, C. (2017). In an imperfect world: Deriving the typology of counterfactual marking. In *Modality Across Syntactic Categories.* (Vol. 1). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198718208.003.0009

Black, J., Williams, D. M., & Ferguson, H. J. (2018). Imagining counterfactual worlds in autism spectrum disorder. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1444–1463.

Bliss, L. S. (1988). Modal usage by preschool children. *Journal of Applied Developmental Psychology*, *9*(3), 253–261.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when, and why. *Cognitive Psychology*, *6*(3), 380–420.

Bloom, P. (2000). *Learning, development, and conceptual change. How children learn the meanings of words. Cambridge, MA, US.*

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (Version 6.2.09) [Computer software]. http://www.praat.org/

Bohannon III, J. N., & Marquis, A. L. (1977). Children's control of adult speech. *Child Development*, 1002–1008.

Bowerman, M. (1986). First steps in acquiring conditionals. In E. C. Traugott, A. G. Meulen, J. S. Reilly, & C. A. Ferguson (Eds.), *On conditionals* (pp. 285–308). Cambridge University Press.

Braunwald, S. R. (1971). Mother-child communication: The function of maternal-language input. *Word*, *27*(1–3), 28–50.

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin et Memoires de La Societe Anatomique de Paris*, *6*, 330–357.

Brodbeck, C. (2017). *Eelbrain*. Zenodo. https://doi.org/10.5281/zenodo.438193

Brown, R. (1973). *A first language: The early stages.* Harvard University Press.

Brownell, H. H., Potter, H. H., Bihrle, A. M., & Gardner, H. (1986). Inference deficits in right brain-damaged patients. *Brain and Language*, *27*(2), 310–321. https://doi.org/10.1016/0093-934X(86)90022-2

Buchsbaum, D., Bridgers, S., Weisberg Skolnick, D., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2202–2212. https://doi.org/10.1098/rstb.2012.0122

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport*, *18*(17), 1851. https://doi.org/10.1097/WNR.0b013e3282f1a999

Burmester, J., Spalek, K., & Wartenburger, I. (2014). Context updating during sentence comprehension: The effect of aboutness topic. *Brain and Language*, *137*, 62–76. https://doi.org/10.1016/j.bandl.2014.08.001

Byrne, R. M. J. (2007a). *The rational imagination: How people create alternatives to reality*. MIT press.

Byrne, R. M. J. (2007b). Précis of *The Rational Imagination: How People Create Alternatives to Reality*. *Behavioral and Brain Sciences*, *30*(5–6), 439–453. https://doi.org/10.1017/S0140525X07002579

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.

Byrne, R. M. J., & Johnson-Laird, P. N. (2009). 'If' and the problems of conditional reasoning. *Trends in Cognitive Sciences*, *13*(7), 282–287. https://doi.org/10.1016/j.tics.2009.04.003

Caplan, R., & Dapretto, M. (2001). Making sense during conversation: An fMRI study. *Neuroreport*, *12*(16), 3625–3632.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science*, *359*(6381), 1263–1266. https://doi.org/10.1126/science.aao3539

Clark, E. V. (1979). Building a vocabulary: Words for objects, actions and relations. *Language Acquisition*, 149–160.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition: The 20th Annual Carnegie Mellon Symposium on Cognition*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/nyulibrary-ebooks/detail.action?docID=1619055

Clark, E. V. (2001). Emergent categories in first language acquisition. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (1st ed., pp. 379–405). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.015

Claus, B. (2008). *Discourse referents for nonspecific entities of described nonfactual situations.* 102–104.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213. https://doi.org/10.1037/h0026256

Collins, P. (2007). Can/could and may/might in British, American and Australian English: A corpus-based account. *World Englishes*, *26*(4), 474–491. https://doi.org/10.1111/j.1467-971X.2007.00523.x

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Cournane, A. (2015). *Revisiting the epistemic gap: Evidence for a grammatical source*. Proceedings of the 39th annual Boston University conference on language development.

Cournane, A. (2021). Revisiting the epistemic gap: It's not the thought that counts. *Language Acquisition*, *28*(3), 215–240. https://doi.org/10.1080/10489223.2020.1860054

Cromer, R. F. (1968). *The development of temporal reference during the acquisition of language* [Unpublished PhD dissertation]. Harvard University.

Crutchley, A. (2004). 'If She Had of Shutted the Cage, the Rabbit Wouldn't Escape': Past Counterfactuals Elicited from 6-to 11-Year-Old Children. *First Language*, *24*(2), 209–240. https://doi.org/10.1177/0142723704044935

Crutchley, A. (2013). Structure of child and adult past counterfactuals, and implications for acquisition of the construction. *Journal of Child Language; Cambridge*, *40*(2), 438–468. http://dx.doi.org/10.1017/S0305000912000049

Cruttenden, A. (1978). Assimilation in child language and elsewhere. *Journal of Child Language*, *5*(2), 373–378.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic Statistical Parametric Mapping: Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity. *Neuron*, *26*(1), 55–67. https://doi.org/10.1016/S0896-6273(00)81138-1

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1199–1211.

Davis, B., Van der Feest, S., & Hoyoung, Y. (2018). Speech sound characteristics of early words: Influence of phonological factors across vocabulary development. *Journal of Child Language*, *45*(3), 673–702.

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*, *51*(12), 2401–2414. https://doi.org/10.1016/j.neuropsychologia.2013.01.015

De Brigard, F., Gessell, B., Yang, B. W., Stewart, G., & Marsh, E. J. (2020). Remembering possible times: Memory for details of past, future, and counterfactual simulations. *Psychology of Consciousness: Theory, Research, and Practice*, *7*(4), 331–339. https://doi.org/10.1037/cns0000220

de Vega, M., & Urrutia, M. (2011). Counterfactual sentences activate embodied meaning: An action–sentence compatibility effect study. *Journal of Cognitive Psychology*, *23*(8), 962–973. https://doi.org/10.1080/20445911.2011.590471

de Vega, M., & Urrutia, M. (2012). Discourse Updating after Reading a Counterfactual Event. *Psicologica: International Journal of Methodology and Experimental Psychology*, *33*(2), 157–173.

de Vega, M., Urrutia, M., & Riffo, B. (2007). Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition*, *35*(6), 1410–1421. https://doi.org/10.3758/BF03193611

de Villiers, J. (2000). Language and theory of mind: What are the developmental relationships? In *Understanding other minds: Perspectives from developmental cognitive neuroscience, 2nd ed* (pp. 83–123). Oxford University Press.

