New Topics in Monolingual and Multilingual Language Acquisition

New Topics in Monolingual and Multilingual Language Acquisition

Edited by

Anca Sevcenco and Veronica Tomescu

Cambridge Scholars Publishing

New Topics in Monolingual and Multilingual Language Acquisition

Edited by Anca Sevcenco and Veronica Tomescu

This book first published 2025

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-0364-5486-9
ISBN (Ebook): 978-1-0364-5487-6

TABLE OF CONTENTS

INTRODUCTION

The current volume is a collection of papers presented at the 8th edition of the Bucharest Colloquium of Language Acquisition in November 2023, at the University of Bucharest, Romania. They present issues that pertain to various contexts of acquisition (monolingual, bilingual, heritage language, second and third language) and focus on an extended range of domains, from syntax to the interfaces between syntax and semantics, pragmatics and discourse.

Bilingualism and second language learning are of particular interest to researchers because they provide a field of study for comparative linguistics as well and an opportunity for incursions into cognitive sciences. An impressive body of evidence has been amassed and is continually being added to prove that bilingualism is an asset not only from a practical point of view, but also from a cognitive perspective. For example, children who are bilingual from birth frequently outperform their monolingual peers in tasks that presuppose the suppression of distractors or require disentangling conflicting pieces of information and have superior metalinguistic ability. Managing their two languages forces them quite early to develop certain non-language specific cognitive control mechanisms (Luk, de Sa & Bialystok, 2011; Barac & Bialystok, 2012; Cuervo & Pérez-Leroux, 2015, a.o.). But bilinguals are also a victim of their own success. The downside of these multiple accomplishments is that under certain conditions the simultaneous management of the two languages can result in delay and processing difficulties (Gawlitzek-Maiwald & Tracy, 2005; Serratrice et al., 2011; Patuto, Repetto & Müller, 2011, a.o.).

One important question that pervades research into bilingualism, as well as second and third language learning, is whether these effects are an inherent result of bilingualism itself or whether language type, relatedness or distance may play a part and to what degree. Whether delay might be due simply to processing difficulties or whether exact, purely linguistic factors can be pinpointed as the main cause. In this sense, it has been successfully argued that for cross-linguistic influence to occur it is necessary that there should be structural overlap between the two languages in question (Hulk & Müller, 2000; Müller & Hulk, 2001; Schmitz, Patuto & Müller, 2012), resulting in qualitative or quantitative differences between the productions of bilingual children compared to their monolingual peers.

Originally developed to account for difficulties in the interpretation of discourse-governed anaphora by speakers of L2 Italian and later extended to other categories of bilinguals, The Interface Hypothesis (Sorace & Filiaci, 2006; Sorace, 2011) predicts that structures at the interface between syntax and other cognitive domains will be acquired later and with greater difficulty, instability and optionality than structures that only require narrow syntactic computation. Importantly, phenomena at the interface can be vulnerable both when cross-linguistic interference can occur, but also when there is little likelihood that one language was influenced by the other, purely as an independent side-effect of bilingualism (Sorace & Serratrice, 2009).

The papers by Avram, Pupp & Tomescu, Vilanova & Figura, Babei-Popa & Toșu in the current volume aim to shed further light on the causes and nature of cross-linguistic influence that accompanies bilingual acquisition as well as language learning, to what extent this influence is beneficial or, on the contrary, hinders progress, and whether it holds true that phenomena at the interface might be more vulnerable than purely syntactic ones. The topics that these authors discuss are quite diverse: differential object marking, anaphora resolution, language change and acquisition in reduced input circumstances.

Differential object marking (DOM) is a phenomenon at the syntax-discourse interface which has been extensively studied in both monolingual and bilingual learning contexts. It involves separate and overt case marking of the direct object constrained by various features such as topicality, specificity, definiteness, affectedness or animacy (Bossong, 1991, 1998; Aissen, 2003). In some languages, such as Romanian, the phenomenon has additional complications because it also interacts with clitic doubling.

Avram, Pupp & Tomescu, in *Input variation under diachronic instability: The acquisition of DOM in L1 and 2L1 Romanian*, study the use of Romanian DOM and clitic doubling in a collection of narratives produced by 2 groups of children of kindergarten age, monolingual Romanian and Romanian-Hungarian bilinguals. Most contexts produced by the children in the two narratives are definite nouns, whose marking is optional. Further, in such contexts, clitic doubling is also optional, although this particular phenomenon is currently undergoing diachronic change: the innovative option no longer accepts the absence of the clitic in the presence of the differential marker PE. The differences between the two groups are twofold. Firstly, the bilinguals have a higher preference for marking definite nouns in optional contexts than the monolinguals. The authors explain the results in terms of a boosting effect from Hungarian, a language which, while lacking a DOM system similar to Romanian, does have a definiteness

agreement marker on the verb, obligatory with definite direct objects (and other similar object categories), as well as identical overt Accusative case marking on all direct objects. It is theorized that a superficial similarity between Romanian DOM and these two properties of Hungarian may have enhanced the acquisition and the frequency of use of Romanian PE. Importantly, the bilinguals do not overextend the use of the marker in non-legitimate contexts, proof that they do have a correct understanding of the pragmatic constraints governing its use.

Regarding optional clitic doubling, however, the bilinguals lag behind the monolinguals, preferring the more traditional option of single PE. The authors propose cross-linguistic influence (Hulk & Müller, 2000; Müller & Hulk, 2001) as an explanation. Since Hungarian has only one option for marking objects (a suffix), and Romanian has two, that is there is structural overlap between Hungarian and Romanian, the bilinguals are more likely to opt for the shared structure, in this case a single marker. We see, therefore, cross-linguistic interference working in two opposite directions: on the one hand enhancing the production of PE in optional contexts, on the other delaying the use of clitic doubling as differential marker. But the differences were only related to preference; the bilinguals produced no errors in obligatory contexts and no overextensions to illicit contexts.

Heritage speakers are a subclass of bilinguals, though certainly the category of bilinguals is not homogeneous. Prototypical or "genuine" bilinguals might be broadly defined as individuals who learned both languages from birth and had more or less equal opportunities of exposure and use, both at home and in the community, all throughout childhood and later in life. Simply put, heritage languages are languages spoken by immigrants and their descendants, different from the societal or official community language (Montrul, 2016). Generally speaking, the heritage speakers have fewer opportunities to use the language as they grow up; therefore, the input that they receive is limited and unbalanced, inescapably leading to the gradual attrition of the heritage language. Additionally, the heritage language may be incompletely acquired, and/or may undergo possible L2 learning effects (Montrul, 2016).

