

ARTICLE

Comprehension of complex sentences containing temporal connectives: How children are led down the event-semantic kindergarten-path

Christos Makrodimitris  and Petra Schulz 

Goethe University Frankfurt

Corresponding author: Petra Schulz; Email: P.Schulz@em.uni-frankfurt.de

(Received 16 June 2023; revised 26 February 2024; accepted 18 April 2024)

Abstract

Children up to school age are known to have difficulty comprehending complex sentences with temporal connectives, but the reasons remain controversial. We tested six- to twelve-year-old children to assess how the iconicity of event-language mapping, type of connective, and clause order mediate the comprehension of temporal sentences. Sixty monolingual Greek-speaking children and 15 adult controls completed a picture-sequence selection task in which they judged *after*- and *before*-sentences in iconic and non-iconic order. Up to age twelve, children did not reach full adult-like comprehension of the connectives; performance in non-iconic *after*-sentences was significantly lower than in the other three conditions across all ages. We conclude that neither iconicity, connective, nor clause order can fully explain these findings and propose an account based on the interaction of iconicity and clause order: non-iconic, sentence-medial *after* requires revision of the initial event representation, resulting in an event-semantic kindergarten-path that children find difficult to overcome.

Keywords: adverbial clauses; clause order; iconicity; temporal connectives; kindergarten-path

Introduction

Imagine a colleague tells you “Alex won the lottery. Her husband quit his job.” Although the temporal order of the two events is not made explicit, you will most likely assume that first Alex won the lottery and then her husband quit his job. This can be traced back to the discourse principle of chronological order stating that, unless marked otherwise, the order in which events are mentioned mirrors the order in which they happened (e.g., Klein, 2009). This discourse principle, often referred to as the iconicity of sequence principle (Greenberg, 1966; Haspelmath, 2008, a.o.), basically says that in the absence of other cues, listeners and readers will assume that what they hear first, happens first. Iconicity is defined as a relationship of resemblance or similarity between a formal property of a sign and a property of its referent (Haiman, 2006). It is typically concerned with the nature of the relationship between the form and the meaning of words. Extending its scope to

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

syntactic structures, iconicity is assumed to also govern speakers' choices of structurally available options in discourse (see the discussion in Newmeyer, 1992). Accordingly, listeners expect the order in which events are mentioned to match the order in which they occur, which leads them to interpret sentence sequences in chronological order.¹ In the following, we use the term "iconicity" to refer to iconicity of sequence.

The iconic order may be marked explicitly via coordinating connectives (e.g., *Alex won the lottery, then her husband quit his job.*) or via subordinating connectives such as *after* and *before* (e.g., *After Alex won the lottery, her husband quit his job.* or *Alex won the lottery, before her husband quit his job.*). Importantly, language also provides means to sidestep the principle of iconicity: specific tense marking, for example, can indicate that the event in a subsequent sentence actually happened before the event in the preceding sentence (e.g., *Alex's husband quit his job. Alex had won the lottery.*, see Hamann et al., 2001). Moreover, a non-iconic order can be expressed through coordinating connectives (e.g., *Alex's husband quit his job, but first Alex won the lottery.*) and via the subordinating temporal connectives *before* and *after* (e.g., *Alex's husband quit his job, after Alex won the lottery. Before Alex's husband quit his job, Alex won the lottery.*).

Acquisition of temporal connectives represents a remarkable example of asymmetry between production and comprehension. Complex sentences with temporal connectives emerge early in child speech, between the ages of two and three (e.g., Bloom et al., 1980; Clark, 2009; Diessel, 2004 for English; Rothweiler, 1993 for German; Baslis, 1994; Stephany, 1997 for Greek). But early production of temporal connectives does not indicate an adult-like understanding of their semantics: many of the early uses of temporal connectives occur in descriptions of common routines (e.g., a child describing that she puts on her socks before her shoes; Clark, 2009). In such descriptions, the temporal relations between the events are known to children through experience and use of the connective does not require knowledge of its temporal semantics (see also Tillman et al., 2017; Zhang & Hudson, 2018). In fact, starting with the seminal study by Clark (1969, 1971), a large body of research has provided strong evidence that children do not

¹The iconicity principle seems to be closely related to Grice's (1975) maxim "Be orderly". This relation was already suggested in the seventies (for a discussion, see Newmeyer, 1992); we are grateful to the anonymous reviewer who pointed us toward this parallel. In Gricean terminology, the maxim 'be orderly' can be characterized as a conversational implicature (i.e., the events occur in the order in which the clauses are uttered), which can be cancelled by the speaker. For example, (i) conversationally implicates, but does not entail, that Alex won the lottery before her husband quit his job, as can be seen by the felicitous continuation *but not in that order* (Newmeyer, 1992, p. 759).

- (i) Alex won the lottery and her husband quit his job... - but not in that order.
- (ii) Alex won the lottery. Her husband quit his job... #? but not in that order.

Note, however, that in our mini-discourse, repeated in (ii), where the two sentences simply follow each other, the temporal inference seems to be much harder to cancel. This is rather unexpected given that the manner maxim 'be orderly' should apply in the same way to these types of discourse. We leave this issue for further research, and in the remainder of the paper, we follow the terminology used in previous acquisition studies and refer to this principle as the iconicity principle.

We also wish to note that both the manner maxim 'Be orderly' and the iconicity principle do not address the difference between events and states, see (iii), taken from Partee (1984):

- (iii) a. John got up. b. He raised the blind. c. It was light out. d. He was not ready to face the day.

While the event sentences (a, b) move the action forward in time, the state sentences (c, d) do not, "they describe how things are at the last-mentioned event" (Partee, 1984, p. 253). In previous acquisition studies as well as in our on study, only event sentences were used; therefore, we do not consider this issue further.

reach full target-like understanding of sentences with temporal connectives up to late childhood.

Since the 70s, researchers have proposed different explanations to account for the results (see de Ruiter et al., 2018a, 2021), but to date, there is no consensus on which linguistic properties of the sentences cause the most difficulty for the language learner. This may be because studies often tested children of different ages and with different methods, making a close comparison across ages and experimental tasks difficult. Moreover, the majority of comprehension studies did not include children over the age of eight, leaving open how comprehension of the connectives develops in older children.

To address these open issues, in this study we investigate the comprehension of sentences containing the temporal connectives *before* and *after* in children aged six to twelve using a carefully designed forced-choice picture-sequence selection task. By testing children across a wide age range, we aim to uncover the possibly different interpretation strategies across age. We focus on how the semantic and syntactic properties of *before*- and *after*-sentences influence children's interpretation of temporal order. Children's responses are analyzed both at the group and at the individual level to discover potential individual differences that may go unnoticed in the group analysis. In addition, we examine to what extent children's performance is influenced by their short-term memory capacity and general language ability.

The remainder of the paper is structured as follows: we first summarize the findings of previous studies with respect to the main linguistic factors argued to affect children's comprehension of *before*- and *after*-sentences. We then provide information about the participants, materials, and methods of the current study and present our findings. To conclude, we suggest a novel account drawing on event-semantic representations to explain our findings. We propose that children's difficulty with non-iconic *after*-sentences is similar to the 'kindergarten-path effect' documented for children's processing of temporary syntactic ambiguities *sensu* Trueswell et al. (1999). Therefore, we refer to our account as the 'event-semantic kindergarten-path'.

Child comprehension of sentences containing temporal connectives

Complex sentences containing temporal connectives (henceforth also temporal sentences) comprise different semantic and syntactic properties. Examples (1) and (2) (taken from Clark's initial study, 1969, 1971), illustrate test sentences typically used in the experiments (*main* and *sub* are shorthand for main clause and subordinate clause).

- (1) a. The girl jumped the gate before she patted the horse. [before, iconic, main–sub]
 b. Before the girl patted the horse, she jumped the gate. [before, non–iconic, sub–main]
- (2) a. The girl patted the horse after she jumped the gate. [after, non–iconic, main–sub]
 b. After the girl jumped the gate, she patted the horse. [after, iconic, sub–main]

Numerous studies across several languages have provided evidence for how the semantic and syntactic properties of temporal sentences affect children's comprehension. Before

summarizing the findings of previous studies with respect to the main linguistic factors investigated, we would like to highlight that different findings may result from the fact that the studies tested children of different ages and, more importantly, from the different methods used. Regarding the age range of the participants, research focused on preschool and young primary school children – that is, up to age five or seven – while studies on children aged eight and older are still scarce (for exceptions, see Overweg et al., 2018; Papakonstantinou, 2015; Wagner & Holt, 2023; for reading: Pyykkönen & Järvikivi, 2012; Karlsson et al., 2019). Notably, none of the studies with older children found fully target-like comprehension of the connectives; even twelve-year-olds had difficulties with one or more of the conditions presented in (1) and (2).

As for the methods used, studies conducted in the 1970s and 1980s often used act-out tasks where children acted out temporal sentences using toys and props (English: Amidon, 1976; Amidon & Carey, 1972; Coker, 1978; Feagans, 1980; French & Brown, 1977; Gorrell et al., 1989; Hatch, 1971; Johnson, 1975; Kavanaugh, 1979; Keller-Cohen, 1987; Richards & Hawpe, 1981; Stevenson & Pollitt, 1987; Danish: Trosborg, 1982; Greek: Natsopoulos & Abadzi, 1986). The wealth of recent studies used variants of the picture-selection method. In one variant, the ‘What happened first/last?’ task, children listened to sentences with *before* and *after* and had to choose the picture of the event that happened first or last (English: Blything & Cain, 2016; Blything et al., 2015; Wagner & Holt, 2023; Dutch: Overweg et al., 2018; Greek: Tsakali & Vamvouka, 2021). A similar design was used in two studies that examined comprehension of the connectives in written language (Finnish: Pyykkönen & Järvikivi, 2012; Dutch: Karlsson et al., 2019). In another variant, the so-called picture-sequence selection task, children had to choose the correct picture sequence out of two sequences, which depicted the actions in the target and in the reversed order (English: de Ruiter et al., 2018a; German: de Ruiter et al., 2018b; Tamil: de Ruiter et al., 2019; Mandarin Chinese: de Ruiter et al., 2020a). We consider the latter picture-sequence selection task to be the most reliable for assessing children’s comprehension of temporal sentences. Act-out tasks, in contrast, provide children with a high degree of freedom in their responses, which makes it difficult to interpret responses regarding the order of the events. For example, children were often found to act out only one of the two clauses (e.g., Amidon & Carey, 1972; Gorrell et al., 1989), which does not inform us about the event order. Finally, the ‘What happened first/last?’ task constrains the range of possible responses to just two, but it is subject to recency effects: children may opt for the event that was more recently activated in their memory, i.e., the event that was mentioned last (Blything & Cain, 2016; Karlsson et al., 2019). The picture-sequence selection task constrains the child’s responses to two and is less likely to cause recency effects, because the picture of the event that happened last appears in both picture-sequences.²

Taking into account the different age groups tested and the methods used, we review below the existing studies on the role of linguistic factors that have been argued to affect children’s comprehension of temporal sentences: iconicity, semantics of the connectives, clause order, and the interaction between iconicity and clause order.