Decety, J., & Lamm, C. (2007). The Role of the Right Temporoparietal Junction in Social Interaction: How Low-Level Computational Processes Contribute to Meta-Cognition. *The Neuroscientist*, *13*(6), 580–593. Gale OneFile: Nursing and Allied Health.

Dehghani, M., Boghrati, R., Man, K., Hoover, J., Gimbel, S. I., Vaswani, A., Zevin, J. D., Immordino-Yang, M. H., Gordon, A. S., Damasio, A., & Kaplan, J. T. (2017). Decoding the neural representation of story meanings across languages. *Human Brain Mapping*, *38*(12), 6096–6106. https://doi.org/10.1002/hbm.23814

Demetras, M. (1989). Changes in parents' conversational responses: A function of grammatical development. *ASHA, St. Louis, MO*.

Demetras, M. J., Post, K. N., & Snow, C. E. (1986). Feedback to first language learners: The role of repetitions and clarification questions. *Journal of Child Language*, *13*(2), 275–292.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, *49*(2), 137–173.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school.* Paul H Brookes Publishing.

Dowty, D. R. (1986). The Effects of Aspectual Class on the Temporal Structure of Discourse: Semantics or Pragmatics? *Linguistics and Philosophy*, *9*(1), 37–61. JSTOR.

Dreher, J.-C., & Grafman, J. (2003). Dissociating the Roles of the Rostral Anterior Cingulate and the Lateral Prefrontal Cortices in Performing Two Tasks Simultaneously or Successively. *Cerebral Cortex*, *13*(4), 329–339. https://doi.org/10.1093/cercor/13.4.329

Drummond, A. (2010). *Ibex-farm*. http://spellout.net/ibexfarm/

Dudman, V. H. (1983). Tense and time in English verb clusters of the primary pattern. *Australian Journal of Linguistics*, *3*(1), 25–44. https://doi.org/10.1080/07268608308599298

Dwivedi, V. D. (1996). Modality and discourse processing. *McGill Working Papers in Linguistics*, *12*, 17–52.

Dwivedi, V. D., Phillips, N. A., Laguë-Beauvais, M., & Baum, S. R. (2006). An electrophysiological study of mood, modal context, and anaphora. *Brain Research*, *1117*(1), 135–153. https://doi.org/10.1016/j.brainres.2006.07.048

Edwards, J., & Beckman, M. E. (2008). Methodological questions in studying consonant acquisition. *Clinical Linguistics & Phonetics*, *22*(12), 937–956.

Erlewine, M. Y., & Kotek, H. (2016). A streamlined approach to online linguistic surveys. *Natural Language & Linguistic Theory*, *34*(2), 481–495. https://doi.org/10.1007/s11049-015-9305-9

Evans, M. A. (1985). Self-initiated speech repairs: A reflection of communicative monitoring in young children. *Developmental Psychology*, *21*(2), 365.

Ezzyat, Y., & Davachi, L. (2011). What Constitutes an Episode in Episodic Memory? *Psychological Science*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Fecica, A. M., & O'Neill, D. K. (2010). A step at a time: Preliterate children's simulation of narrative movement during story comprehension. *Cognition*, *116*(3), 368–381. https://doi.org/10.1016/j.cognition.2010.05.014

Feiman, R., Mody, S., Sanborn, S., & Carey, S. (2017). What Do You Mean, No? Toddlers' Comprehension of Logical "No" and "Not." *Language Learning and Development*, *13*(4), 430–450. https://doi.org/10.1080/15475441.2017.1317253

Ferguson, H. J., Sanford, A. J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. *Brain Research*, *1236*, 113–125. https://doi.org/10.1016/j.brainres.2008.07.099

Ferstl, E. C., Neumann, J., Bogler, C., & Cramon, D. Y. von. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, *29*(5), 581–593. https://doi.org/10.1002/hbm.20422

Ferstl, E. C., Rinck, M., & Cramon, D. Y. von. (2005). Emotional and Temporal Aspects of Situation Model Processing during Text Comprehension: An Event-Related fMRI Study. *Journal of Cognitive Neuroscience*, *17*(5), 724–739. https://doi.org/10.1162/0898929053747658

Fillenbaum, S. (1974). Information amplified: Memory for counterfactual conditionals. *Journal of Experimental Psychology*, *102*(1), 44.

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128. https://doi.org/10.1016/0010-0277(95)00692-R

Forrester, M. A. (2002). Appropriating cultural conceptions of childhood: Participation in conversation. *Childhood*, *9*(3), 255–276.

Francis, G. A., & Gibson, J. L. (2021). *Pretense, Executive Functions, and Counterfactual Reasoning: Evaluating the Case for a 'Unified Theory of Imaginative Processes.'* PsyArXiv. https://doi.org/10.31234/osf.io/skxb8

Friese, U., Rutschmann, R., Raabe, M., & Schmalhofer, F. (2008). Neural Indicators of Inference Processes in Text Comprehension: An Event-related Functional Magnetic Resonance Imaging Study. *Journal of Cognitive Neuroscience*, *20*(11), 2110–2124. https://doi.org/10.1162/jocn.2008.20141

Gajardo-Vidal, A., Lorca-Puls, D. L., Hope, T. M. H., Parker Jones, O., Seghier, M. L., Prejawa, S., Crinion, J. T., Leff, A. P., Green, D. W., & Price, C. J. (2018). How right hemisphere damage after stroke can impair speech comprehension. *Brain*, *141*(12), 3389–3404. https://doi.org/10.1093/brain/awy270

Ganea, P. A. (2005). Contextual Factors Affect Absent Reference Comprehension in 14-Month-Olds. *Child Development*, *76*(5), 989–998. https://doi.org/10.1111/j.1467-8624.2005.00892.x

Ganea, P. A., Shutts, K., Spelke, E. S., & DeLoache, J. S. (2007). Thinking of Things Unseen: Infants' Use of Language to Update Mental Representations. *Psychological Science*, *18*(8), 734–739. https://doi.org/10.1111/j.1467-9280.2007.01968.x

Garvey, C., & Hogan, R. (1973). Social speech and social interaction: Egocentrism revisited. *Child Development*, 562–568.