Bianca Babei-Popa follows up on the acquisition of Romanian DOM but this time in a heritage language context, Romanian-Dutch and Romanian-Italian. Her results differ from Avram, Pupp & Tomescu (this volume) with respect to DOM, since both groups of heritage Romanian-speaking children underuse overt marking of the direct object, but not with respect to clitic doubling, which is dispreferred across the board.

The Romanian-Dutch bilinguals in *Differential object marking at the syntax-pragmatics interface: Data from child heritage Romanian in the*

Netherlands underuse both DOM and clitic doubling compared to adults and monolingual Romanian children, but make no errors and overextensions. Cross-linguistic influence (Hulk & Müller, 2000) might account for the underuse of DOM, in that Dutch has only unmarked direct objects, and the bilinguals preferred the shared unmarked option in both their languages.

In *Differential object marking in child heritage Romanian in contact with Italian: An experimental study*, Babei-Popa investigates DOM in obligatory, optional and illicit contexts in an elicited production task with a group of heritage Romanian-Italian bilinguals. Like the Romanian-Dutch heritage speakers, the Romanian-Italian children underused DOM in obligatory contexts and they also dispreferred clitic doubling. No overextensions were recorded. Babei-Popa accounts for her results in terms of DOM being an interface phenomenon, hence vulnerable in bilingual acquisition; differences are explained as being an effect of bilingualism itself. Italian (or the varieties spoken in the communities where the children in the study live) does not have DOM, therefore cross-linguistic influence is ruled out. But often, irrespective of the language combination, bilinguals fare worse or at least there are visible quantitative differences compared to monolingual production, especially with respect to phenomena that necessitate interpretation or constant updating of contextual information.

This is precisely the result in Alicia Toșu's paper, *Anaphora resolution by child heritage speakers of Romanian living in Italy*. She accounts for the delay exhibited by Romanian-Italian children in the interpretation of anaphora in terms of vulnerability at the syntax-discourse interface. Child heritage Romanian speakers did not behave like monolingual adults and children in the interpretation of null subjects. This could not have been transfer from Italian, since, regarding null subjects, Romanian and Italian are not dissimilar. Strikingly, the bilinguals failed to behave target-like with demonstrative pronouns as well, which are cross-linguistically prone to be interpreted as coindexed with the closest antecedent. Since cross-linguistic interference once again could not have had a negative effect, and it failed to have a positive effect, the non-target behaviour of the bilinguals could only have been a result of bilingualism itself.

The study on heritage speakers of Judeo-Spanish living in Bulgaria by Vilanova & Figura analyses three separate phenomena: clitics, adjective placement and aspect. Clitics, a purely syntactic phenomenon, have not been found to be problematic, in line with the Interface Hypothesis (Sorace & Filiaci, 2006). Adjective placement and aspect, however, are more vulnerable. Adjective placement shows definite transfer from the dominant language (Bulgarian): a shift towards postnominal adjectives. This is a

phenomenon situated at the syntax-semantics interface in Judeo-Spanish, since the position of the adjective can affect its semantics; but these heritage speakers were no longer aware of these semantic restrictions, preferring the Bulgarian word-order. Therefore, the authors rely on the Interface Hypothesis to account for the results. Input conditions may also have played a part, since this is a phenomenon usually acquired later in childhood, and the acquisition path for these heritage speakers was incomplete. Aspect is another vulnerable phenomenon. But the authors affirm that in this case cross-linguistic interference is not the causing factor, since there is little overlap between the two languages. The authors attribute the changes in Judeo-Spanish not necessarily to the influence of Bulgarian, but to an overgeneralization – in the sense of Yang (2016)'s Tolerance Principle – that the heritage speakers resort to, extending the use of various forms to non-target-like contexts.

The volume also approaches three topics in the domains of monolingual second and third language acquisition with Romanian-speaking learners as participants.

Comprehending which-*questions in real-time: An eye-tracking study with Romanian-speaking children and adults* by Anamaria Bentea focuses on monolingual Romanian. The author investigates the well-known asymmetry in the comprehension of subject and object *which*-questions in child Romanian. She revisits the role of number agreement and DOM in the comprehension of *which*-questions. Number turns out to have a facilitative effect in comprehension, observed incrementally during processing. Case marking also supports the process, but not to the extent that it cancels the subject-object asymmetry.

Getting the dative alternation in L2 English right: A study on Romanian learners by Anca Sevcenco casts a glance at a complex topic in second language learning. Complexity arises because many factors need to be considered before they get internalized by the learners' grammatical system. Above and beyond the syntax of the double object and prepositional object constructions, semantics (verb properties such as caused possession, in the case of double object verbs, versus caused motion, for prepositional object verbs) and discourse aspects (the organization of information in the two syntactic frameworks) guide the learning process. In addition, the acquisition of the dative alternation in both L1 and L2 English poses a problem of inductive learning (see Yang & Montrul, 2017). The crux of the matter for inductive learning lies in how learners fall back from a permissive grammar, which rules in double object constructions with illicit verbs, to the target system, which constrains this structure. Various patterns of transfer effects are also at play in the learning process. Anca Sevcenco

brings up a discussion about transfer effects and the existence of a prepositional object – double object asymmetry (observed in other language pairs too). The main outcome is that transfer applies selectively because it targets members of specific verb classes (for instance verbs such as *explain* and *demonstrate*, from the Latinate verb class). The asymmetry also comes down to transfer of the morphological properties of Romanian double object constructions.

The learning of a third language (L3) has received increased attention in recent years (De Angelis, 2007, a.o.). It has been found that L3 learning differs from both L1 (and 2L1) and L2 learning in several ways. It should be evident that the L3 learner relies not only on his or her native language, but on the other non-native language that he or she has acquired (the level of proficiency is not relevant), greatly increasing opportunities for cross-linguistic influence. The L3 learner can additionally rely on explicit learning strategies previously employed in L2 learning. One important question that has captured the interest of researchers in the field has been which language (L1 and L2) – and to what extent – is, or can be, the source for cross-linguistic influence. While factors such as proficiency in the second language or frequency and recency of use are not trivial, of greater importance is the similarity or distance between the three languages in question (Westergaard, 2019; Kolb, Mitrofanova & Westergaard, 2022). Typological and structural closeness may determine which structures are chosen as model, or which structures may hinder the acquisition process, irrespective of whether they belong to L1 or L2. Importantly, superficial similarities between unrelated phenomena may be identified by the learner and may help the acquisition process, or, conversely, trick and lure the learner – so to speak – down an erroneous path.