²The order in which the visual and the auditory stimuli are presented has sometimes been argued to play a role for children’s response patterns. In ‘What happened first/last?’ tasks, children saw the pictures before they listened to the sentences, while in picture-sequence selection tasks, children generally listened to the sentences before they saw the pictures.

Iconicity

Clark (1969, 1971) tested English-speaking children (aged 3;0–5;0) on *before* and *after* in an act-out task: children were asked to perform the events described by sentences such as (1) and (2), using toys. In the iconic sentences (1a) and (2b), the linear order of the clauses follows the chronological order of the events, whereas in the non-iconic sentences (1b) and (2a), the clause order is incongruent with the chronological order. Performance gradually improved with age; accuracy was higher in iconic than in non-iconic sentences (and higher in sentences with *before*(1a,b) than in sentences with *after*(2a,b), see Section *Semantics of the connectives*). This pattern was also present in the individual data: most three-year-olds and some four-year-olds responded correctly to iconic and incorrectly to non-iconic sentences. Clark called this pattern order-of-mention strategy and argued that children employing this strategy treat the clauses of the complex sentence as independent.

Subsequent act-out studies provided mixed evidence for iconicity, i.e., higher accuracy for iconic than for non-iconic sentences. Many studies found positive effects of iconicity (Coker, 1978; Feagans, 1980; French & Brown, 1977; Hatch, 1971; Johnson, 1975; Natsopoulos & Abadzi, 1986; Richards & Hawpe, 1981; Trosborg, 1982), while others did not confirm the advantage of iconic over non-iconic sentences (Amidon, 1976; Amidon & Carey, 1972; Gorrell et al., 1989; Stevenson & Pollitt, 1987) or even found better comprehension of non-iconic sentences in the case of *after* (Papakonstantinou, 2015). The majority of ‘What happened first/last?’ studies found either no effects of iconicity (Karlsson et al., 2019; Overweg et al., 2018; Tsakali & Vamvouka, 2021) or partial effects (for *before*: Blything et al., 2015, for *after*: Pyykkönen & Järvikivi, 2012; Wagner & Holt, 2023). Only one study reported a general advantage of iconic over non-iconic sentences (Blything & Cain, 2016).

Finally, the studies using picture-sequence selection found that in English and German, languages with clause-initial adverbial connectives, iconic sentences are understood better than non-iconic ones (de Ruiter et al., 2018a, 2018b). Two pilot studies on Tamil and Mandarin Chinese, languages with adverbial connectives in clause-final position (expressed via a bound morpheme in Tamil and via a free morpheme in Mandarin Chinese), found a mixed pattern: iconic sentences were easier than non-iconic ones in the case of *after*, but non-iconic sentences were easier than iconic ones in the case of *before* (de Ruiter et al., 2019, 2020a). However, the results of these two pilot studies are difficult to interpret because it is not clear whether the order main-subordinate is as acceptable as the order subordinate-main in Tamil (Shanmugasundaram Rajamathangi, p.c.) and in Mandarin Chinese (Le’an Luo, p.c.).³

In short, evidence for iconicity effects seems to be present to some extent, though not very robustly, across different tasks, including picture-sequence selection. Considering the age of the children tested, while iconicity may play some role in comprehension in younger children, no general iconicity effects have been found in studies with children over the age of eight, regardless of the task used (Karlsson et al., 2019; Overweg et al., 2018; Papakonstantinou, 2015; Pyykkönen & Järvikivi, 2012; Wagner & Holt, 2023).

³The acceptability of the main-subordinate order in Tamil and Mandarin Chinese is unclear; accordingly, de Ruiter and colleagues’ (2019, 2020a) results for these languages are difficult to interpret in terms of the factors clause order and semantics of the connective. For this reason, we do not report further on these findings.

Semantics of the connectives

Another important finding of Clark (1969, 1971) was that performance at the group level was better for *before*-sentences than for *after*-sentences. The individual data showed that some children responded correctly to *before*- and incorrectly to *after*-sentences, both in their iconic and their non-iconic version. According to Clark, children exhibiting this pattern treated *after* as if it meant *before*. In addition, some children, mainly four-year-olds, responded correctly to iconic and non-iconic sentences with *before* and to iconic sentences with *after*. To account for the developmental advantage of *before* suggested by her data, Clark (1971) proposed the so-called ‘semantic components hypothesis’. According to this hypothesis, the meaning of *before* and *after* is represented as a set of three features. *Before* is specified as +Time, -Simultaneous, +Prior (‘at a time preceding the time as which...’), whereas *after* as +Time, -Simultaneous, -Prior (‘at a time following the time at which...’). The positive value of each feature is assumed to be acquired prior to the negative value, resulting in earlier mastery of *before*.

Subsequent act-out studies provided varied evidence for the semantic components hypothesis: some confirmed Clark’s finding that *before* was understood better than *after* (Coker, 1978; Feagans, 1980; Hatch, 1971; Papakonstantinou, 2015; Richards & Hawpe, 1981; Trosborg, 1982), while others found an advantage of *before* only in the iconic condition (Johnson, 1975) or reported no differences between the connectives (Amidon, 1976; Amidon & Carey, 1972; French & Brown, 1977; Gorrell et al., 1989; Kavanaugh, 1979; Natsopoulos & Abadzi, 1986). Studies using the ‘What happened first/last?’ task provided an even more diverse picture. Two studies found a general advantage of *before* (Overweg et al., 2018; Wagner & Holt, 2023), two studies found a partial advantage of *before*, in the iconic (Blything et al., 2015) or in the non-iconic condition (Tsakali & Vamvouka, 2021), and one study found no difference between the connectives (Blything & Cain, 2016). Notably, the findings of a recent reading study with nine- to twelve-year-old Dutch-speaking children (Karlsson et al., 2019) suggest that difficulty of the connective may depend on the specific instruction: when asked to choose the picture of the event that happened first, children performed better with *before*-sentences; but when asked to choose the picture of the event that happened last, children performed better with *after*-sentences. The studies using picture-sequence selection tasks found that *before* was easier than *after* for five-year-old children in English and German (de Ruiter et al., 2018a, 2018b).

In summary, Clark’s (1971) semantic components hypothesis has been challenged by a number of studies. The studies based on picture-sequence selection tasks, which we consider to be the most reliable, have provided some evidence in favor of this for languages with clause-initial connectives, at least for five-year-olds.

Clause order

Diessel (2004, 2005, 2008) proposed that from a processing perspective, complex sentences should be easier to understand if they occur in main-subordinate (1a, 2a) than in subordinate-main order (1b, 2b). In the order main-subordinate, the connective establishes the link between the two clauses, after the main clause has been heard, i.e., the biclausal sentence can be parsed stepwise. The connective in the order subordinate-main appears at the beginning of the complex sentence and the subordinate clause has to be kept in memory until the main clause has been parsed; only then can the link between the two clauses be formed. Note that this formulation of the account assumes that the

language has clause-initial adverbial connectives, which is the case for many of the languages investigated in previous studies as well as for the language of the current study, Greek. Importantly, in languages with clause-initial connectives, the predictions of clause order and iconicity differ only regarding *after*-sentences: (2a) should be easier than (2b) according to the clause order account, but according to the iconicity account, (2a) should be more difficult than (2b). In *before*-sentences, the easier order main-subordinate coincides with the easier iconic order. To date, the majority of research on languages with clause-initial connectives has not found general effects of clause order on children's comprehension, regardless of the method used: iconic *after*-sentences were understood either as well as or better than their non-iconic variants (e.g., Clark, 1971; Feagans, 1980; Blything et al., 2015; Overweg et al., 2018, de Ruiter et al., 2018a, 2018b, see also Section *Iconicity*). The study by Papakonstantinou on Greek (2015) indicated that *after*-sentences in the order main-subordinate were easier than in the order subordinate-main, but no differences were found for *before*-sentences, speaking against a general effect of clause order (or of iconicity).

In summary, the clause order account of Diessel (2004, 2005, 2008) has not received clear empirical support from previous studies on languages with clause-initial connectives, including the studies that used picture-sequence selection tasks.

Interaction of iconicity and clause order

Pyykkönen and Järvikivi (2012) argue that effects of iconicity in concert with clause order mediate children's comprehension. Although their study targeted reading comprehension, it is reviewed here in some detail, as our event-semantic kindergarten-path account – conceived independently – is similar in spirit to their explanation. The authors conducted a paper-and-pencil reading experiment with eight-, ten-, and twelve-year-old Finnish-speaking children, in which the participants had to read sentences with *before*- and *after*-clauses. A typical item is given in (3) in English translation:

- (3) Ilkka read the letter before he went to school.
 READ WENT AT THE SAME TIME

The participants' task was to indicate which of the events occurred earlier or whether the events occurred at the same time, by marking one of the three options (the option 'at the same time' was the correct choice for another test condition⁴). Children's accuracy did not differ for iconic and non-iconic sentences in the order subordinate-main (see (1b) and (2b)), but in the order main-subordinate, iconic sentences were easier than non-iconic ones (see (1a) and (2a)), even in the group of twelve-year-olds. According to Pyykkönen and Järvikivi (2012), this pattern suggests that effects of iconicity are mediated by the position of the connective: when the connective appears sentence-initially, children can construct the correct situation model of the sentence, even if iconicity is violated. In contrast, when the connective appears sentence-medially, a non-iconic description is challenging, because it requires children to revise the situation model they are building, as soon as they encounter the connective. It is open to what extent this written task is comparable to oral tasks. By providing the test items in a written form without time

⁴Pyykkönen and Järvikivi (2012) also tested temporal converb constructions, which are not relevant for the present study.

constraints, participants could go back to their answers and – in principle – revise answers after noticing certain patterns in the test items. On the other hand, children had to translate their judgment into one of the three ‘verbal’ answer choices, which may have added a complex metalinguistic aspect which is not present in oral tasks.