Gathercole, V. (1986). The acquisition of the present perfect: Explaining differences in the speech of Scottish and American children. *Journal of Child Language*, *13*, 537–560. https://doi.org/10.1017/S0305000900006875

Gathercole, V. C. M. (1980). *BIRDIES LIKE BIRDSEED THE BESTER THAN BUNS: A STUDY OF RELATIONAL COMPARATIVES AND THEIR ACQUISITION.* [Unpublished PhD dissertation]. University of Kansas.

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., Pappas, A., & Keil, F. C. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development*, i–157.

Gelman, S. A., Taylor, M. G., Nguyen, S. P., Leaper, C., & Bigler, R. S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development*, i–142.

Gelman, S. A., Ware, E. A., Kleinberg, F., Manczak, E. M., & Stilwell, S. M. (2014). Individual differences in children's and parents' generic language. *Child Development*, *85*(3), 924–940.

German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, *6*(5), 514–523. https://doi.org/10.1111/1467-7687.00309

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 430. https://doi.org/10.1037/0278-7393.16.3.430

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard Words. *Language Learning and Development*, *1*(1), 23–64. https://doi.org/10.1207/s15473341lld0101_4

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83. https://doi.org/10.1016/0749-596X(87)90063-5

Gopnik, A., & Walker, C. M. (2013). Considering Counterfactuals The Relationship between Causal Learning and Pretend Play. *American Journal of Play*, *6*(1), 15–28.

Gopnik, M. (1989). Reflections on challenges raised and questions asked. In P. R. Zelazo, R. G. Barr, & P. D. Zelazo (Eds.), *Challenges to Developmental Paradigms: Implications for Theory, Assessment and Treatment* (pp. 259–273). Taylor & Francis Group. http://ebookcentral.proquest.com/lib/nyulibrary-ebooks/detail.action?docID=1639232

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027

Grosu, I., & Cournane, A. (forthcoming). The changing facts of children's counterfactual reasoning. *Proceedings of the 46th Annual Boston University Conference on Language Development (BUCLD46)*.

Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology*, *27*(3), 681–702. https://doi.org/10.1348/026151008X357886

Guajardo, N. R., & Turley-Ames, K. J. (2004a). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*(1), 53–80. https://doi.org/10.1016/j.cogdev.2003.09.002

Guajardo, N. R., & Turley-Ames, K. J. (2004b). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*(1), 53–80. https://doi.org/10.1016/j.cogdev.2003.09.002

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, *61*(1), 29–48. https://doi.org/10.1348/000711006X126600

Hacquard, V. (2006). *Aspects of modality*.

Hacquard, V. (2013). *On the grammatical category of modality*. 19–26.

Hacquard, V. (2020). Actuality Entailments. In *The Wiley Blackwell Companion to Semantics* (pp. 1–26). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118788516.sem052

Hacquard, V., & Lidz, J. (2018). Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind & Language*, *34*(1), 73–96. https://doi.org/10.1111/mila.12192

Haggerty, L. C. (1930). What a two-and-one-half-year-old child said in one day. *The Pedagogical Seminary and Journal of Genetic Psychology*, *37*(1), 75–101.

Hall, W. S., & Tirre, W. C. (1979). The communicative environment of young children: Social class, ethnic, and situational differences. *Center for the Study of Reading Technical Report; No. 125*.

Harris, P. L. (2001). The Veridicality Assumption. *Mind & Language*, *16*(3), 247–262. https://doi.org/10.1111/1468-0017.00168

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259. https://doi.org/10.1016/S0010-0277(96)00715-9

Harris, P. L., Kavanaugh, R. D., Wellman, H. M., & Hickling, A. K. (1993). Young Children's Understanding of Pretense. *Monographs of the Society for Research in Child Development*, *58*(1), i–107. JSTOR. https://doi.org/10.2307/1166074

Hartig, F. (2022). *DHARMa: Residual Diagnostics for Hierarchical (Multi-level / Mixed) Regression Models.* (R package version 0.4.5) [Computer software]. http://florianhartig.github.io/DHARMa/

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Hauptman, M., Blank, I., & Fedorenko, E. (2022). *Non-literal language processing is jointly supported by the language and Theory of Mind networks: Evidence from a novel meta-analytic fMRI approach* (p. 2022.03.08.481056). bioRxiv. https://doi.org/10.1101/2022.03.08.481056

Henry, A. (1995). *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford University Press on Demand.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*(4), 131–138. https://doi.org/10.1016/S1364-6613(00)01463-7

Hicks, D. (1991). Kinds of texts: Narrative genre skills among children from two communities. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure*. Hillsdale, N.J. : L. Erlbaum.

Higginson, R. P. (1985). *Fixing: Assimilation in language acquisition*.

Hill, A., Collier-Baker, E., & Suddendorf, T. (2011). Inferential reasoning by exclusion in great apes, lesser apes, and spider monkeys. *Journal of Comparative Psychology*, *125*(1), 91. https://doi.org/10.1037/a0020867

Hockett, C., F. (1959). Animal "Languages" and Human Language. *Human Biology*, *31*(1), 32–39.

Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: Theme organization. *Brain and Language*, *38*(2), 253–277. https://doi.org/10.1016/0093-934X(90)90114-V

Huang, Y. T., & Hollister, E. (2019). Developmental parsing and linguistic knowledge: Reexamining the role of cognitive control in the kindergarten path effect. *Journal of Experimental Child Psychology*, *184*, 210–219. https://doi.org/10.1016/j.jecp.2019.04.005

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, *31*(2), 231–270.

Iatridou, S., & Embick, D. (1994). *Conditional inversion*. *24*, 189–203.

Inkelas, S., & Rose, Y. (2007). Positional neutralization: A case study from child language. *Language*, 707–736.