Sorana Banu, in *-Te i- in the written narratives of Romanian learners of L3 Japanese,* uses a collection of frog stories produced by L1 Romanian speakers, with English as L2. She identifies cross-linguistic influence (Hulk & Müller, 2000; Müller & Hulk, 2001) from Romanian into Japanese as cause for the quantitative differences in the production of the overt aspectual marker *-te iru-*. The marker is underused by the L3 speakers, probably under the influence of Romanian, where aspect is not overtly expressed by morphology. Since their L1 has only one option to express aspect, and English and Japanese two, the learners prefer the simpler system and hesitate to employ the overt aspectual marker in their L3. Influence from Romanian also appears to consist in a preference for the use of past tense forms in narratives, even though present tense is the preferred option for Japanese (Inaba, 2003). A comparison with previous studies making use of

English (Tănase, 2023) and Romanian (Buja, 2009) frog stories guides the author to this conclusion.

References

Aissen, Judith (2003). Differential Object Marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21: 435–483.

Barac, Raluca and Ellen Bialystok (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development*, 83–2: 413–422

Bossong, Georg (1991). Differential object marking in Romance and beyond. In D. Wanner & D. A. Kibbee (eds.), *New Analyses in Romance Linguistics. Selected Papers from the XVIII Linguistic Symposium on Romance Languages 1988*, Amsterdam: John Benjamins, 143–170.

—. (1998). Le marquage de l'experient dans les langues d'Europe. In J. Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, Berlin : Mouton de Gruyter, 259–294.

Buja, Elena (2009). *Relating events in narrative: A Case Study of Romanian*. Saarbrücken: VDM Dr. Müller Verlag AG & Co.KG.

Cuervo, Maria Cristina and Ana Teresa Pérez-Leroux (2015). The sum is more than its parts: Acquisition of clitic clusters in Spanish. In R.Klassen, J.M.Liceras and E.Valenzuela (eds.) *Hispanic Linguistics at the Crossroads: Theoretical linguistics, language acquisition and language contact, Proceedings of the Hispanic Linguistics Symposium*, Amsterdam: John Benjamins, 145–168.

De Angelis, Gessica (2007). *Third or additional language acquisition*. Clevedon: Multilingual Matters.

Gawlitzek-Maiwald, Ira and Rosemarie Tracy (2005). All children start out as multilinguals. In J. Cohen, K. T. McAlister, K. Rolstad and J. MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*, Somerville MA: Cascadilla Press, 875–889.

Hulk, Aafke and Natascha Müller (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism*: *Language and Cognition* 3(3): 227–244.

Inaba, Midori (2003). Anchor tense in Japanese narrative. *The Bulletin of Aichi University of Education (Humanities/Societies)* 52: 51–61.

Kolb, Nadine, Natalia Mitrofanova and Marit Westergaard (2022). Crosslinguistic influence in child L3 English: An empirical study on Russian-German heritage bilinguals. *International Journal of Bilingualism* 26(4): 1–26.

Luk, Gigi, Eric de Sa and Ellen Bialystok (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition* 14(4): 588–595.

Müller, Natascha and Aafke Hulk (2001). Cross-linguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition* 4(1): 1–21.

Montrul, Silvina (2016). *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.

Patuto, Marisa, Valentina Repetto and Natascha Müller (2011). Delay and acceleration in bilingual first language acquisition: really the same source? In E. Rinke and T. Kupisch (eds.), *The development of Grammar: Language Acquisition and Diachronic Change. In honour of Jürgen M. Meisel,* Amsterdam: John Benjamins, 231–261.

Schmitz, Katrin, Marisa Patuto and Natascha Müller (2012). The null-subject parameter at the interface between syntax and pragmatics: Evidence form bilingual German-Italian, German-French and Italian-French children. *First Language* 32(1–2): 205–238.

Serratrice, Ludovica, Antonella Sorace, Francesca Filiaci and Michela Baldo (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness. *Bilingualism: Language and Cognition* 11(2): 239–257.

Sorace, Antonella (2011). Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism* 1(1): 1–33.

Sorace, Antonella and Francesca Filiaci (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22(3): 339–368.

Sorace, Antonella and Ludovica Serratrice (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *International Journal of Bilingualism* 13(2): 195–210.

Tănase, Alexandru (2023). The Progressive in L2 English in an L1 Romanian Context. MA dissertation, University of Bucharest.

Yang, Charles (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press.

Yang, Charles and Silvina Montrul (2017). Learning datives: The Tolerance Principle in monolingual and bilingual acquisition. *Second Language Research* 33(1): 119–144.

Westergaard, Marit (2019). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37: 379–407.

CHAPTER ONE

COMPREHENDING *WHICH*-QUESTIONS IN REAL-TIME: AN EYE-TRACKING STUDY WITH ROMANIAN-SPEAKING CHILDREN AND ADULTS

ANAMARIA BENTEA

1. Introduction

A common finding across languages is that subject *which*-questions (1) are generally easier to comprehend than object *which*-questions (2).

(1) Which grandmother is hugging the girl?

(2) Which grandmother is the girl hugging?

Both structures involve syntactic movement of a wh-phrase (*which grandmother*) to the beginning of the sentence, but this movement takes place from the subject position of the verb *hug* in (1) and from the object position in (2). In languages that instantiate a Subject-Verb-Object order, this movement gives rise to a non-canonical word order in object *which*-questions, as the object (*which grandmother*) precedes the subject (*the girl*). Successful sentence comprehension therefore requires to reconstruct the correct syntactic and interpretative dependency between the moved wh-phrase, also called "filler", and the position where it is generated in the structure, also known as "gap" (e.g., Bever, 1970).