Other factors

In addition to the large body of research reviewed above that has examined the role of linguistic properties of temporal sentences, some studies have examined the role of factors that go beyond these linguistic properties. Here we summarize three factors relevant to the current study: world knowledge, cognitive abilities, and language abilities.

Addressing the role of world knowledge, some studies compared sentences in which the event order was arbitrary with sentences in which the event order was constrained by world knowledge (e.g., *Ann fills the bottle before she feeds the baby.*, see French & Brown, 1977). Arbitrary sentences were reported to be more difficult in early act-out studies (French & Brown, 1977; Kavanaugh, 1979; Keller-Cohen, 1987; Natsopoulos & Abadzi, 1986; Trosborg, 1982). However, the two ‘What happened first/last?’ studies to include world knowledge in their experimental design found either no evidence for this asymmetry (Blything et al., 2015) or only for *after* (Wagner & Holt, 2023).

More recently, the role of cognitive abilities and of language abilities for children’s comprehension of temporal sentences has been investigated using picture-selection tasks. Some studies found memory-related measures to be significant predictors of children’s performance (forward digit recall: Blything & Cain, 2016; Blything et al., 2015; word sequences and non-words: de Ruiter et al., 2018b; sentence span, updating: Karlsson et al., 2019). Two of these studies in fact reported that short-term memory was a stronger predictor of children’s performance than age (Blything & Cain, 2016; Blything et al., 2015). Another study (Overweg et al., 2018) reported a positive effect of receptive vocabulary ability on children’s comprehension of *before*- and *after*-sentences. Other studies (de Ruiter et al., 2018a, 2020a), however, found that neither cognitive nor language measures influenced their results. In short, there is some, albeit weak, evidence that event sequences constrained by world knowledge are better understood than arbitrary sequences, and that cognitive abilities as well as language abilities may mediate children’s comprehension of temporal sentences.

Summary

Previous research agrees that comprehension of sentences involving the connectives *before* and *after* may remain challenging even in late childhood. Assuming that mastery of a phenomenon involves mastery of all its aspects (e.g., Schulz, 2003), temporal sentences can be characterized as very late acquired (see Schulz & Grimm, 2019; Tsimpli, 2014 for a general discussion of late phenomena). Using a range of different tasks in different languages, existing studies have explored different accounts, emphasizing the factors iconicity, semantics of the connectives, and/or clause order. We have argued that the results from picture-sequence selection tasks are particularly reliable when evaluating children’s comprehension of temporal sentences. The two studies using picture-sequence selection (de Ruiter et al., 2018a, 2018b) suggest a role for iconicity and for the semantics of the connective, but not for clause order, in five-year-olds. The few studies of children aged eight or older used either ‘What happened first/last?’ tasks (Karlsson et al., 2019;

Overweg et al., 2018; Pyykkönen & Järvikivi, 2012; Wagner & Holt, 2023) or act-out tasks (Papakonstantinou, 2015). These studies found that clause order had no effect and that iconicity and the semantics of the connective played only a minor role for children's performance, suggesting that children's comprehension strategies may change with age. However, because the studies with older children used less reliable methods than the studies with younger children, it is not clear how comparable results are. Thus, the question of whether iconicity, semantics of the connectives, and/or clause order play different roles in younger and older children's comprehension of sentences with temporal connectives remains open.

The present study

To address these open issues, the present study investigates the comprehension of sentences containing the temporal connectives *prin* ('before') and *afu* ('after') in children between the ages of six and twelve years who acquire Greek as their native language. Greek is similar to many of the languages studied so far in that it uses clause-initial adverbial connectives and it allows the temporal clause to either precede or follow the main clause. Accordingly, the results of our study are directly comparable with previous results from English, German, and Finnish.

To evaluate all factors reviewed above, we adopted a design that crosses the factors Connective and Iconicity. Clause order (as well as the interaction of iconicity and clause order) is also manipulated as a result of presenting *before* or *after* in iconic or non-iconic order, as shown in examples (1) and (2) in Section *Child comprehension of sentences containing temporal connectives*. Specifically, the study addresses the following research questions:

- Q1: How does comprehension of sentences with *after* and *before* develop between the ages of six and twelve?
- Q2: How do iconicity, semantics of the connective, and clause order influence children's comprehension?
- Q3: To what degree do verbal short-term memory and general language ability modulate children's comprehension?

Group-level and individual response patterns were calculated to answer the first two questions. As for (Q1), we expected children's performance to be non-adult-like at least until age eight. Regarding (Q2), the specific accounts make different predictions: under the iconicity account, children's performance on iconic sentences (1a, 2b) should be easier than on non-iconic sentences (1b, 2a). If only the semantics of the connective is relevant, *before*-sentences (1a,b) should be easier than *after*-sentences (2a,b). Under the clause order account, main-subordinate sentences (1a, 2a) should be easier than sentences in the order subordinate-main (1b, 2b). Finally, if the interaction of iconicity and clause order influences children's comprehension, i.e., if non-iconic sentences cause more difficulty than iconic sentences only in the order main-subordinate, non-iconic *after* (2a) should be more difficult than the other sentence types (1a,b, 2b). Given our discussion of the previous findings for languages with clause-initial connectives, effects of iconicity in children younger than age eight could be expected as well as an advantage of *before* over *after* in all ages. Importantly, clause order should not affect children's performance. To the extent that the findings from the reading study by Pyykkönen and Järvikivi (2012) are

- b. Prin vali to piato sto trapezi,
 before put–NPST.PFV.3SG the plate on–the table
 eklise to parathiro.
 closed–PST.PFV.3SG the window
 ‘Before she put the plate on the table, she closed the window.’
- (5) a. Eklise to parathiro, afu evale
 closed–PST.PFV.3SG the window after put–PST.PFV.3SG
 to piato sto trapezi.
 the plate on–the table
 ‘She closed the window, after she put the plate on the table.’
- b. Afu evale to piato sto trapezi,
 after put–PST.PFV.3SG the plate on–the table
 eklise to parathiro.
 closed–PST.PFV.3SG the window
 ‘After she put the plate on the table, she closed the window.’

The factor clause order was not manipulated directly, but its two levels (main-subordinate, subordinate-main) varied systematically as a result of crossing the factors Iconicity and Connective, allowing us to address the role of clause order. In half of the *before*- and half of the *after*-sentences, the temporal clause preceded the main clause ((4b), (5b)), and in the other half the temporal clause followed the main clause ((4a), (5a)).

Each sentence described a different pair of events, which could not be ordered based on world knowledge, i.e., the two events were logically unrelated, rendering both orders of occurrence equally plausible. The use of logically unrelated events also excluded an ambiguous temporal-causal reading of *afu*-sentences (see Tsimpli et al., 2010), which may affect children’s comprehension negatively. Note that *afu* is only used as a connective, whereas *prin* also functions as a temporal adverb and as a temporal or spatial preposition (Holton et al., 1999), similar to English *before*.

The make-up of our test sentences differed from the one used by de Ruiter and colleagues (de Ruiter et al., 2018a, 2018b, 2019, 2020a) in two respects. Whereas in their task both telic and atelic predicates were used in present tense, our test sentences all contained telic predicates in the past tense. Telic predicates were used to avoid ambiguity of the temporal relation between the events described in the main and the subordinate clause. The verbs were either inherently telic (e.g., *to close*) or formed telic predicates in combination with a quantized object NP (e.g., *to eat an apple*); accordingly, we excluded readings where the first event continues, while the second event starts (see Makrodimitris, 2024; Rett, 2020). Telic predicates are mastered in comprehension between the ages of three and five (for recent overviews, see Schulz, 2018; van Hout, 2018). Moreover, all test sentences exhibited past reference to emphasize completion of the events. The perfective past form (aorist) was used in main clauses and in *after*-clauses. *Before*-clauses obligatorily appeared in the so-called dependent form (Holton et al., 1999). This form is specified for perfective aspect, but does not have an independent time reference; hence, the time reference of the sentence is defined by the tense of the main clause. Given that Greek is a pro-drop language, the subject pronouns were omitted in both clauses, establishing identity of the agent.

To control for possible item effects, two lists were created: sentences that contained *after* in list A contained *before* in list B and *vice versa*. Half of the participants from each

of the six school grades were randomly assigned to one of the two lists. The task contained eight control sentences, which served to ensure that children paid attention to the task. These sentences consisted of a single clause (e.g., *He put on his jacket.*) and had the same grammatical features as the main clauses of the test sentences. Overall, the picture-sequence selection task comprised 32 sentences, divided in four blocks of six test items and two control items each. The order of the test and control items within each block was pseudorandomized (for the complete list of sentences in list A and B, see [Appendix](#)). A female native speaker of Greek recorded the test and control sentences; a pause of 0.5s was inserted between the end of the first and the onset of the second clause in all test sentences.

Short-term memory

The forward digit-recall task from the Greek WISC-III (Georgas et al., 1997) was used to assess children's verbal short-term memory. The task includes nine levels of increasing length, from two to ten digits. Each level consists of two trials. Every correctly repeated trial receives 1 point, and incorrect repetitions receive 0 points. The child has to repeat at least one of the two trials correctly to proceed to the next level. The task stops, if the child fails in both trials of the same level. In the present study, the digit sequences were prerecorded and were presented to children over headphones. The total score was calculated by summing up the points of all sequences up to the highest level that the child had reached. Accordingly, the score ranged from 0 to 18 points.

Language ability

The Greek LITMUS sentence repetition task (Chondrogianni et al., 2013) was used to assess children's general language ability. The task consists of 32 sentences targeting eight different syntactic structures: SVO, negation, coordinated clauses, clitic doubling/clitic left dislocation, complement clauses, adverbial clauses, relative clauses, and object *wh*-questions. Sentences are matched for length and word frequency; they are prerecorded and presented auditorily to the children via a PowerPoint presentation.

For the present study, children's responses were recorded and transcribed independently by two native speakers of Greek, the first author and a research assistant familiar with transcribing child data; disagreements were resolved via discussion. The final transcripts were scored for accuracy of repetition: each sentence received a score of 3 points, if it was repeated verbatim, a score of 2 points, if there was one error, a score of 1 point, if there were two errors, and a score of 0 points, if there were three or more errors. Accordingly, the score ranged from 0 to 96 points (see Prentza et al., 2022, for details on the task and its validation with monolingual Greek children).