Ippolito, M. (2003). Presuppositions and Implicatures in Counterfactuals. *Natural Language Semantics*, *11*(2), 145–186. https://doi.org/10.1023/A:1024411924818

Ippolito, M. (2006). Semantic Composition and Presupposition Projection in Subjunctive Conditionals. *Linguistics and Philosophy*, *29*(6), 631–672. https://doi.org/10.1007/s10988-006-9006-2

Ippolito, M., & Keyser, S. J. (2013). *Subjunctive Conditionals: A Linguistic Analysis*. MIT Press. http://ebookcentral.proquest.com/lib/nyulibrary-ebooks/detail.action?docID=3339677

Jacoby, N., & Fedorenko, E. (2020). Discourse-level comprehension engages medial frontal Theory of Mind brain regions even for expository texts. *Language, Cognition and Neuroscience*, *35*(6), 780–796. https://doi.org/10.1080/23273798.2018.1525494

James, D. (1982). Past Tense and the Hypothetical a Cross-Linguistic Study. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language," 6*(3), 375–403.

Jiang, Y. (2019). Ways for expressing counterfactual conditionals in Mandarin Chinese. *Linguistic Vanguard*.

Jipson, J. L., Gülgöz, S., & Gelman, S. A. (2016). Parent–child conversations regarding the ontological status of a robotic dog. *Cognitive Development*, *39*, 21–35.

Johns, C. L., Tooley, K. M., & Traxler, M. J. (2008). Discourse Impairments Following Right Hemisphere Brain Damage: A Critical Review. *Language and Linguistics Compass*, *2*(6), 1038–1062. https://doi.org/10.1111/j.1749-818X.2008.00094.x

Johnson, M. G. (1986). *A computer-based approach to the analysis of child language data*. University of Reading.

Johnson-Laird, P. N. (1994). Mental models and probabilistic thinking. *Cognition*, *50*(1), 189–209. https://doi.org/10.1016/0010-0277(94)90028-0

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*(4), 646–678. https://doi.org/10.1037/0033-295X.109.4.646

Johnson-Laird, P. N., & Ragni, M. (2019). Possibilities as the foundation of reasoning. *Cognition*, *193*, 103950. https://doi.org/10.1016/j.cognition.2019.04.019

Judd, A. (2018). *Exploring relationships between phonological awareness and phonological productive abilities of kindergarten-aged children* [Master Thesis, Memorial University of Newfoundland]. https://research.library.mun.ca/13353/

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, *9*(11), 512–518. https://doi.org/10.1016/j.tics.2005.09.009

Jung-Beeman, M., & Chiarello, C. (1998). Complementary Right- and Left-Hemisphere Language Comprehension. *Current Directions in Psychological Science*, *7*(1), 2–8. https://doi.org/10.1111/1467-8721.ep11521805

Karawani, H. (2014). *The real, the fake, and the fake fake in counterfactual conditionals, crosslinguistically*. LOT.

Karawani, H., & Zeijlstra, H. (2013). The semantic contribution of the past tense morpheme *kaan* in Palestinian counterfactuals. *Journal of Portuguese Linguistics*, *12*(1), 105–119. https://doi.org/10.5334/jpl.79

Katis, D. (1997). The emergence of conditionals in child language: Are they really so late? *Current Issues in Linguistic Theory*, *143*, 355–386.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, *38*(7), 1033–1050. https://doi.org/10.1016/j.pragma.2005.09.012

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *The Quarterly Journal of Experimental Psychology*, *60*(7), 976–990. https://doi.org/10.1080/17470210600823512

Kay, K., Chung, J. E., Sosa, M., Schor, J. S., Karlsson, M. P., Larkin, M. C., Liu, D. F., & Frank, L. M. (2020). Constant Sub-second Cycling between Representations of Possible Futures in the Hippocampus. *Cell*, *180*(3), 552-567.e25. https://doi.org/10.1016/j.cell.2020.01.014

Kazanina, N., Baker, S., & Seddon, H. (2019). *Actuality bias in verb learning: The case of sublexically modal transfer verbs*. https://doi.org/10.17863/CAM.35350

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology*, *37*(2), 216–223. https://doi.org/10.1111/1469-8986.3720216

Koster-Hale, J., Richardson, H., Velez, N., Asaba, M., Young, L., & Saxe, R. (2017). Mentalizing regions represent distributed, continuous, and abstract dimensions of others' beliefs. *NeuroImage*, *161*, 9–18. https://doi.org/10.1016/j.neuroimage.2017.08.026

Kovács, Á. M., Tauzin, T., Téglás, E., Gergely, G., & Csibra, G. (2014). Pointing as Epistemic Request: 12-month-olds Point to Receive New Information. *Infancy*, *19*(6), 543–557. https://doi.org/10.1111/infa.12060

Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, *1*(3), 337–355. https://doi.org/10.1007/BF00353453

Kratzer, A. (1981). The Notional Category of Modality. In *Words, Worlds and Contexts. New approaches in World (Eikmeyer H-J & Rieser H, eds)* (pp. 38–74). Berlin: Mouton de Gruyter.

Kratzer, A. (2012). *Modals and conditionals: New and revised perspectives* (Vol. 36). Oxford University Press.

Kratzer, A. (2013). *Modality for the 21st century*. 181–201.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 589–600. https://doi.org/10.1016/S0022-5371(77)80021-2

Kuczaj, S. A., & Daly, M. J. (1979). The development of hypothetical reference in the speech of young children*. *Journal of Child Language*, *6*(3), 563–579. https://doi.org/10.1017/S0305000900002543

Kulakova, E., Aichhorn, M., Schurz, M., Kronbichler, M., & Perner, J. (2013). Processing counterfactual and hypothetical conditionals: An fMRI investigation. *NeuroImage*, *72*, 265–271. https://doi.org/10.1016/j.neuroimage.2013.01.060

Kulakova, E., & Nieuwland, M. S. (2016a). Pragmatic skills predict online counterfactual comprehension: Evidence from the N400. *Cognitive, Affective & Behavioral Neuroscience*, *16*(5), 814–824. https://doi.org/10.3758/s13415-016-0433-4

Kulakova, E., & Nieuwland, M. S. (2016b). Understanding Counterfactuality: A Review of Experimental Evidence for the Dual Meaning of Counterfactuals. *Language and Linguistics Compass*, *10*(2), 49–65. https://doi.org/10.1111/lnc3.12175

LaCroix, A. N., James, E., & Rogalsky, C. (2021). Neural Resources Supporting Language Production vs. Comprehension in Chronic Post-stroke Aphasia: A Meta-Analysis Using Activation Likelihood Estimates. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/article/10.3389/fnhum.2021.680933

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child* (pp. xi, 250). Harvard University Press.