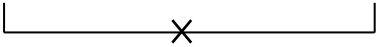
A subject-object asymmetry in comprehension has been observed in numerous studies with typically-developing children, primarily using offline measures such as sentence-picture matching or character-selection tasks, which assess children's final interpretation of *wh*-questions (for Dutch: Metz, van Hout & van der Lely, 2010; Schouwenaars, van Hout &

Hendriks, 2014; for English: O'Grady, 1997; for French: Bentea, Durrleman & Rizzi, 2016; Jakubowicz & Gutierez, 2007; for German: Biran & Ruigendijk, 2015; Roesch & Chondrogianni, 2015, 2016; for Greek: Stavrakaki, 2006; for Hebrew: Friedmann & Novogrodsky, 2011; Friedmann, Belletti & Rizzi, 2009; Friedmann, Rizzi & Belletti, 2017; for Italian: De Vincenzi et al., 1999; for Romanian, Bentea, 2017; a.o).

From a processing perspective, it has been argued that the difficulty with object questions like in (2) stems from a preference for interpreting the wh-phrase as a subject rather than an object (e.g. van Gompel, 2013). According to the "active filler hypothesis" (Clifton & Frazier, 1989; Frazier & Flores d'Arcais, 1989), adapted to language acquisition (Avrutin, 2000; Deevy & Leonard, 2004), gaps are interpreted at the first possible location in the sentence, in an effort to simplify sentence structure, thus allowing for faster, more efficient processing. However, if the initial interpretation proves incorrect as more of the sentence unfolds, reanalysis based on the incoming linguistic information is required to arrive at the correct interpretation. This approach explains why both children and adults encounter difficulties with object-first structures like in (2): the fronted wh-phrase is initially interpreted as subject during incremental processing, because the subject position is the first possible gap available in the structure. Upon encountering the second noun phrase/NP, this initial interpretation needs to be revised. While adults can revise their early interpretive commitment to a subject interpretation, it has been argued that children struggle more with revising the initial interpretation they assigned to the sentence (Trueswell, et al., 1999) when the disambiguating cues appear in a sentence-final position (Choi & Trueswell, 2010; Omaki et al., 2014).

The difficulty with object *which*-questions can also be explained by the similarity in morphosyntactic features between the fronted wh-phrase, and the subject NP that intervenes in the syntactic dependency between the wh-phrase and the position where it originates in the sentence. The presence of this intervening element hinders the assignment of the correct thematic roles and leads to intervention effects similar to those found in adult grammars and accounted for through the syntactic locality principle of Relativized Minimality/RM (Rizzi, 1990, 2004). This account argues that intervention effects arise when both the object and the subject are lexically specified (they have an NP feature) and the object crosses over the subject. Thus, in object *which*-questions, the lexically-restricted subject acts as an intervener and disrupts the establishment of the correct syntactic dependency between the wh-object and the gap, because the intervener could be considered as a potential filler for the gap. This is schematized in (3) below:

[NP] [NP]
(3) Which grandmother is the girl hugging ____?



A closer investigation of the acquisition of these structures has shown that the intervention effects present in object *which*-questions are weakened when the wh-phrase and the subject mismatch in features (e.g., gender, number), because these mismatches act as cues that can facilitate comprehension. However, these effects are selective, as not all features are relevant for the computation of similarity (Friedmann, Belletti & Rizzi, 2009; Belletti et al., 2012; Friedmann, Rizzi & Belletti, 2017). For example, while a mismatch in number plays a role in Italian (Adani et al., 2010), Belletti et al. (2012) found that a mismatch in gender features alleviates comprehension difficulties in Hebrew, but not in Italian. The authors conclude that only syntactically active features, meaning features expressed in the verbal inflection of the finite verb and able to trigger phrasal movement, are taken into account for the computation of similarity (see also Biondo et al., 2023, for the selective effect of number and gender features on the processing of relative clauses in Italian adults).

In this paper I focus on *which*-questions in Romanian, where object wh-phrases are marked with a differential object marker (DOM) and number agreement is present on the verb in order to assess (i) whether mismatches in morphosyntactic features like DOM and number reduce intervention effects in offline comprehension in monolingual children and (ii) whether children use these morphosyntactic features as cues to incrementally interpret sentences during online processing. The paper is structured as follows: section 2 describes previous findings from monolingual acquisition regarding the effect of number agreement, case and DOM on the offline and online comprehension of *which*-questions, section 3 outlines the research questions, sections 4 and 5 present the study design and the results, and section 6 discusses the main findings.

2. The role of number and case/DOM in the comprehension of *which*-questions

Various studies have examined the role that mismatches in number or case features and the presence of DOM have on both the offline and the online comprehension of *which*-questions. Research shows that number mismatches facilitate the comprehension of object *which-questions*, thus adding to findings on how number mismatches impact on the comprehension of object

relative clauses (e.g., Adani et al., 2010; Bentea & Durrleman, 2017; Contemori & Marinis, 2014).

In a visual-world eye-tracking study with 31 monolingual English-speaking children aged 5 to 7 years (average age: 6;4, SD: 0.1), Contemori, Carlson & Marinis (2018) found that children had a strong preference to interpret object questions as subject questions, as revealed by their comprehension accuracy scores. However, when comparing object *which-questions* with number mismatch to those without, children performed significantly better with number-mismatched questions in offline accuracy measures. The gaze data also showed that children took longer than adults and had more difficulty revising an initial subject interpretation when processing object *which-questions*. Nonetheless, children displayed a faster increase in looks to the target image matching the question when the number marking on the auxiliary clearly indicated an object interpretation (e.g., *Which cow are the goats pushing?* vs. *Which cow is the goat pushing?*). The results thus suggest that the mismatch in number facilitated children's real-time processing of object *which*-questions and reduced intervention effects.

While number mismatch assists English-speaking children with the comprehension of object *which*-questions, various studies have shown that Dutch-speaking children do not rely on number agreement between the wh-phrase and the verb to overcome a subject-first interpretation bias. This was reported in two studies by Metz, van Hout & van der Lely (2010, 2012), which used an offline picture-selection task with children aged 5 to 9 years old. Their results clearly show that all groups found object questions more challenging and that comprehension improved gradually with age. However, even by age 9, children's understanding of these questions did not match adult levels. The authors concluded that number agreement alone is insufficient to overcome the preference for a subject-first interpretation.