Procedure

Children were tested individually in a quiet room at their school. During the familiarization phase, children first saw all pictures used in the main experiment to ensure that they were familiar with the actions and the predicates used later. The pictures were presented on separate cards, and the child was asked to name the action depicted on each card. The actions shown in the pictures were performed by four different child characters, two boys and two girls. If the child did not recognize the depicted action, the experimenter

provided the description and asked the child again to name the action at the end of the familiarization phase. No child had difficulty naming the actions.

The familiarization phase was followed by a practice phase in which children were introduced to the experimental set-up. Five practice items were presented on a laptop, and the child listened to the prerecorded sentences over headphones. The first two items had the same structure as the control sentences of the main test phase (e.g., *He put on his coat.*). The purpose of these items was to familiarize the child with the procedure and the layout of the slides: first, the child listened to a sentence, while the screen depicted an empty page of a picture book. After a pause of 1.5s, two pictures appeared on the page (see Fig. 1). After 1s, the sentence was repeated and the child had to indicate the picture matching the sentence by naming and/or pointing at the respective number. By presenting the test sentence before the pictures/picture-sequences, we wanted to prevent the child from creating a mental representation of the visual stimuli that would possibly influence the processing of the test sentence. No child exhibited difficulties with the first two test items.

Another three practice items consisted of juxtaposed main clauses with temporal adverbs in clause-initial position (*prota* 'first' and *meta* 'then', see example (6)). The layout was the same as in the test sentences, i.e., the child had to choose one of two picture sequences (see Fig. 2). These practice items served to establish the left to right 'reading' direction of the pictures; this is crucial as the two picture sequences only differed regarding the order of the pictures. If a child followed a different 'reading' direction, the experimenter asked: "Look, did things happen in this way or in this way?", while moving her finger over each picture sequence from left to right. This feedback was repeated until the child gave the correct answer. If necessary, the sentence was played a second time. Four children initially followed the wrong reading direction in the first of these three items; after receiving feedback they chose the correct picture sequence and responded correctly to the second and third item.

- (6) *Prota* efaje mia banana. *Meta* eklise
 First ate–PST.PFV.3SG a banana. Then closed–PST.PFV.3SG
 tis kurtines.
 the curtains.
 'First she ate a banana. Then she closed the curtains.'



Figure 1. Visual display for the first two practice sentences (Example: He put on his coat) and the control sentences.

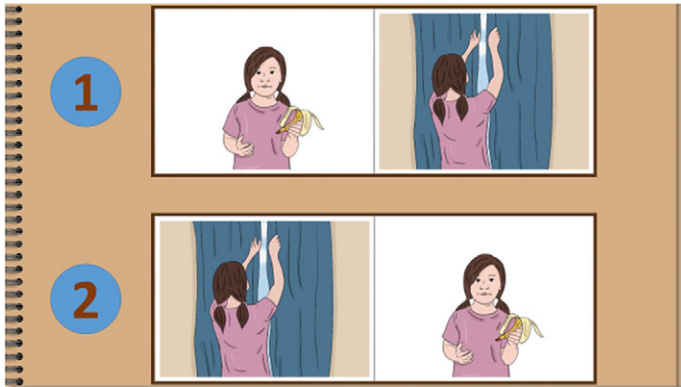


Figure 2. Visual display for example (6). The same visual set-up was used for the test sentences.

Presentation order of the experimental items was randomized as follows: after the practice phase, a picture with the four actors appeared (see Fig. 3) and the child could choose the first actor. A block of six test and two control items was associated with each actor; after completion of the first block, the picture with the four actors appeared again and the child chose the next actor. During the main test phase, the experimenter noted the child's answers on an answer sheet, without providing response-contingent feedback.

The picture-sequence selection task was administered before the sentence repetition and the short-term memory tasks. The entire testing session lasted approximately 30 min. Children received a sticker as a reward for their participation.

Results

Group analysis

The control sentences were answered correctly by all children, and were not considered further. Each correct response to a test item in the picture-sequence selection task received 1 point, and each incorrect response received 0 points, yielding a maximum



Figure 3. Introductory slide of the four test blocks.

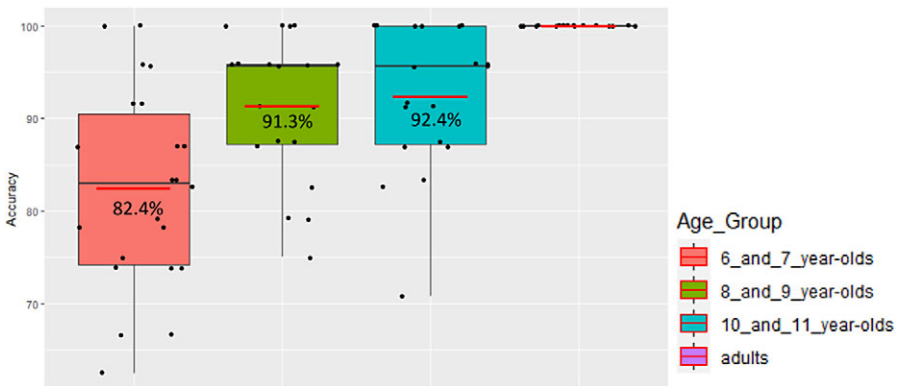


Figure 4. Boxplots depicting overall accuracy (in %) per age group. Red lines indicate the mean.

score of 6 points per condition and of 24 points in total; these raw points were converted to percentages. Wilcoxon tests for independent samples⁶ revealed no difference between children assigned to list A ($N = 32$) and children assigned to list B ($N = 28$), in both the overall score ($W = 415.5$, $p = 0.633$) and the scores per condition (iconic *after*: $W = 431$, $p = 0.740$; non-iconic *after*: $W = 464.5$, $p = 0.806$; iconic *before*: $W = 438.5$, $p = 0.877$; non-iconic *before*: $W = 403.5$, $p = 0.409$). Adult participants performed at ceiling, independent of list. Accordingly, results of the two experimental lists are presented together.

Q1: How does comprehension of sentences with *after* and *before* develop between the ages of six and twelve?

Children achieved a mean overall accuracy of 88.4% ($min = 62.5$, $max = 100$, $SD = 10$). There was a moderate correlation between accuracy and age (in months) ($\rho = 0.507$, $p < 0.001$), indicating that overall comprehension improves with age. For our first analysis based on overall accuracy, we divided the children into three age groups: 6;1–7;10 ($N = 22$), 8;1–9;11 ($N = 19$), and 10;0–11;11 ($N = 19$). This allowed closer comparison with the adult-control group and with the age ranges tested previously. Fig. 4 shows the mean overall accuracy in each of the three child groups and the adult control group; individual dots represent the overall accuracy of individual participants. The youngest age group exhibited the lowest accuracy ($M = 82.4\%$), with large interindividual variation ($min = 62.5$, $max = 100$, $SD = 10.7$). Performance of the two older child groups was similar (8- and 9-year-olds: $M = 91.3\%$, $SD = 7.8$; 10- and 11-year-olds: $M = 92.4\%$, $SD = 7.9$). A Kruskal-Wallis test revealed a significant effect of group ($\chi^2 = 32.836$, $df = 3$, $p < 0.001$); pairwise comparisons with Bonferroni correction showed that all three child groups performed significantly different from the adults (all $ps < 0.001$). The 6- and 7-year-olds differed from the 8- and 9-year-olds ($p = 0.038$) and the 10- and 11-year-olds ($p = 0.019$), while the two older child groups did not differ significantly from each other ($p = 1.0$).

⁶Non-parametric tests were applied, because the data obtained from the picture-sequence selection task violated the normality assumption in the majority of cases.

Q2: How do iconicity, semantics of the connective, and clause order influence children’s comprehension?

Addressing Q2, we calculated the mean accuracies per condition (see Fig. 5) and the development across age for each condition (see Fig. 6). In both figures, individual dots indicate the data points of individual participants. The width of the violin plots varies by the density of data points in a specific region. As can be seen in Fig. 5, the upper part of the four violins (iconic *after*, iconic *before*, non-iconic *before*) is much wider than the lower part, indicating that the majority of children scored at or close to ceiling in these conditions (iconic *after*: $M = 94.2\%$, $SD = 13.7$; iconic *before*: $M = 91.4\%$, $SD = 12.8$; non-iconic *before*: $M = 93.1\%$, $SD = 13.5$). The violin for non-iconic *after* is longer than the other three and exhibits less variation regarding width, indicating that individual scores are more evenly dispersed across the accuracy scale. Children’s mean performance in this

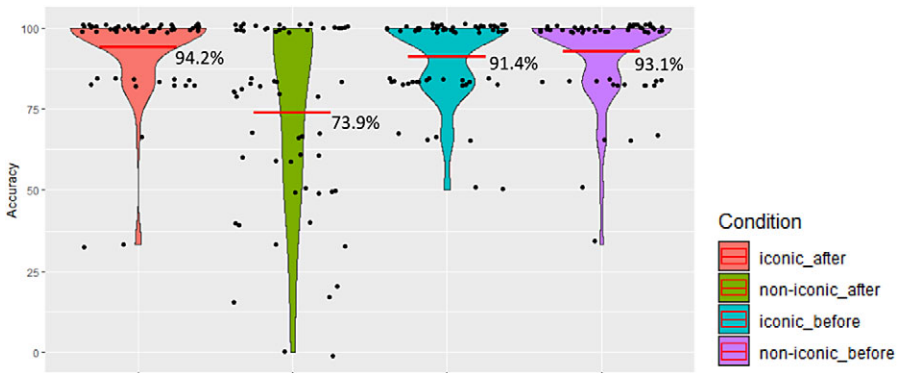


Figure 5. Violin plots depicting accuracy (in %) per condition in the group of child participants. Red lines indicate the mean.