Landau, B., Gleitman, L. R., & Landau, B. (2009). *Language and experience: Evidence from the blind child* (Vol. 8). Harvard University Press.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. JSTOR. https://doi.org/10.2307/2529310

Leahy, B. P., & Carey, S. E. (2019). The Acquisition of Modal Concepts. *Trends in Cognitive Sciences*, *24*(1), 65–78. https://doi.org/10.1016/j.tics.2019.11.004

Lewis, D. (1973). Counterfactuals and comparative possibility. In *Ifs* (pp. 57–85). Springer.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*(3), 481–507.

Lillard, A. S. (2013). *Fictional Worlds, the Neuroscience of the Imagination, and Childhood Education*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195395761.013.0010

Lillard, A. S. (2017). Why Do the Children (Pretend) Play? *Trends in Cognitive Sciences*, *21*(11), 826–834. https://doi.org/10.1016/j.tics.2017.08.001

Lin, N., Yang, X., Li, J., Wang, S., Hua, H., Ma, Y., & Li, X. (2018). Neural correlates of three cognitive processes involved in theory of mind and discourse comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(2), 273–283. https://doi.org/10.3758/s13415-018-0568-6

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic Infants, but Not Chimpanzees, Communicate About Absent Entities. *Psychological Science (0956-7976)*, *20*(5), 654–660. https://doi.org/10.1111/j.1467-9280.2009.02346.x

MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk. Hillsdale, NJ, US*.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. Transcription format and programs* (Vol. 1). Psychology Press.

MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of Child Language*, *17*(2), 457–472.

Mahy, C. E. V., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience*, *9*, 68–81. https://doi.org/10.1016/j.dcn.2014.01.002

Male, A. G., Smith, C. A., & Gouldthorp, B. (2021). An ERP study of hemispheric differences in perceptual representations of language reveals meaning attribution in the right hemisphere and constituents of the N400-effect. *Brain and Language*, *219*, 104963. https://doi.org/10.1016/j.bandl.2021.104963

Mano, Y., Harada, T., Sugiura, M., Saito, D. N., & Sadato, N. (2009). Perspective-taking as part of narrative comprehension: A functional MRI study. *Neuropsychologia*, *47*(3), 813–824. https://doi.org/10.1016/j.neuropsychologia.2008.12.011

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language*, *85*(3), 451–466. https://doi.org/10.1016/S0093-934X(03)00070-1

McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1–9. https://doi.org/10.1016/j.cogdev.2017.10.001

Minga, J., Johnson, M., Blake, M. L., Fromm, D., & MacWhinney, B. (2021). Making Sense of Right Hemisphere Discourse Using RHDBank. *Topics in Language Disorders*, *41*(1), 99–122. https://doi.org/10.1097/tld.0000000000000244

Mitchell, J. P. (2008). Activity in Right Temporo-Parietal Junction is Not Selective for Theory-of-Mind. *Cerebral Cortex*, *18*(2), 262–271. https://doi.org/10.1093/cercor/bhm051

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The Link between Social Cognition and Self-referential Thought in the Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315. https://doi.org/10.1162/0898929055002418

Morales, M., Patel, T., Tamm, A., Pickering, M. J., & Hoffman, P. (2022). Similar Neural Networks Respond to Coherence during Comprehension and Production of Discourse. *Cerebral Cortex*, bhab485. https://doi.org/10.1093/cercor/bhab485

Morisset, C. E., Barnard, K. E., Greenberg, M. T., Booth, C. L., & Spieker, S. J. (1990). Environmental influences on early language development: The context of social risk. *Development and Psychopathology*, *2*(2), 127–149. https://doi.org/10.1017/S0954579400000663

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, *28*(3), 292–312. https://doi.org/10.1016/0749-596X(89)90035-1

Moscati, V., Zhan, L., & Zhou, P. (2017). Children's on-line processing of epistemic modals*. *Journal of Child Language*, *44*(5), 1025–1040. https://doi.org/10.1017/S0305000916000313

Mullally, E. A., & Maguire, Sinéad L. (2014). Memory, Imagination, and Predicting the Future: A Common Brain Mechanism? *The Neuroscientist*, *20*(3), 220–234.

Nelson, K. (1989). *Narratives from the crib*. Cambridge, Mass. : Harvard University Press.

Newman, R. S., Rowe, M. L., & Ratner, N. B. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173.

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, *184*, 161–170. https://doi.org/10.1016/j.neuroimage.2018.09.010

Nieuwland, M. S. (2012). Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri. *NeuroImage*, *59*(4), 3433–3440. https://doi.org/10.1016/j.neuroimage.2011.11.018

Nieuwland, M. S., & Martin, A. E. (2012). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition*, *122*(1), 102–109. https://doi.org/10.1016/j.cognition.2011.09.001

Ninio, A., Snow, C. E., Pan, B. A., & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of Communication Disorders*, *27*(2), 157–187.

Nippold, M. A., Nehls-Lowe, A., & Lee, D. (2020). Development of Past Tense Counterfactual Sentences: Examining Production and Comprehension in Adolescents and Adults. *Journal of Speech, Language, and Hearing Research : JSLHR*, *63*(3), 764–773. https://doi.org/10.1044/2019_JSLHR-19-00036

Nippold, M. A., Shinham, K., & LaFavre, S. (2020). Mastering the Grammar of Past Tense Counterfactual Sentences: A Pilot Study of Adolescents Attending Low-Income Schools. *Folia Phoniatrica et Logopaedica*, 1–10. https://doi.org/10.1159/000511902

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Nyhout, A., & Ganea, P. A. (2019). Mature counterfactual reasoning in 4- and 5-year-olds. *Cognition*, *183*, 57–66. https://doi.org/10.1016/j.cognition.2018.10.027

Nyhout, A., Henke, L., & Ganea, P. A. (2017). Children's Counterfactual Reasoning About Causally Overdetermined Events. *Child Development*, *90*(2), 610–622. https://doi.org/10.1111/cdev.12913

Ogihara, T. (2000). Counterfactuals, Temporal Adverbs, and Association with Focus. *Semantics and Linguistic Theory*, *10*(0), 115–131. https://doi.org/10.3765/salt.v10i0.3106

O'Neill, D. K., & Atance, C. M. (2000). "Maybe my daddy give me a big piano": The development of children's use of modals to express uncertainty. *First Language*, *20*(58), 029–052. https://doi.org/10.1177/014272370002005802

O'Neill, D. K., & Shultis, R. M. (2007). The emergence of the ability to track a character's mental perspective in narrative. *Developmental Psychology*, *43*(4), 1032–1037. https://doi.org/10.1037/0012-1649.43.4.1032

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258.