Schouwenaars, van Hout & Hendriks (2014) also observed that number mismatch had no impact on Dutch-speaking children's understanding of object *which*-questions. Their study assessed 23 children aged 6;6 to 7;10 (mean age: 7;1) and revealed that, although the children could use number agreement to identify subjects in simple declarative sentences, they had difficulties integrating number agreement in the comprehension of more complex structures like *which*-questions. The children displayed very low comprehension accuracy for object questions disambiguated by number agreement and often misinterpreted them as subject questions.

As for the effect of case marking on the comprehension of *which*-questions, Friedmann, Belletti & Rizzi (2009) found that the presence of *et*, an accusative case marker, before the wh-element in Hebrew object questions did not reduce children's difficulty with the comprehension of

which-questions. The 22 children (aged 3;7–4;10, average age 4;3) included in their study performed at chance level (58% accuracy) with object *which*-questions during an offline sentence-picture matching task, whereas they were more accurate (84% correct responses) with both subject *who* and *which*-questions, as well as with object *who*-questions.

In a subsequent study with 28 Hebrew-speaking children aged 3;0 to 6;5, Friedmann, Rizzi & Belletti (2017) showed that the presence of the overt case marker *et* did not help with the interpretation of object *which*-questions. The children found both questions with and questions without *et* difficult to interpret and often selected the picture in which the roles of the arguments were reversed. In other words, children misinterpreted object questions as subject questions, regardless of the presence of case marking, indicating that they had difficulties associating the two NPs with the proper thematic roles even when case was overtly marked on the object. Friedman, Rizzi & Belletti (2017) interpreted these findings as evidence for the idea that only syntactically active features enter the computation of similarity and can relieve intervention effects in object dependencies like *which*-questions. Case marking, much like gender in Italian (see Belletti et al., 2012), is not a relevant feature as it does not trigger phrasal movement. As a result, children have more difficulties using case mismatches to establish the syntactic dependency between the moved wh-filler and the object gap in the presence of a lexically-restricted intervening subject, which leads to difficulties with the correct interpretation of object *which*-questions.

For German, Biran & Ruigendijk (2015) have shown that case morphology does not eliminate the subject-object asymmetry observed in the comprehension of *which*-questions. They assessed the comprehension of German object questions, which share a similar word order with Romanian, namely fronted wh-word, verb, and post-verbal subject. The participants, 49 children between the ages of 3;6 and 6;8 (average age 5;2, SD = 10.2), gave only 51% accurate answers to object questions. Interestingly, children performed better on object *which*-questions where both NPs shared the same gender compared to questions with NPs of different genders. This improvement was attributed to the presence of two masculine NPs which are clearly distinguished by case marking in German contrary to feminine and neuter nouns which are ambiguous. Based on these findings, the authors concluded that case marking positively affected comprehension of *which*-questions in German. However, while German-speaking children were able to use case marking to perform to above-chance levels of accuracy, they were still unable to fully overcome the intervention effects associated with object *which*-questions.

Roesch & Chondrogianni (2015) explored how case marking influences the comprehension of subject and object *who* and *which*-questions in 4 to 5-year-old German-speaking children, using a character-selection task. In their study, case was marked either on both the wh-element and the embedded NP (mirroring the same-gender condition from Biran & Ruigendijk's study), or only on the wh-word. Their findings revealed that children showed better accuracy in comprehending subject questions compared to both object *who* and *which*-questions. Whereas the presence of case marking did not affect the comprehension of object *who*-questions, children were more accurate with *which*-questions when both the wh-word and the post-verbal subject were marked for case as compared to when only the sentence-initial *wh*-phrase was marked for case. The authors argued that preschool-aged German-speaking children use case marking as a cue to assign thematic roles in complex sentences, aligning with findings from earlier studies on relative clauses in older children (e.g., Arosio et al., 2012). Nonetheless, children's response accuracy was higher when case information appeared both at the beginning and at the end of the question.

Another study that revealed a facilitating role for case on the comprehension of object *which*-questions was Sauerland et al. (2016). This cross-linguistic study with 392 children aged 4;10 to 6;0 across 18 languages examined the comprehension of subject and object *who*- and *which*-questions and found that overt case marking on the moved wh-phrase was among the language-specific factors that significantly improved children's comprehension of object questions, with languages featuring overt case marking yielding higher accuracy than those without. Notably, the authors also reported no effect for the NP feature, potentially due to the design of the study, which included a consistent number mismatch: the moved wh-phrase was singular and the second NP was plural in all the items. As a mismatch in number can ameliorate intervention effects (Adani et al., 2010; Contemori, Carlson & Marinis, 2018), the difference in number may have overridden any effect of the NP feature. However, Sauerland et al.'s data indicate that children's difficulties with object *which*-questions was greatly reduced in the presence of both case marking and number agreement.

Finally, in a study combining the visual world eye-tracking paradigm with a picture-selection task, Schouwenaars, Hendriks & Ruigendijk (2018) analyzed the comprehension patterns of subject and object *which*-questions disambiguated by case and number agreement in monolingual German children aged between 7;05 and 10;09 (average age 9;01, SD = 12.7). For offline accuracy, Schouwenaars et al. found no significant difference in responses between object questions disambiguated by case and those

disambiguated by number agreement (so when the two NPs mismatched in number, but were ambiguous with respect to case marking). However, individual accuracy patterns suggested that case marking was a more effective cue than number agreement on the verb for offline comprehension. Specifically, four children relied on case marking but not verb agreement when interpreting the questions, while no child showed the opposite pattern. Eye-tracking data indicated that, while case and number agreement helped reanalysis of initially misinterpreted object questions, children were slower when disambiguation was based only on number agreement compared to when both case and number cues were available. That is, children initially misinterpreted object questions as subject questions when only number agreement on the verb was used for disambiguation. In contrast, when the *which*-phrase was marked for case, children did not make these initial incorrect interpretations. They displayed a preference to look at the target picture, although this preference emerged more slowly and later in time compared to adults.