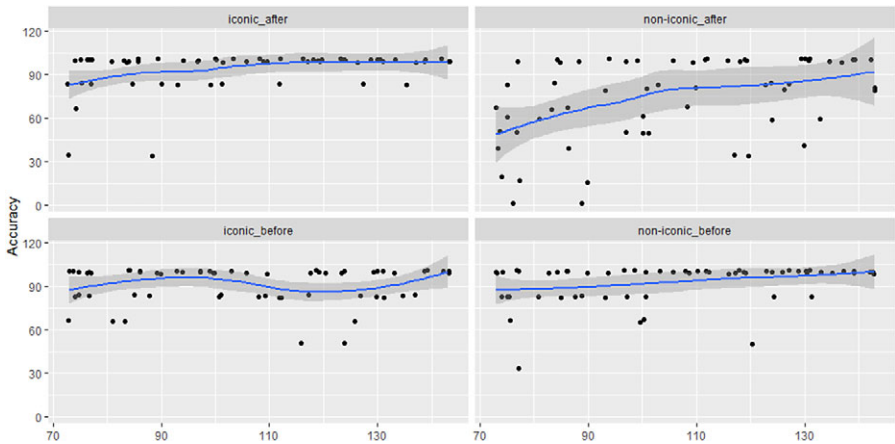


Figure 6. Correlations between accuracy and age by condition. Age of the participants in months is plotted on the X-axis. The shaded areas indicate the 95% confidence intervals.

condition was lower ($M = 73.9\%$, $SD = 28.6$), but above chance (one-sample Wilcoxon signed-rank test: $V = 1351.5$, $p < 0.001$).

As can be inferred from Fig. 6, there was a weak correlation between accuracy and age for iconic *after* ($\rho = 0.341$, $p = 0.004$) and non-iconic *before* ($\rho = 0.336$, $p = 0.004$); accuracy and age were not correlated for iconic *before* ($\rho = 0.035$, $p = 0.394$). Regarding non-iconic *after*, the correlation between accuracy and age was moderate ($\rho = 0.411$, $p < 0.001$), suggesting that performance improves with age. Several children scored around or below chance level, as illustrated in Fig. 6 upper-right panel (see Section *Individual analysis* for a detailed look at children's responses at the individual level).

To test the effects of iconicity and connective on children's performance, we built a generalized linear mixed-effects model using the *lme4* package (Bates et al., 2015) of R (R Core Team, 2022). The model included Iconicity (iconic, non-iconic), Connective (*before*, *after*), and their interaction as fixed effects; both variables were sum-coded (iconic = 0.5, non-iconic = -0.5; *before* = 0.5, *after* = -0.5). The maximal random-effects structure for which the model converged included random intercepts and slopes for participants, and random intercepts for items. The model revealed no main effects of Iconicity and Connective, but an interaction between Iconicity and Connective (see Table 1).

Pairwise comparisons with Tukey adjustment were computed using the package *emmeans* (Lenth et al., 2021) to explore the two-way interaction. The results showed that iconic and non-iconic *before* did not differ significantly ($p = 0.650$), but iconic and non-iconic *after* did ($p < 0.001$). That is, violation of iconicity negatively affected comprehension of *after*, but not of *before*. There was no difference between iconic *before* and iconic *after* ($p = 0.167$), whereas non-iconic *before* was easier than non-iconic *after* ($p < 0.001$), which indicates that *before*-sentences were easier than *after*-sentences only when iconicity was violated. Note that performance on non-iconic *after* was also significantly lower than on iconic *before* ($p = 0.036$), whereas performance on non-iconic *before* and iconic *after* did not differ significantly ($p = 0.934$).

To test whether these response patterns change with age, we divided our sample via median-split into a younger (age range: 6;1–8;7) and an older (age range: 8;10–11;11) subgroup and ran the model separately for each subgroup. Both models revealed the same interaction between Iconicity and Connective as the model for the whole sample, suggesting that the response patterns do not change between the ages of 6 and 12. The models for the subgroups can be found in the Appendix.

The group analyses presented here indicate that non-iconic *after* is the most difficult condition. Inspection of the plots in Fig. 6 points to quite some individual variation in each condition, which raises the question of whether children master sentences with temporal connectives in a specific order, e.g., iconic earlier than non-iconic orders or

Table 1. Output of baseline model

	Estimate	Standard error	z-value	Pr(> z)
(Intercept)	2.754	0.200	13.757	<0.001
Iconicity	0.734	0.394	1.863	0.062
Connective	0.401	0.301	1.332	0.183
Iconicity*Connective	-2.644	0.598	-4.423	<0.001

Model formula: `glmer(Result ~ Iconicity*Connective + (1 + Iconicity + Connective| IDNumber) + (1| Item), control = glmerControl(optimizer = "bobyqa"), family = binomial(link = "logit"), data)`.

before earlier than *after*. These questions are addressed in the individual analysis presented below.

Individual analysis

Analysis of children's response patterns at the individual level was based on calculating mastery. Given that responses to the test items were binary and that each condition included six items, above-chance performance is achieved with five or six correct responses (binomial test). A child with above-chance performance was classified as a 'masterer' of that condition, and a child with less than five correct responses as a 'non-masterer' of that condition. All children showed mastery of at least one condition.

Overall, 29 of the 60 children (48.3%) had mastered all four conditions (MASTERY), with the youngest child aged 6;3. The 31 children who did not show mastery of all conditions were classified according to the different factors discussed above: iconicity, semantics of connective, clause order, and interaction of iconicity and clause order. Five children exhibited only mastery of the iconic conditions, reflecting an order of mention strategy (ICONICITY). Only one child showed mastery of *before*, but not of *after*, exhibiting a developmental advantage of *before* (BEFORE). No child had mastered only sentences in the order main-subordinate, while three children had mastered only sentences in the order subordinate-main, reflecting an advantage of sentence-initial cues (SUB-MAIN ORDER, see Section Discussion). Notably, 18 children had difficulty only with non-iconic *after* (INTERACTION OF ICONICITY AND CAUSE ORDER). The remaining four children did not exhibit a specific pattern (NON-CLASSIFIABLE). The distribution of response patterns across the three age groups is depicted in Fig. 7.

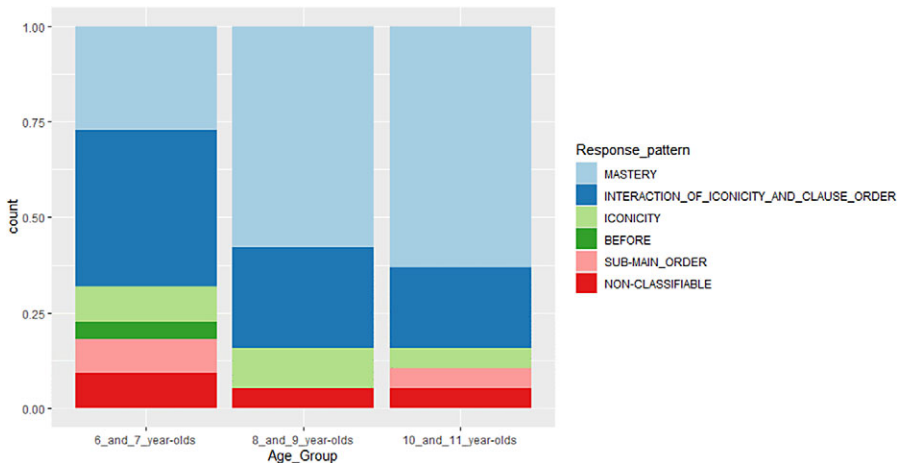


Figure 7. Barplot depicting the distribution (in %) of response patterns in each age group.

Notes. MASTERY: +iconic *before*, +non-iconic *before*, +iconic *after*, +non-iconic *after*; INTERACTION OF ICONICITY AND CLAUSE ORDER: +iconic *before*, +non-iconic *before*, +iconic *after*, -non-iconic *after*; ICONICITY: +iconic *before*, -non-iconic *before*, +iconic *after*, -non-iconic *after*; BEFORE: +iconic *before*, +non-iconic *before*, -iconic *after*, -non-iconic *after*; SUB-MAIN ORDER: -iconic *before*, +non-iconic *before*, +iconic *after*, -non-iconic *after*

Table 2. Distribution of children who showed (no) mastery of non-iconic *before* and non-iconic *after*

		Non-iconic <i>after</i>	
		Mastery	No mastery
Non-iconic <i>before</i>	Mastery	32	23
	No mastery	0	5

Note. Mastery: correct answers for at least five of the six test items of a given condition

The proportion of children who mastered all conditions increased with age: 27.3% (6/22 children) in the group of 6- and 7-year-olds, 57.9% (11/19 children) in the group of 8- and 9-year-olds, and 63.1% (12/19 children) in the group of 10- and 11-year-olds. The proportion of children failing mastery in one or more conditions gradually decreased; this is particularly visible in the difference between the two younger age groups. These results are in line with the findings for the group level: there, the youngest age group scored lower than the other two groups, which did not differ from one another.

Across all three age groups, the most frequent error pattern was mastery of all conditions except for non-iconic *after*; this pattern was exhibited by 40.9% (9/22 children) of the 6- and 7-year-olds, 26.3% (5/19 children) of the 8- and 9-year-olds, and 21.0% (4/19 children) of the 10- and 11-year-olds. Therefore, the individual analysis confirmed that non-iconic *after* is the most difficult of the four conditions to acquire.

To see whether mastery of non-iconic *after* indeed follows mastery of non-iconic *before* across all children, in a final step we compared individual mastery of non-iconic *after* and non-iconic *before*. As presented in Table 2, 32 of the 60 children showed mastery of both conditions, and five children had difficulties with both conditions. Notably, 23 children showed mastery of non-iconic *before* and had difficulties with non-iconic *after*, while no child showed the opposite pattern, which confirms that non-iconic *after*-sentences are most difficult.

Q3: To what degree do general language ability and verbal short-term memory modulate children's comprehension?

Addressing Q3, we first calculated the correlations of scores in the forward digit-recall task ($M = 6.25$, $SD = 1.6$), measuring short-term memory, and in the sentence repetition task ($M = 63.8$, $SD = 22.6$), assessing general language ability, with overall accuracy in the picture-sequence selection task. Both correlations were significant (forward digit recall: $\rho = 0.427$, $p < 0.001$; sentence repetition: $\rho = 0.619$, $p < 0.001$); hence, the two variables were added as continuous predictors to the baseline model (see Table 1). Age was included as a factor to ensure that possible effects of memory and language ability are not proxies for the role of age. Table 3 summarizes the correlations between the background variables; the output of the full model is reported in Table 4.