Ozturk, O., & Papafragou, A. (2015). The Acquisition of Epistemic Modality: From Semantic Meaning to Pragmatic Interpretation. *Language Learning and Development*, *11*(3), 191–214. https://doi.org/10.1080/15475441.2014.905169

Palmer, F. R. (2001). *Mood and modality*. Cambridge university press.

Papafragou, A. (1998). The Acquisition of Modality: Implications for Theories of Semantic Representation. *Mind & Language*, *13*(3), 370–399. https://doi.org/10.1111/1468-0017.00082

Papafragou, A. (2002). Modality and theory of mind-Perspectives from language development. *S. Barbiers, F. Beukema, W. Van Der Wurff, Modality and Its Interaction with the Verbal System*, 185–205.

Parsons, J. M. (2006). *Positional effects in phonological development: A case study*. Memorial University of Newfoundland.

Pascual-Leone, A., Gates, J. R., & Dhuna, A. (1991). Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. *Neurology*, *41*(5), 697–702. https://doi.org/10.1212/WNL.41.5.697

Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition*, *6*(3), 201–253.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Pepperberg, I. M., Koepke, A., Livingston, P., Girard, M., & Hartsfield, L. A. (2013). Reasoning by inference: Further studies on exclusion in grey parrots (Psittacus erithacus). *Journal of Comparative Psychology*, *127*(3), 272. https://doi.org/10.1037/a0031641

Peterson, C., & McCabe, A. (1983). *Developmental Psycholinguistics: Three Ways of Looking at a Child's Narrative*. Springer. http://ebookcentral.proquest.com/lib/nyulibrary-ebooks/detail.action?docID=3084310

Pettijohn, K. A., & Radvansky, G. A. (2016). Narrative event boundaries, reading times, and expectation. *Memory & Cognition*, *44*(7), 1064–1075. https://doi.org/10.3758/s13421-016-0619-6

Phillips, J., & Knobe, J. (2018). The psychological representation of modality. *Mind & Language*, *33*(1), 65–94. https://doi.org/10.1111/mila.12165

Phillips, J., Morris, A., & Cushman, F. (2019). How We Know What Not To Think. *Trends in Cognitive Sciences*, *23*(12), 1026–1040. https://doi.org/10.1016/j.tics.2019.09.007

Phillips, J., & Norby, A. (2019). Factive theory of mind. *Mind & Language*, 1–24. https://doi.org/10.1111/mila.12267

Portner, P. (2009). *Modality*. OUP Oxford.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Prolific (December 2020). (2022). [Computer software]. https://www.prolific.co

Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology*, *52*(5), 1038–1038. https://doi.org/10.1212/WNL.52.5.1038

Quon, E., & Atance, C. M. (2010). A Comparison of Preschoolers' Memory, Knowledge, and Anticipation of Events. *Journal of Cognition and Development*, *11*(1), 37–60. https://doi.org/10.1080/15248370903453576

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of “nearest possible world.” *Child Development*, *81*(1), 376–389.

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404.

Raichle, M. E. (2015). The Brain’s Default Mode Network. *Annual Review of Neuroscience*, *38*(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Rall, J., & Harris, P. L. (2000). In Cinderella’s slippers? Story comprehension from the protagonist’s point of view. *Developmental Psychology*, *36*(2), 202–208. https://doi.org/10.1037/0012-1649.36.2.202

Redshaw, J., & Suddendorf, T. (2016). Children’s and Apes’ Preparatory Responses to Two Mutually Exclusive Possibilities. *Current Biology*, *26*(13), 1758–1762. https://doi.org/10.1016/j.cub.2016.04.062

Reilly, J., Snitzer. (1982). *The Acquisition of Conditionals in English* [Unpublished PhD dissertation]. University of California.

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cerebral Cortex (New York, NY)*, *25*(11), 4374–4391. https://doi.org/10.1093/cercor/bhv024

Ries, S. K., Dronkers, N. F., & Knight, R. T. (2016). Choosing words: Left hemisphere, right hemisphere, or both? Perspective on the lateralization of word retrieval. *Annals of the New York Academy of Sciences*, *1369*(1), 111–131. https://doi.org/10.1111/nyas.12993

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *13*(1), 73–90. https://doi.org/10.1016/S0885-2014(98)90021-1

Ritter, E., & Wiltschko, M. (2014). The composition of INFL. *Natural Language & Linguistic Theory*, *32*(4), 1331–1386. https://doi.org/10.1007/s11049-014-9248-6

Roberts, C. (1987). *Modal subordination, anaphora, and distributivity* [Ph.D.]. University of Massachusetts.

Roberts, W. A. (2006). The Questions of Temporal and Spatial Displacement in Animal Cognition. In *Comparative cognition: Experimental explorations of animal intelligence* (pp. 145–163). Oxford University Press.

Robertson, D. A., Gernsbacher, M. A., Guidotti, S. J., Robertson, R. R. W., Irwin, W., Mock, B. J., & Campana, M. E. (2000). Functional Neuroanatomy of the Cognitive Process of Mapping During Discourse Comprehension. *Psychological Science*, *11*(3), 255–260. https://doi.org/10.1111/1467-9280.00251

Robinson, E. J., & Beck, S. (2000). What is difficult about counterfactual reasoning. *Children's Reasoning and the Mind*, 101–119.

Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children's Sensitivity to Their Own Relative Ignorance: Handling of Possibilities under Epistemic and Physical Uncertainty. *Child Development*, *77*(6), 1642–1655. JSTOR.

Romero, M. (2014). *'Fake Tense'in Counterfactuals: A Temporal Remoteness Approach*. *2*, 47–63.

Rothmayr, C., Sodian, B., Hajak, G., Döhnel, K., Meinhardt, J., & Sommer, M. (2011). Common and distinct neural networks for false-belief reasoning and inhibitory control. *NeuroImage*, *56*(3), 1705–1713. https://doi.org/10.1016/j.neuroimage.2010.12.052

Rouvoli, L., Tsakali, V., & Kazanina, N. (2019). If They Had Been More Transparent, the Child Would Have Discovered Them More Easily: How Counterfactuals Develop. In *Proceedings of the 43rd Boston University Conference on Language Development (BUCLD 43)* (pp. 549–560). Cascadilla Press. http://www.cascadilla.com/bucld43toc.html

Rowland, C. F., & Fletcher, S. L. (2006). The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language*, *33*(4), 859–877.

Russell, V. L. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* (1.7.2.) [Computer software]. https://CRAN.R-project.org/package=emmeans

Sachs, J., & Nelson, K. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. *Children's Language*, *4*, 1–28.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*(4), 1928–1941. https://doi.org/10.3758/s13428-018-1176-7

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842.

Saxe, R., & Powell, L. J. (2006). It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saylor, M. M., & Baldwin, D. A. (2004). Discussing those not present: Comprehension of references to absent caregivers. *Journal of Child Language*, *31*(3), 537–560. https://doi.org/10.1017/S0305000904006282

Schneider, P., Hayward, D., & Dubé, R. V. (2006). Storytelling from pictures using the Edmonton narrative norms instrument. *Journal of Speech Language Pathology and Audiology*, *30*(4), 224.

Schoot, M. van der, Horsley, T. M., & Lieshout, E. C. D. M. van. (2010). The effects of instruction on situation model construction: An eye fixation study on text comprehension in primary school children. *Educational Psychology*, *30*(7), 817–835. https://doi.org/10.1080/01443410.2010.510600

Schulz, K. (2014). Fake Tense in conditional sentences: A modal approach. *Natural Language Semantics*, *22*(2), 117–144. https://doi.org/10.1007/s11050-013-9102-0

Schurz, M., & Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Frontiers in Psychology*, *6*, 1610. https://doi.org/10.3389/fpsyg.2015.01610

Shetreet, E., Alexander, E. J., Romoli, J., Chierchia, G., & Kuperberg, G. (2019). What we know about knowing: Presuppositions generated by factive verbs influence downstream neural processing. *Cognition*, *184*, 96–106. https://doi.org/10.1016/j.cognition.2018.11.012

Shi, E. R., & Zhang, Q. (2021). Displacement and Evolution: A Neurocognitive and Comparative Perspective. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43). https://escholarship.org/uc/item/2hv2f27d

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge University Press.

Snyder, W. (2007). *Child Language: The Parametric Approach*. OUP Oxford.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human Brain Activity Time-Locked to Narrative Event Boundaries. *Psychological Science*, *18*(5), 449–455. https://doi.org/10.1111/j.1467-9280.2007.01920.x

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510. https://doi.org/10.1162/jocn.2008.21029

Sprott, R. A. (1992). Children's use of discourse markers in disputes: Form-function relations and discourse in child language. *Discourse Processes*, *15*(4), 423–439.

Stalnaker, R. (1996). On the Representation of Context. *Semantics and Linguistic Theory*, *6*(0), 279–294. https://doi.org/10.3765/salt.v6i0.2768

Stalnaker, R. C. (1975). Indicative Conditionals. In *IFS* (pp. 193–210). Springer, Dordrecht. https://doi.org/10.1007/978-94-009-9117-0_9

Stromswold, K. J. (1990). *Learnability and the acquisition of auxiliaries* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/13715

Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1317–1324. https://doi.org/10.1098/rstb.2008.0301

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*(3), 299–313. https://doi.org/10.1017/S0140525X07001975

Suppes, P. (1974). The semantics of children's language. *American Psychologist*, *29*(2), 103.

Tamura, K., & Hashimoto, T. (2012). Displacement in communication. In *The Evolution of Language* (pp. 352–359). WORLD SCIENTIFIC. https://doi.org/10.1142/9789814401500_0046

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure Reasoning in 12-Month-Old Infants as Probabilistic Inference. *Science*, *332*(6033), 1054–1059. https://doi.org/10.1126/science.1196404

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*(1), 127–152. https://doi.org/10.1017/S0305000900004608

Thomason, R. H. (1984). Combinations of Tense and Modality. In D. Gabbay & F. Guenthner (Eds.), *Handbook of Philosophical Logic: Volume II: Extensions of Classical Logic* (pp. 135–165). Springer Netherlands. https://doi.org/10.1007/978-94-009-6259-0_3

Thompson, V. A., & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1154–1170. https://doi.org/10.1037/0278-7393.28.6.1154

Tommerdahl, J., & Kilpatrick, C. D. (2014). The reliability of morphological analyses in language samples. *Language Testing*, *31*(1), 3–18.

Traugott, E. C. (2006). Historical aspects of modality. *The Expression of Modality*, *1*, 107–139.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3

Tulling, M., & Cournane, A. (2019). The role of "fake" past tense in acquiring counterfactuals. *Proceedings of the Amsterdam Colloquium*, 387–396.