The findings for Romanian also suggest that marking case on the wh-pronoun does not improve comprehension with object *which*-questions. Like Hebrew, Romanian has the differential object marker/DOM *pe* that precedes the wh-object. Contrary to Hebrew, the subject appears in a post-verbal position in *wh*-questions in Romanian, such that only DOM *pe* on the fronted *wh*-word indicates that the structure is an object, rather than a subject question, when the subject and the object match in number. Bentea (2017) tested two groups of monolingual Romanian-speaking children, 21 children aged 3;8 to 5;2 (average age 4;4, SD = 0.6) and 23 children aged 5;8 to 7;2 (average age 6;5, SD = 0.6) on the comprehension of subject and object *who* and *which*-questions with a match in number. The author found a subject-object asymmetry only in the comprehension of *which*-questions, but not in the comprehension of *who*-questions, although DOM *pe* appears in both types of questions.

As for the effect of number agreement in Romanian, Avram & Sevcenco (2020) explored the role of number mismatch on the offline comprehension of subject and object *who*- and *which*-questions in a study involving 23 monolingual 5-year-old Romanian-speaking children (average age 5;4, SD = 4.7). Their findings revealed no significant difference between subject and object *which*-questions. While the initial results suggested that number mismatches reduced the intervention effects observed in object *which*-questions, the authors proposed that this performance pattern might also reflect a specific vulnerability associated with subject *which*-questions. This interpretation was further supported by a separate study with another group of 23 monolingual Romanian-speaking children (mean age = 5;4, SD = 2.5).

In this study, children demonstrated higher comprehension accuracy for subject *which*-questions in a number match task compared to the number mismatch task. On the other hand, their performance on object *which*-questions was similar across both match and mismatch conditions.

Summarising, the effect of number agreement and case marking/DOM on the comprehension of *which*-questions seems to be subject to language variation. Most of the studies reported in this section used offline tasks to probe children's comprehension of subject and object questions. The only two studies so far that have analysed children's real-time processing of *which*-questions (Contemori, Carlson & Marinis, 2018, for English and Schouwenaars, Hendriks & Ruigendijk, 2018, for German) suggest that children show a subject-initial interpretation of object *which*-questions, but are subsequently able to revise this towards an object interpretation by making use of morphosyntactic cues like case and number marking on the NPs and on the verb during online processing. Importantly, Schouwenaars, Hendriks & Ruigendijk (2018) showed that the presence of case marking on the fronted wh-phrase in German seems to prevent children from misanalysing object questions as subject questions.

3. The current study

Building on the visual-world eye-tracking study by Contemori, Carlson & Marinis (2018), in this study I focus on Romanian and examine the processing of *which*-questions in monolingual child and adult Romanian speakers. To my knowledge, this is the first eye-tracking study with wh-questions in Romanian-speaking children. The study aimed to investigate the real-time and offline processing of subject and object *which*-questions in both children and adults and to explore the role of DOM and of number agreement for successful parsing and interpretation.

The specific research questions were:

1. Do morphosyntactic cues like DOM and number agreement facilitate intervention effects in the offline comprehension of *which*-questions?

2. What impact do these morphosyntactic cues have on the real-time comprehension of *which*-questions in children and adults? Do children and adults use them for incremental interpretation?

For the offline interpretation, the prediction was that a mismatch in number would alleviate the difficulties that children might encounter with object *which*-questions. This is based on previous findings for the effect of

number in monolingual acquisition (Avram & Sevcenco, 2020; Contemori, Carlson & Marinis, 2018). Furthermore, number is also a syntactically active feature in Romanian, as it is manifested in the agreement on the verb. As such, the featural intervention account (Friedmann, Belletti & Rizzi, 2009; Belletti et al., 2012) predicts that a mismatch in number between the moved object and the intervening subject, which is also manifested in the verbal inflection, should facilitate the intervention effects found with object *which*-questions. The presence of DOM, on the other hand, should not eliminate the subject-object asymmetry found with *which*-questions in Romanian. Given that DOM, like morphological case, is not part of the features that act as attractors for movement (Friedmann, Rizzi & Belletti, 2017), the prediction is that children would have difficulties using DOM for overcoming intervention effects in object *which*-questions. Note, however, that the predictions of the intervention account only hold for the final interpretation, so once listeners reach the end of the sentence and are computing its meaning.

For online interpretation, based on the findings of Schouwenaars, Hendriks & Ruigendijk (2018) for German, the prediction was that the presence of DOM on the fronted *which*-object in Romanian will guide children and adults' interpretation, such that they would start with an object-first interpretation for object questions. In other words, after hearing the *which*-object phrase, participants would direct their gaze to the target image corresponding to an object interpretation. Moreover, if children and adults are sensitive to the presence of number mismatch between the fronted wh-word and the verb in an object *which*-question and integrate this into the online interpretation (similar to the results reported by Contemori Carlson & Marinis, 2018), then there should be an increase in looks to the target image corresponding to an object interpretation after the verb.

4. Method

4.1 Participants

The study included thirty typically-developing children with no language, hearing, or speech pathologies (as reported by the parents), aged between 6 and 10 (21 male, 6;4–10;4, M = 7;11 years, SD = 13.9 months). Thirty-two adults were tested as control group (9 male, M = 24 years old, SD = 44.6 months). Children were recruited in northern Romania either through schools, or via e-mail or word-of-mouth, while adults were recruited at various Romanian universities through e-mail, word-of-mouth and social media posts. The study was approved by the Ethical Committee of the

University of Konstanz, Germany, and in accordance with the declaration of Helsinki. Informed consent was obtained prior to the experiment. In the case of child participants, the parents received information about the study and gave their written consent for their children's participation in the study.

4.2 Method

The participants took part in a picture-selection task with eye tracking (adapted from Contemori, Carlson & Marinis, 2018). The study manipulated two independent variables: Question Type and Number, each with two levels. Participants were thus assessed on four conditions, exemplified in Table 1, and each condition included eight sentences, resulting in thirty-two experimental items. The Question Type variable distinguished between subject and object *which*-questions, while the Number variable distinguished between number match and number mismatch conditions. The questions with number match contained two NPs in the singular and involved the same number agreement between the first NP and the verb (examples (4) and (5)). The questions with number mismatch contained a first NP in the plural and a second NP in the singular and involved the same number agreement between the first NP and the verb (for subject questions, as shown in (6)), or different number agreement

Table 1. Examples of test sentences

Question Type	Number	Test Item
Subject	match	1. Care doctor împinge regele? which doctor pushes king.the.SG.M Which doctor is pushing the king?
	mismatch	2. Care doctori împing regele? which doctors push king.the.SG.M Which doctors are pushing the king?
Object	match	3. Pe care doctor îl împinge regele? DOM which doctor him pushes king.the.SG.M Which doctor is the king pushing?'
	mismatch	4. Pe care doctori îi împinge regele? DOM which doctors them pushes king.the.SG.M Which doctors is the king pushing?