As summarized in Table 4, there was a main effect of Iconicity and an interaction between Iconicity and Connective. Pairwise comparisons with Tukey adjustment revealed the same patterns as in the baseline model: iconicity affected the comprehension of sentences with *after* but not of sentences with *before*, and sentences with *before* were easier than sentences with *after* only in the non-iconic condition. Among the continuous predictors, sentence repetition had an effect, but forward digit recall and age did not.

Table 3. Spearman's *rho* correlations (one-sided) between background variables.

	Age	Forward Digit Recall
Forward Digit Recall	0.681***	
Sentence Repetition	0.593***	0.671***

Note. *** $p < 0.001$

Table 4. Output of the model including Age, Forward Digit Recall, and Sentence Repetition as predictors.

	Estimate	Standard error	z-value	Pr(> z)
(Intercept)	2.890	0.210	13.749	<0.001
Iconicity	1.075	0.449	2.391	0.017
Connective	0.439	0.399	1.098	0.272
Age	0.294	0.190	1.549	0.121
Forward Digit Recall	-0.041	0.204	-0.202	0.840
Sentence Repetition	0.524	0.187	2.796	0.005
Iconicity*Connective	-2.869	0.776	-3.698	<0.001

Model formula: `glmer(Result ~ Iconicity*Connective + Age + Forward_Digit_Recall + Sentence_Repetition + (1 + Iconicity * Connective | IDNumber) + (1 | Item), control = glmerControl(optimizer = "bobyqa"), family = binomial(link = "logit"), data)`.

Discussion

The present study investigated how the comprehension of complex sentences containing the temporal connectives *before* and *after* develops throughout childhood. We asked to what extent iconicity, the semantics of the connective, and/or clause order can explain children's difficulties and whether what children struggle with (e.g., with non-iconicity, with *after*) remains the same across ages. Sixty monolingual six- to twelve-year-old typically developing Greek-speaking children were tested with a carefully designed picture-sequence selection task, which was similar to the design used by de Ruiter and colleagues (2018a, 2018b, 2019, 2020a) but improved on the aspect of event completion by using only telic predicates, in the past tense. Below we discuss our findings, addressing our three research questions in turn.

Comprehension of sentences with "after" and "before" develops up to age twelve (Q1)

Overall performance in the group of six- to twelve-year-old children was quite high (82–92%). At the same time, even the ten- and eleven-year-olds did not perform as well as adults, confirming previous findings that full comprehension of temporal connectives remains a challenge into late childhood (Karlsson et al., 2019; Overweg et al., 2018; Papakonstantinou, 2015; Pyykkönen & Järvikivi, 2012; Wagner & Holt, 2023). The overall accuracy scores of our three age groups are comparable to those reported by Papakonstantinou (2015) for Greek-speaking children of approximately the same age (seven-year-olds: 77% correct, nine-year-olds: 89% correct, eleven-year-olds: 92% correct). Notably, we found that performance improved significantly between the group of six- and seven-year-olds and the group of eight- and nine-year-olds, but did not improve further in the

oldest age group. The individual analysis confirmed this pattern: the proportion of children who mastered all four conditions of the experiment increased from the youngest to the middle age group, but then remained unchanged. Thus, age alone cannot fully explain the results.

Children's comprehension is influenced by iconicity in interaction with clause order (Q2)

The design of our study allowed us to examine the four main accounts to explain children's comprehension of temporal sentences: (A) iconicity, arguing that sentences following the chronological order of the events are understood better than non-iconic sentences (e.g., Clark, 1971; de Ruiter et al., 2018a, 2018b); (B) the semantics of the connective, positing a developmental advantage of *before* over *after* (e.g., Clark, 1971); (C) clause order, claiming that sentences in main-subordinate order are easier to comprehend than sentences in the order subordinate-main (Diessel, 2004, 2005, 2008); and (D) interaction of iconicity and clause order, arguing that non-iconic sentences are difficult only in main-subordinate clause order (Pyykkönen & Järviö, 2012). Our review of previous research indicated that some effects were found for iconicity in children younger than eight years and for the semantics of the connective, independent of age, while no effects were documented for clause order. The one study to find an interaction of iconicity and clause order was a reading study.

Our own findings provide partial support for account (A), no support for account (B), counterevidence for account (C), and clear evidence for account (D). Turning first to the clause order account (C), our data clearly contradict its prediction – that is, the clause order main-subordinate facilitates comprehension. Clause order did not affect performance of *before*-sentences. Moreover, *after*-sentences in the order subordinate-main in fact had an advantage over *after*-sentences in the order main-subordinate both at the group and at the individual level. Three children even exhibited better comprehension of the order subordinate-main for both connectives. Our findings are in line with the majority of previous research on languages with clause-initial adverbial connectives, which did not find effects of clause order (e.g., Blything & Cain, 2016; Clark, 1971; de Ruiter et al., 2018a, 2018b; Feagans, 1980). Papakonstantinou (2015), who tested monolingual Greek-speaking children of roughly the same age as we did, found better comprehension of *after*-sentences in the order main-subordinate than in the order subordinate-main, but no effects of clause order for *before*-sentences. In our view, task factors could explain this difference: we employed a picture-sequence selection task, which excluded simultaneous responses, whereas Papakonstantinou (2015) used an act-out task and included simultaneous responses in her coding scheme.

Regarding the iconicity account (A), our group results showed that non-iconic *after*-sentences were more difficult than iconic *after*-sentences, but that there was no difference between iconic and non-iconic *before*-sentences. Put differently, iconicity influenced comprehension of *after* but not of *before*, arguing against a general role for iconicity. Moreover, the individual analysis revealed that iconicity accounted only for a small subset of the children: in line with the order-of-mention strategy proposed by Clark (1971), five children (aged between 6;4 and 10;0) exhibited mastery of the iconic and non-mastery of the non-iconic conditions. Most of the children who did not perform at ceiling (18/31), however, had difficulty only with non-iconic *after*. In addition, as noted above, three children mastered the order subordinate-main but not the order main-subordinate for both connectives; that is, non-iconic *before* was in fact understood better than iconic

before. Accordingly, our data clearly show that children between the ages of six and twelve do not generally find the violation of iconicity a challenge for comprehension.

As for Account (B), the semantics of the connective *per se* did not play a major role; that is, *before* had no general advantage over *after*. At the group level, *before* was easier than *after* only in the non-iconic condition. Moreover, in the individual analysis only one child showed mastery of both variants of *before*, but not of *after*. Consequently, Clark's (1971) semantic components hypothesis cannot explain the comprehension difficulties of children between the ages of six and twelve.

Account (D) predicts that children's comprehension is mediated by the interaction of iconicity and clause order (see Pyykkönen & Järvikivi, 2012, for reading). This is exactly what we found in our data: in the order subordinate-main, iconic *after*-sentences and non-iconic *before*-sentences were understood equally well, but in the order main-subordinate, non-iconic *after*-sentences were more difficult than iconic *before*-sentences. This pattern was present at the group level and was also the predominant individual response pattern, exhibited by 58% of those children who were not-adult-like. Similar in spirit to Pyykkönen and Järvikivi's (2012) account, we propose that children's difficulty with non-iconic *after*-sentences is comparable to the 'kindergarten-path effect' documented for children's processing of temporary syntactic ambiguities (e.g., Trueswell et al., 1999).⁷

We refer to our account as the "event-semantic kindergarten-path". Let us illustrate this account first informally. Imagine two events *x* and *y* expressed by the clauses *X* and *Y*, respectively, and imagine that the order of events is *x, y*. In the sequence "*after X, Y*", sentence-initial *after* signals that the sentence is in line with iconicity: *after X, Y* => *x, y*. In the sequence "*before Y, X*", sentence-initial *before* signals that reordering is required to build the correct event representation of the complex sentence: *before Y, X* => *x, y*. In this case, the clause order violates iconicity, but the cue occurs early. Given that iconicity is assumed as the default for event ordering, reanalysis would occur upon encountering *before* and would be resolved very quickly.

In contrast, if the connective appears sentence-medially, explicit information about the order of the events is provided after the main clause has been processed. In the sequence "*X, before Y*", sentence-medial *before* requires subordinating this clause under the main clause, but the order of events (*x, y*) remains unchanged and is in line with iconicity: *X, before Y* => *x, y*. Note that the cue signaling the continuation of the sentence appears late; therefore, listeners may expect the sentence to be completed when they encounter the connective, and this may in principle impede comprehension (Dan Parker, p.c.). The children tested in the current study did not show lower performance with *before* in this condition, which suggests that 'surprisal' *per se* does not negatively impact children's performance, at least not in off-line tasks. Sentence-medial *after*, on the other hand, requires revision of the initial event representation: the listener establishes an event *x* for the first clause; after encountering *after*, another event has to be established that is prior to event *x*. As a result, *x* has to be revised as the second event and the second event has to be established as the first one: *Y, after X* => *x, y*. We suggest that children find this revision most difficult. The rationale is as follows: the proposed main clause provides no information about the order of events and by default creates the expectation that the next event sentence will move the action forward in time. When encountering the connective, children cannot easily overcome their initial expectation and, as a result, they often

⁷See Makrodimitris and Schulz (2021b) for suggesting this explanation for bilingual children.

maintain the incorrect iconic interpretation of the event order (y, x).⁸ Further support for this account comes from our individual analysis, which showed that children did not master non-iconic *after*-sentences until they had mastered non-iconic *before*-sentences.

Importantly, we assume that 'reordering' of events occurs at the conceptual level, yielding a correct mental representation of the events in the world. At the discourse-semantic level, discourse representations (DR, see Kamp, 1984) are built sequentially as sentences are encountered. To model the revision process, we assume that each DR contains event variables and a default ordering relation $e_x < e_{x+1}$ (" $<$ " for the relation of complete precedence between events, see Partee, 1984). When a non-iconic clause is encountered, this ordering relationship must be canceled and reversed.⁹

Our cross-sectional findings suggest a uniform developmental path for comprehension of temporal sentences. We assume that children initially use the order-of-mention strategy, because they have poor knowledge of the connectives. At some point, they master comprehension of non-iconic *before* but still face difficulty with non-iconic *after*, i.e., they are in the event-semantic kindergarten-path stage. Our data suggest that this transition from the iconicity stage to the event-semantic kindergarten-path stage takes place before the age of six. Given that this was our youngest age group, we have to leave open when exactly this transition occurs. Finally, mastery of non-iconic *after* is achieved very late, after age eleven. Structures with early revision cues such as non-iconic *before*-sentences may help children realize that it is possible to sidestep iconicity in discourse, and may prepare them to cope with cases in which violations of iconicity are signaled late, i.e., sentence-medially.