Tulling, M., Law, R., Cournane, A., & Pylkkänen, L. (2021). Neural Correlates of Modal Displacement and Discourse-Updating under (Un)Certainty. *Eneuro*, *8*(1), ENEURO.0290-20.2020. https://doi.org/10.1523/ENEURO.0290-20.2020

Turkeltaub, P. E. (2019). A Taxonomy of Brain–Behavior Relationships After Stroke. *Journal of Speech, Language, and Hearing Research*, *62*(11), 3907–3922. https://doi.org/10.1044/2019_JSLHR-L-RSNP-19-0032

Urrutia, M., de Vega, M., & Bastiaansen, M. (2012). Understanding counterfactuals in discourse modulates ERP and oscillatory gamma rhythms in the EEG. *Brain Research*, *1455*, 40–55. https://doi.org/10.1016/j.brainres.2012.03.032

Urrutia, M., Gennari, S. P., & de Vega, M. (2012). Counterfactuals in action: An fMRI study of counterfactual sentences describing physical effort. *Neuropsychologia*, *50*(14), 3663–3672. https://doi.org/10.1016/j.neuropsychologia.2012.09.004

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1–2), 21–81.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* Academic Press.

van Dooren, A., Dieuleveut, A., Cournane, A., & Hacquard, V. (2017). *Learning what must and can must and can mean*. 225–234.

van Dooren, A., Tulling, M., Cournane, A., & Hacquard, V. (2019). *Discovering Modal Polysemy: Lexical Aspect Might Help*.

Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (2013). Counterfactual thinking: An fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience*, *8*(5), 556–564. https://doi.org/10.1093/scan/nss031

Van Hoeck, N., Watson, P. D., & Barbey, A. K. (2015). Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00420

Van Houten, L. J. (1986). *The Role of Maternal Input in the Acquisition Process: The Communicative Strategies of Adolescent and Older Mothers with the Language Learning Children.*

van KLEECK, A., Maxwell, M., & Gunter, C. (1985). A methodological study of illocutionary coding in adult-child interaction. *Journal of Pragmatics*, *9*(5), 659–681.

Vistoli, D., Brunet-Gouet, E., Baup-Bobin, E., Hardy-Bayle, M.-C., & Passerieux, C. (2011). Anatomical and temporal architecture of theory of mind: A MEG insight into the early stages. *NeuroImage*, *54*(2), 1406–1414. https://doi.org/10.1016/j.neuroimage.2010.09.015

von Fintel, K. (2006). Modality and Language. In D. Borchert (Ed.), *Encyclopedia of Philosophy* (pp. 20–27). Macmillan Reference.

von Frisch, K. V. (1967). Honeybees: Do They Use Direction and Distance Information Provided by Their Dancers? *Science*, *158*(3804), 1072–1076. https://doi.org/10.1126/science.158.3804.1072

von Prince, K. (2017). Counterfactuality and past. *Linguistics and Philosophy*, 1–39.

Walker, C. M., & Gopnik, A. (2013). Pretense and possibility—A theoretical proposal about the effects of pretend play on development: Comment on Lillard et al. (2013). *Psychological Bulletin*, *139*(1), 40. https://doi.org/10.1037/a0030151

Wapner, W., Hamby, S., & Gardner, H. (1981). The role of the right hemisphere in the apprehension of complex linguistic materials. *Brain and Language*, *14*(1), 15–33. https://doi.org/10.1016/0093-934X(81)90061-4

Warren, A. R. (1982). *Sex differences in speech to children*.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*(2), 75–95. https://doi.org/10.1016/j.jecp.2007.10.004

Weisberg, D. S., & Gopnik, A. (2013). Pretense, Counterfactuals, and Bayesian Causal Models: Why What Is Not Real Really Matters. *Cognitive Science*, *37*(7), 1368–1381. https://doi.org/10.1111/cogs.12069

Weisberg, D. S., & Gopnik, A. (2016). Which Counterfactuals Matter? A Response to Beck. *Cognitive Science*, *40*(1), 257–259. https://doi.org/10.1111/cogs.12241

Weismer, S. E., Venker, C. E., Evans, J. L., & Moyle, M. J. (2013). Fast mapping in late-talking toddlers. *Applied Psycholinguistics*, *34*(1), 69–89.

Weist, R. M., & Zevenbergen, A. A. (2008). Autobiographical memory and past time reference. *Language Learning and Development*, *4*(4), 291–308.

Wells, G. (1981). Language as interaction. *Learning through Interaction: The Study of Language Development*, 22–72.

Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Cohn.

Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., & Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. *NeuroImage*, *47*(1), 360–366. https://doi.org/10.1016/j.neuroimage.2009.04.037

Wilson, E. O. (1962). Chemical communication among workers of the fire ant Solenopsis saevissima (Fr. Smith) 1. The Organization of Mass-Foraging. *Animal Behaviour*, *10*(1), 134–147. https://doi.org/10.1016/0003-3472(62)90141-0

Woolley, J. D., & Cox, V. (2007). Development of beliefs about storybook reality. *Developmental Science*, *10*(5), 681–693. https://doi.org/10.1111/j.1467-7687.2007.00612.x

Woolley, J. D., & Ghossainy, M. E. (2013). Revisiting the Fantasy–Reality Distinction: Children as Naïve Skeptics. *Child Development*, *84*(5), 1496–1510. https://doi.org/10.1111/cdev.12081

Woolley, J. D., & Wellman, H. M. (1993). Origin and truth: Young children's understanding of imaginary mental representations. *Child Development*, *64*(1), 1–17. https://doi.org/10.2307/1131434

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *NeuroImage*, *41*(4), 1408–1425. https://doi.org/10.1016/j.neuroimage.2008.03.062

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: Where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, *22*(3), 181–192. https://doi.org/10.1038/s41583-020-00420-w

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same Story, Different Story: The Neural Representation of Interpretive Frameworks. *Psychological Science*, *28*(3), 307–319. https://doi.org/10.1177/0956797616682029

Yong, Q. (2016). A corpus-based study of counterfactuals in Mandarin. *Language and Linguistics*, *17*(6), 891–915.

Zehr, J., & Schwarz, F. (2018). *Penncontroller for Internet Based Experiments (IBEX)*. https://doi.org/10.17605/OSF.IO/MD832

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage*, *34*(3), 1310–1316. https://doi.org/10.1016/j.neuroimage.2006.08.047

Ziegler, F. V., & Acquah, D. K. (2013). Stepping into someone else's shoes: Children create spatial mental models from the protagonist's point of view. *European Journal of Developmental Psychology*, *10*(5), 546–562. https://doi.org/10.1080/17405629.2012.744689

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, *44*, 35–62.

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*(4), 1028–1034. https://doi.org/10.3758/s13423-015-0864-x

Zwaan, R. A., & Madden, C. J. (2004). Updating situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 283–288. https://doi.org/10.1037/0278-7393.30.1.283

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162