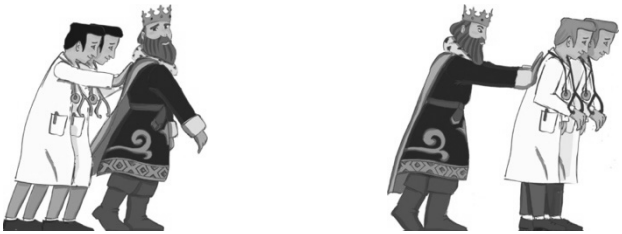
between the first NP and the verb (for object questions, as shown in (7)). Therefore, in the number match conditions, object questions were disambiguated from subject questions through the presence of DOM, marking accusative case, and of the clitic that doubled the wh-object; in the number mismatch conditions, object questions were disambiguated through the presence of DOM, clitic doubling and number agreement on the verb.

Nine transitive verbs were chosen (*a acoperi* 'to cover', *a fugări* 'to chase', *a împinge* 'to push', *a lovi* 'to kick', *a atinge* 'to touch', *a mușca* 'to bite', *a pieptăna* 'to comb', *a spăla* 'to wash', *a trage* 'to pull'). All the verbs denoted semantically reversible actions. Four verbs (*a acoperi* 'to cover', *a mușca* 'to bite', *a pieptăna* 'to comb', *a spăla* 'to wash') were associated with inaudible verb agreement, and five verbs instantiated audible number agreement with the subject (*a fugări* 'to chase', *a împinge* 'to push', *a lovi* 'to kick', *a atinge* 'to touch', *a trage* 'to pull'). Thirty-two filler items were also included. These were subject and object *who*-questions (e.g. *Who is covering the grandfather*?), subject *which*-questions with an inanimate patient (e.g. *Which boy is eating an ice-cream*?) and adjunct questions (e.g. *Where is the laughing* fish?). Each item was preceded by a lead-in introducing the characters (e.g. *Here are some doctors and kings.*) and was aurally presented together with two pictures depicting two animate entities (e.g. either two animals or two people) with one performing an action on the other (e.g., pushing). One of the pictures (the "target") matched the argument structure of the question, while the other (the "competitor") depicted the reverse structure. Figures 1 and 2 illustrate the pictures corresponding to the four experimental conditions.

A native speaker of Romanian recorded all the sentences, as well as the task instructions, in a sound-proof booth. The items were pseudo-randomized using a Latin-square design and divided across four lists, such that each participant saw an item only once (in one of the four conditions). The experimental items were interspersed with fillers, such that no two test items were presented after each other. The position of the pictures and the direction of the actions was also counterbalanced: the correct picture appeared on the left for half of the items, and on the right, for the other half. The action was realized from left to right in half of the pictures and from right to left in the other half.

Figure 1. Example of pictures[1] associated with subject and object questions in the number match condition (examples (4) and (6)).

Figure 2. Example of pictures associated with subject and object questions in the number mismatch condition (examples (5) and (7)).



4.3 Procedure

The study employed a visual world eye-tracking task, which was set up and run using the Gorilla Experiment Builder platform (Anwyl-Irvine et al., 2020; Anwyl-Irvine et al., 2021) Participants were tested individually in a quiet room and were recommended to wear headphones. For the children, each testing session started with a video call in which the experimenter (the author) introduced the task to the children and also explained in detail what they had to do during the task. For adults, the session started with video instructions outlining the purpose and the general procedure of the experiment. Afterwards, participants proceeded to the eye-tracking task, beginning with more detailed audio instructions and the initial webcam calibration. During calibration, they saw five red points, one at a time, and had to fixate their gaze on each. This was followed by a validation phase in which five green points appeared one after the other at different locations on the screen and participants had to fixate their gaze on them. Once the

[1] The pictures were created by Jovana Kačavenda (www.youwannaux.com) as part of the Marie Skłodowska-Curie grant No 101026216.

participants completed the calibration successfully, they began the eye-tracking experiment, which comprised sixty-eight trials: four practice trials, thirty-two experimental trials, and thirty-two fillers. Three additional calibration phases were included – one after the practice trials, one after twenty-four trials, and another after the next twenty trials. Each trial started with a fixation cross, and participants clicked a "NEXT" button below the cross to proceed to the visual display, which consisted of two pictures. A preview of 1000 milliseconds was provided, during which participants listened to the lead-in associated with each sentence. They then listened to a *which*-question, while the two pictures remained on the screen. At the end of the question, a cursor appeared on the screen and participants had to click on the picture that answered the question. They could only click to select an image once the whole audio was played. There was no time limit for providing an answer and there was no feedback given on the accuracy of the response. Children completed the task in about thirty minutes, while adults took about twenty minutes to finish the task.

5. Results

5.1 Analysis

The data were analysed in R (R Core Team, 2022). *Regression* analyses were performed *for the accuracy data* using the lme4 package (version 1.1–33; Bates et al., 2015). The DHARMa package (version 0.4.6; Hartig, 2020) was used to run model diagnostics and the ggplot2 package (Wickham, 2016) for visualizations. Given that the dependent variable (response accuracy) had a binomial distribution, binomial generalized linear mixed-effects models were fit to the data. These model the likelihood of a correct (1) or an incorrect response (0) as a function of a set of predictors. The analysis began by specifying the maximal random-effects structure, which had to be reduced to sustain model convergence. The final models included a random intercept for Participant and one for Item, as well as random slopes for Question Type by Participant. The fixed effects structure initially included all the predictors of interest, namely Group (Adults vs. Children), Question Type (subject vs. object), and Number (match vs. mismatch), as well as the three-way interaction between Group, Question Type, and Number. A repeated-contrast specification was used for all three categorical fixed effects, such that: “Adults” (from the Group predictor), “subject” (from the Question Type predictor), and “Mismatch” (from the Number predictor) were coded as –0.5, and “Children” (from the Group predictor), “object” (from the Question Type predictor), and “Match” (from the

Number predictor) were coded as 0.5. This contrast specification compares the two levels of the factor against each other and the estimates represent the difference between these two levels. The Intercept corresponds to the global grand mean. Each non-significant interaction term or predictor was then eliminated step by step using backwards selection. Numerical predictors were scaled and centred to circumvent issues related to the use of different scales. Model comparisons were performed using likelihood ratio tests. Throughout the text only the results of the final models are reported.