Note that the proposed stages are only loosely tied to age: although mastery was more frequently found among the older children, some six-year-olds had already reached mastery, and some eleven-year-olds were still in the event-semantic kindergarten-path stage. The developmental path proposed here differs from that of Clark (1971) in one important respect: Clark assumes that some children may proceed from the order-of-mention strategy to a stage at which they treat *after* as if it meant *before*. Our data provide no strong evidence for such a stage, as only one out of 60 children showed this pattern. The order-of-mention strategy was the predominant response pattern in Clark's (1971) study, but it was used by only few children (5/60) in the current study. This difference may be due to the different age ranges examined (6;1 to 11;11 on our study, 3;0 to 5;0 in Clark, 1971). It is likely that the majority of the children in our study had already proceeded to the next developmental stages.

General language ability is a stronger predictor of children's performance than verbal short-term memory (Q3)

Our statistical analysis showed that both general language ability, measured via sentence repetition, and verbal short-term memory, measured via forward digit recall, were positively correlated to overall accuracy in the comprehension of temporal sentences. When entered into the same model along with age, only general language ability was a significant predictor of performance. This suggests first, that morpho-syntactic abilities are more crucial for processing complex sentences at the semantic level than short-term working memory. This is

⁸Note that empirically, this response is indistinguishable from ignoring the connective *after* in this position, which would also result in an incorrect iconic interpretation.

⁹Thanks to an anonymous reviewer for asking about the concrete representation and to Cornelia Ebert for discussion of the issue of cancellation.

not to say that memory does not play any role: sentence repetition tasks inevitably draw on children's memory capacity, as evidenced by the strong correlation between the two variables in our data ($\rho = 0.671$, $p < 0.001$, see also Marinis & Armon-Lotem, 2015). The effect of sentence repetition may have been stronger than the effect of forward digit recall, because the former captures an aspect of verbal memory that is more similar to processing of the sentences in our picture-sequence selection task. Note that most previous studies either found memory measures to be stronger predictors of performance than language measures (Blything et al., 2015; Blything & Cain, 2016; de Ruiter et al., 2018b) or that neither memory nor language measures predicted performance (de Ruiter et al., 2018a, 2020a). However, these studies used language assessments with different demands than sentence repetition, mainly vocabulary tasks, which may explain the difference with our findings. Second, our results suggest that understanding complex temporal sentences is driven by morpho-syntactic abilities rather than simply by age: although age certainly plays some role, as indicated by the moderate correlation between age and overall comprehension accuracy ($\rho = 0.507$, $p < 0.001$), sentence repetition seems to be a better way to capture children's developmental stage regarding grammar.

Limitations

The event-semantic processing account proposed here refers to the ability to revise event representations. This ability may require several memory resources, but the current study included only short-term memory measured via forward digit recall. As a case in point, Karlsson et al. (2019) found that comprehension of non-iconic *after*-sentences in reading was predicted by updating abilities. Future studies should explore links between different cognitive measures (e.g., complex working memory, inhibition, updating) and specific child response patterns. Moreover, the present study – just like previous studies – used a cross-sectional design. Future longitudinal studies could help to substantiate the developmental path we have proposed. In particular, data from children over the age of twelve would help to determine at what age language learners finally become fully adult-like.

The test sentences of our experiment were presented in out-of-the-blue contexts: children would hear a sentence (e.g., *After he cleaned his bicycle, he wrote a postcard.*) and then see two picture sequences that only differed in the order of the two events. The actor was introduced before, but none of the events could be viewed as given information. Manipulating the factor givenness, de Ruiter et al. (2020b) found that providing a context sentence (e.g., *Sue crawls on the floor. Before she crawls on the floor, she hops up and down.*) improved performance compared to presenting the test sentences in isolation (see also Gorrell et al., 1989). Accordingly, future research could include the factor givenness to see if the event-semantic kindergarten-path effect can be modulated by presenting the sentences in given-before-new contexts. Finally, future research should examine the role of frequency of the connectives in both clause orders in children's input across languages (for *before* and *after* in English, see de Ruiter et al., 2021). Studies of connective use based on child corpora of Greek (Baslis, 1994; Stephany, 1997) have analyzed children's productions but not the adult speech directed to children, so we could not incorporate this aspect in our analysis.

Outlook

We would like to suggest two avenues for further research: acquisition studies involving languages with clause-final temporal morphemes and, second, adult processing studies.

To uncover to what extent the event-semantic kindergarten-path account holds cross-linguistically, we need to examine languages with clause-initial temporal connectives, such as Greek, as well as languages with clause-final temporal morphemes, such as Mandarin Chinese or Tamil. To arrive at robust conclusions, it is crucial to first determine the presence of both clause orders (subordinate-main and main-subordinate) in the respective language. As a case in point, the order main-subordinate in Mandarin Chinese has been reported to be less frequent than the order subordinate-main (de Ruiter et al., 2020a) or even ungrammatical (Le'an Luo, p.c.). Similarly, Tamil may not allow the order main-subordinate for temporal sentences (Shanmugasundaram Rajamathangi, p.c.). More cross-linguistic studies are needed to determine which languages allow both orders, subordinate-main and main-subordinate, for temporal sentences. Languages that express the meaning of *before* and *after* via clause-final morphemes, such as Tamil and Mandarin Chinese, may prove especially relevant in this regard.

As for the second point, under the assumption that what is difficult for children slows down processing in adults (see Friedmann et al., 2009, for comprehension of object relative clauses), adults are predicted to show an event-semantic kindergarten-path under the right conditions. First evidence comes from a reading study with *before*- and *after*-sentences by Scholman et al. (2022): the authors found non-iconic *after* to exhibit the longest regression path duration, which is argued to reflect the process of integrating the linguistic material with the previous context (Rayner, 1998). To substantiate this point, further studies on adult processing are needed.

Conclusion

The current study showed that six-to-twelve-year-old monolingual children have mastered some but not all aspects of understanding complex sentences containing *after* and *before*, with comprehension of non-iconic *after*-sentences being the most difficult. We argued that this selective difficulty cannot be explained by iconicity, the semantics of *after*, or clause order alone, but by an event-semantic kindergarten-path. According to this processing account, children have difficulty reanalyzing the event-semantic representation of a complex sentence when the cue requiring this event-semantic reanalysis appears late; in this case, children may maintain their initial incorrect discourse representation. The current results show that children acquiring Greek behave similarly to children acquiring Germanic languages, such as German or English. Much more cross-linguistic research, especially on non-Indo-European languages, is needed to shed light on the boundaries between universal and language-specific processing and acquisition principles (see also Kidd & Garcia, 2022).

Acknowledgments. We are grateful to the children who participated in the study and to their parents, as well as to their teachers for their support. We would also like to thank Alexandra Lowles for creating the pictures of our experiment, Ilona-Eirini Pistopoulou for recording the sentences of the experiment, Alexandra Karousou and Tarsi Christodoulou for helping us with data collection and transcription, Jacopo Torregrossa for statistical advice, and Merle Weicker for discussion of adverbial clauses.

Competing interest. The authors declare none.

Ethics statement. The study was conducted according to the guidelines of the Declaration of Helsinki, and its procedures were approved by the Institute of Educational Policy of the Greek Ministry of Education and Religious Affairs (protocol code: 44/10-09-2020, date of approval: 10 September 2020), and the Ethics Committee of the DIPF|Leibniz Institute for Research and Information in Education (protocol code: DIPF_EK_DazabSechs, date of approval: 24 February 2020).

Funding. The research presented here was conducted in the framework of the project “DaZ ab Sechs: The role of language knowledge for the acquisition of German by child second language learners” (PI: Petra Schulz). The project was part of the Research Center IDEA of the DIPF|Leibniz Institute for Research and Information in Education and was supported by the Hessian Ministry of Higher Education, Research, Science and the Arts.

References

- Amidon, A.** (1976). Children’s understanding of sentences with contingent relations: Why are temporal and conditional connectives so difficult? *Journal of Experimental Child Psychology*, *22*(3), 423–437. doi: [https://doi.org/10.1016/0022-0965\(76\)90106-5](https://doi.org/10.1016/0022-0965(76)90106-5)
- Amidon, A., & Carey, P.** (1972). Why five-year-olds cannot understand *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, *11*(4), 417–423. doi: [https://doi.org/10.1016/S0022-5371\(72\)80022-7](https://doi.org/10.1016/S0022-5371(72)80022-7)
- Baslis, Y.** (1994). The development of subordinate clauses in the language of Greek children. In I. Philippaki-Warbuton, K. Nicolaidis, & M. Sifianou (Eds.), *Themes in Greek linguistics: Papers from the First International Conference on Greek Linguistics, Reading, September 1993* (pp. 333–340). Amsterdam, Philadelphia: John Benjamins.
- Bates, D., Mächler, M., Bolker, B., & Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K.** (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*, 235–261. doi: <https://doi.org/10.7916/D8X06785>
- Blything, L. P., & Cain, K.** (2016). Children’s processing and comprehension of complex sentences containing temporal connectives: The influence of memory on the time course of accurate responses. *Developmental Psychology*, *52*(10), 1517–1529. <https://doi.org/10.1037/dev0000201>
- Blything, L. P., Davies, R., & Cain, K.** (2015). Young children’s comprehension of temporal relations in complex sentences: The influence of memory on performance. *Child Development*, *86*(6), 1922–1934. <https://doi.org/10.1111/cdev.12412>
- Chondrogianni, V., Andreou, M., Nerantzini, M., Varlokosta, S., & Tsimpli, I. M.** (2013). *The Greek Sentence Repetition Task*. COST Action IS0804.
- Clark, E. V.** (1969). *Language acquisition: The child’s spontaneous descriptions of events in time* (Doctoral dissertation). University of Edinburgh, UK.
- Clark, E. V.** (1971). On the acquisition of the meaning of *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, *10*, 266–275. [https://doi.org/10.1016/S0022-5371\(71\)80054-3](https://doi.org/10.1016/S0022-5371(71)80054-3)
- Clark, E. V.** (2009). *First language acquisition* (2. ed., 3. printing). Cambridge: Cambridge University Press.
- Coker, P. L.** (1978). Syntactic and semantic factors in the acquisition of *before* and *after*. *Journal of Child Language*, *5*(2), 261–277. <https://doi.org/10.1017/S0305000900007467>
- de Ruiter, L. E., Chen, S., Etz, A., & Wen, P.** (2020a). Is children’s iconicity preference universal? Evidence from Mandarin adverbial sentences. Paper presented at the 33rd Annual CUNY Human Sentence Processing Conference, March 19–21, Amherst, MA, USA.
- de Ruiter, L. E., Lemen, H. C. P., Lieven, E. V. M., Brandt, S., & Theakston, A. L.** (2021). Structural and interactional aspects of adverbial sentences in English mother–child interactions: An analysis of two dense corpora. *Journal of Child Language*, *48*(6), 1150–1184. <https://doi.org/10.1017/S0305000920000641>
- de Ruiter, L. E., Lieven, E. V. M., Brandt, S., & Theakston, A. L.** (2020b). Interactions between givenness and clause order in children’s processing of complex sentences. *Cognition*, *198*: 104130. <https://doi.org/10.1016/j.cognition.2019.104130>
- de Ruiter, L. E., Priyadharshini, V., Etz, A., & Kuppuraj, S.** (2019). *Tamil-speaking children do not prefer iconic adverbial sentences over non-iconic ones*. Paper presented at the 44th Boston University Conference on Language Development (BUCLD 44), November 7–10, Boston, MA, USA.
- de Ruiter, L. E., Theakston, A. L., Brandt, S., & Lieven, E. V. M.** (2018a). Iconicity affects children’s comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition*, *171*, 202–224. <https://doi.org/10.1016/j.cognition.2017.10.015>
- de Ruiter, L. E., Theakston, A. L., Lieven, E. V. M., Hilton, M., & Brandt, S.** (2018b). *Is before easier than after in German, too? Testing the relative influence of iconicity, ambiguity, and language-specific frequencies*