The *eye-movement data* was analysed with *generalised additive mixed models* (GAMMs) using the packages *itsadug* and *gamm4* (*bam, mgcv, compareML*, Wood, 2006, 2011). GAMMs are a type of nonlinear regression analysis, well-suited for anaylzing time-course data obtained from visual world eye-tracking studies (Ito & Knoeferle, 2023). Similar to generalized linear mixed-effects models, GAMMs can take both fixed and random factors into account. Their key advantage, however, is their ability to handle nonlinear datasets by modeling the relationships between predictors and the dependent variable as smooth functions.

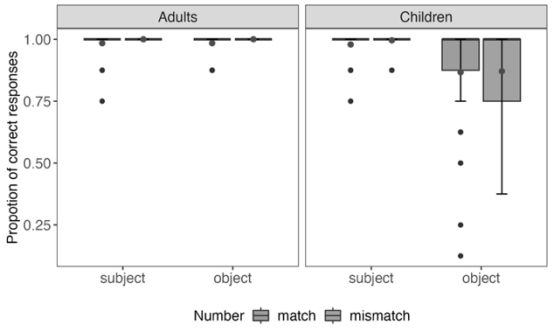
Incorrect trials were removed from the analyses consistent with Contemori, Carlson & Marinis (2018) and standard practice with this type of data. This resulted in the loss of 10 trials for subject questions (around 0.01% of the total subject question trials) and 67 trials for object questions (around 7% of the total object question trials) due to lower comprehension accuracy in the latter. From the eye-tracking record, gaze position was determined in 100ms bins and fixation was coded binomially (1 = fixated, 0 = not fixated) for each time bin (Ito & Knoeferle, 2023). Eye-movements were analysed during a 3000 ms window from the onset of the wh-word until the end of the sentence. Two models were fit for each group, children and adults. The baseline model included time as smooth term and Participant as random effect. The second model also included Question Type and Number, as well as their interaction, as a categorical variable transformed into an ordered factor. The two fixed factors were coded using treatment contrast coding. A model comparison was run with the *compareML* function to assess the fit and differences between the two generalized additive models. The model comparison results indicated a significant improvement in fit for the full model compared to the null model ($p < .001$). The packages *ggplot2* (Wickham, 2016) and *tidymv* (*plot_smooths* and *plot_difference* functions, Coretta, 2020) were employed for data visualization. The difference plots obtained from the *plot_difference* function were used to compute differences between looking behaviour for subject and object questions with or without a mismatch in number in children and adults.

The results for comprehension accuracy are presented first and then those for the eye-gaze data.

5.2 Accuracy results

To answer RQ1, the comprehension of *which*-questions in children and adults was compared in order to determine the impact that the type of question (subject vs. object) and the presence or absence of a mismatch in number have on offline comprehension in both groups. Figure 3 illustrates the descriptive results for children and adults' response accuracy for subject and object *which*-questions matching or mismatching in Number. The results for subject questions show that both adults and children comprehended these very well, displaying ceiling or near ceiling performance (M_{adults} = 0.992, SD_{adults} = 0.088; $M_{children}$ = 0.987, $SD_{children}$ = 0.111). The results for object questions indicated a higher degree of variability in children, but not in adults, who reached ceiling performance for object questions as well (M_{adults} = 0.992, SD_{adults} = 0.088; $M_{children}$ = 0.868, $SD_{children}$ = 0.338).

Figure 3. Response accuracy as proportion score (0–1) by Question type (subject and object) and by Number (matching and mismatching) for the two groups. The red dots represent the mean for each condition.

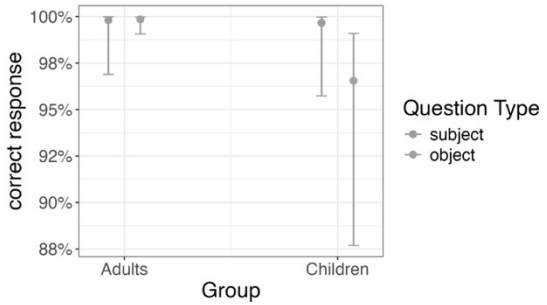


The lower black points in the adult group represent outlier responses values. Further inspection of the data showed that one adult participant gave two incorrect responses for subject questions with a match in number (75% accuracy), while six others incorrectly interpreted one item in one of the match conditions. Importantly, the lower accuracy scores for subject and object questions in the number match condition do not come from the same person and all participants performed above chance in the picture selection

task. Figure 3 also points to the presence of outlier values in the child group. These are mainly present in the object match condition, which also has a smaller spread than the mismatch condition. The outlier values indicate that some children performed much worse than the majority with object questions containing a match in number. A closer look at the data reveals that these children are among the younger ones (aged 6 to 7) in the group. Two children had below-chance performance in the object match condition and one of them also displayed below-chance performance in the object mismatch condition. Three other children performed at chance with both types of object questions.

The final model investigating offline comprehension of *which*-questions included as fixed predictors Group (Adults vs. Children), Question Type (subject vs. object), Number (match vs. mismatch) as well as the interaction between Group and Question Type. Only the significant effects are reported in the text. The model showed that, overall, children were less accurate than adults (*est.* = -1.863, *SE* = 0.656, *z* =- 2.840, *p* = .005). There was, however, a significant interaction between Group and Question Type (*est.* = -2.598, *SE* = 1.317, *z* = -1.972, *p* = .049), indicating that child group had more pronounced comprehension difficulties with object questions but not with subject questions compared to the adult group (Figure 4).

Figure 4. Predicted values given the interaction between Group (Adults vs. Children) and Question Type (subject vs. object) in the model predicting performance with *which*-questions according to group of participants, type of question, match or mismatch in number features, and the interaction between question type and number.



Given the wide age range of the child group, the effect of age on the comprehension of subject and object questions with or without a mismatch in number was probed further by focusing only on the child data. The final model included Question Type, Number and Age in months as fixed predictors. The analysis only revealed a main effect of age (*est.* = 1.016, *SE*