- on the processing of adverbial sentences in German. Paper presented at the 43rd Boston University Conference on Language Development (BUCLD 43), November 2–4, Boston, MA, USA.
- Diessel, H.** (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.
- Diessel, H.** (2005). Competing motivations for the ordering of main and adverbial clauses. *Linguistics*, **43**(3), 449–470. <https://doi.org/10.1515/ling.2005.43.3.449>
- Diessel, H.** (2008). Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics*, **19**(3), 465–490. <https://doi.org/10.1515/COGL.2008.018>
- Feagans, L.** (1980). Children's understanding of some temporal terms denoting order, duration, and simultaneity. *Journal of Psycholinguistic Research*, **9**(1), 41–57. <https://doi.org/10.1007/bf01067301>
- French, L. A., & Brown, A. L.** (1977). Comprehension of *before* and *after* in logical and arbitrary sequences. *Journal of Child Language*, **4**(2), 247–256. <https://doi.org/10.1017/S0305000900001641>
- Friedmann, N., Belletti, A., & Rizzi, L.** (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, **119**(1), 67–88. <https://doi.org/10.1016/j.lingua.2008.09.002>
- Georgas, J., Paraskevopoulos, I. N., Besevegis, E., & Giannitsas, N.** (1997). *Greek WISC-III: Wechsler intelligence scale for children*. Athens: Ellinika Grammata.
- Gorrell, P., Crain, S., & Fodor, J. D.** (1989). Contextual information and temporal terms. *Journal of Child Language*, **16**(3), 623–632. <https://doi.org/10.1017/S0305000900010758>
- Greenberg, J. H.** (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of Grammar* (2nd ed., pp. 73–113). Cambridge, MA: MIT Press.
- Grice, H. P.** (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds.), *Speech Acts* (pp. 41–58). Leiden: Brill.
- Haiman, J.** (2006). Iconicity. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (2nd ed., Vol. 5, pp. 457–461). Amsterdam: Elsevier. doi: <https://doi.org/10.1016/B0-08-044854-2/00194-2>
- Hamann, C., Lindner, K., & Penner, Z.** (2001). Tense, reference time, and language impairment in German children. In C. Féry & W. Sternefeld (Eds.), *Audiatur vox sapientiae: A Festschrift for Arnim von Stechow* (pp. 182–213). Berlin: Akademie Verlag.
- Haspelmath, M.** (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, **19**(1), 1–33. <https://doi.org/10.1515/COG.2008.001>
- Hatch, E.** (1971). The young child's comprehension of time connectives. *Child Development*, **42**(6), 2111–2113. <https://doi.org/10.2307/1127618>
- Holton, D., Mackridge, P., & Philippaki-Warbuton, I.** (1999). *Greek: A comprehensive grammar of the modern language* (2nd ed.). London, New York: Routledge.
- Johnson, H. L.** (1975). The meaning of *before* and *after* for preschool children. *Journal of Experimental Child Psychology*, **19**(1), 88–99. [https://doi.org/10.1016/0022-0965\(75\)90151-4](https://doi.org/10.1016/0022-0965(75)90151-4)
- Kamp, H.** (1984). A Theory of Truth and Semantic Representation. In J. Groenendijk, T. M. V. Janssen, & M. Stokhof (Eds.), *Truth, Interpretation and Information* (pp. 1–42). Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110867602.1>
- Karlsson, J., Jolles, D., Koornneef, A., van den Broek, P., & van Leijenhof, L.** (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating. *Journal of Experimental Child Psychology*, **185**, 1–18. <https://doi.org/10.1016/j.jecp.2019.04.007>
- Kavanaugh, R. D.** (1979). Observations on the role of logically constrained sentences in the comprehension of 'before' and 'after'. *Journal of Child Language*, **6**(2), 353–357. <https://doi.org/10.1017/S0305000900002348>
- Keller-Cohen, D.** (1987). Context and strategy in acquiring temporal connectives. *Journal of Psycholinguistic Research*, **16**(2), 165–183. <https://doi.org/10.1007/BF01072000>
- Kidd, E., & Garcia, R.** (2022). How diverse is child language acquisition research? *First Language*, **42**(6), 703–735. <https://doi.org/10.1177/01427237211066405>
- Klein, W.** (2009). How time is encoded. In W. Klein, & P. Li (Eds.), *The expression of time* (pp. 39–81). Berlin, New York: Mouton de Gruyter.
- Lenth, R. V., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., & Singmann, H.** (2021). *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*. R Package Version 1.5.4. URL: <https://cran.r-project.org/web/packages/emmeans/index.html> (accessed on 9 May 2024).

- Makrodimitis, C.** (2024). *Bilingual children's comprehension of temporal connectives: The role of age of onset* (Doctoral dissertation). Goethe University Frankfurt, Germany.
- Makrodimitis, C., & Schulz, P.** (2021a). DaZ ab Sechs: Ein Projekt zum Erwerb temporaler Subjunkoren bei griechisch-deutschen Lernern. In A.-L. Scherger, B. Lütke, E. Montanari, A. Müller, & J. Ricart Brede (Eds.), *Deutsch als Zweitsprache - Forschungsfelder und Ergebnisse* (pp. 147–161). Stuttgart: Fillibach bei Klett.
- Makrodimitis, C., & Schulz, P.** (2021b). Iconic sentences are not always easier: Evidence from bilingual Greek-German children. In D. Dionne & L.-A. Vidal Covas (Eds.), *Proceedings of the 45th annual Boston University Conference on Language Development* (pp. 528–541). Somerville, MA: Cascadilla Press. URL: <http://www.lingref.com/bucltd/45/BUCLD45-40.pdf>
- Marinis, T., & Armon-Lotem, S.** (2015). Sentence Repetition. In S. Armon-Lotem, J. d. Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 95–121). Bristol, Buffalo, Toronto: Multilingual Matters. <https://doi.org/10.21832/9781783093137-007>
- Natsopoulos, D., & Abadzi, H.** (1986). Understanding linguistic time sequence and simultaneity: A literature review and some new data. *Journal of Psycholinguistic Research*, *15*(3), 243–273.
- Newmeyer, F.** (1992). Iconicity and generative grammar. *Language*, *68*(4), 756–796. <https://doi.org/10.2307/416852>
- Overweg, J., Hartman, C. A., & Hendriks, P.** (2018). Temporarily Out of Order: Temporal Perspective Taking in Language in Children With Autism Spectrum Disorder. *Frontiers in Psychology*, *9*: 1663. <https://doi.org/10.3389/fpsyg.2018.01663>
- Papakonstantinou, M.** (2015). *Temporal connectives in child language: A study of Greek* (Doctoral dissertation). Aristotle University of Thessaloniki, Greece.
- Partee, B.** (1984). Nominal and temporal anaphora. *Linguistics and Philosophy*, *7*(3), 243–286.
- Prentza, A., Tafiadis, D., Chondrogianni, V., & Tsimpli, I. M.** (2022). Validation of a Greek Sentence Repetition Task with Typically Developing Monolingual and Bilingual Children. *Journal of Psycholinguistic Research*, *51*, 373–395. <https://doi.org/10.1007/s10936-022-09853-z>
- Pyykkönen, P., & Järvikivi, J.** (2012). Children and situation models of multiple events. *Developmental Psychology*, *48*(2), 521–529. <https://doi.org/10.1037/a0025526>
- Rayner, K.** (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- R Core Team** (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Retz, J.** (2020). Eliminating EARLIEST: A general semantics for *before* and *after*. In M. Franke, N. Kompa, M. Liu, J. L. Mueller, & J. Schwab (Eds.), *Proceedings of Sinn und Bedeutung 24* (Vol. 2, pp. 201–218). Osnabrück University. Retrieved October 06, 2022, from <https://semanticsarchive.net/Archive/mZhNDA4Y/SuB24-twovolume.pdf>
- Richards, M. M., & Hawpe, L. S.** (1981). Contrasting patterns in the acquisition of spatial/temporal terms. *Journal of Experimental Child Psychology*, *32*(3), 485–512. [https://doi.org/10.1016/0022-0965\(81\)90110-7](https://doi.org/10.1016/0022-0965(81)90110-7)
- Rothweiler, M.** (1993). *Der Erwerb von Nebensätzen im Deutschen*. Berlin, New York: Max Niemeyer Verlag. <https://doi.org/10.1515/9783111355856>
- Scholman, M. C. J., Blything, L., Cain, K., Hoek, J., & Evers-Vermeul, J.** (2022). Discourse rules: The effects of clause order principles on the reading process. *Language, Cognition and Neuroscience*, 1–15. <https://doi.org/10.1080/23273798.2022.2077971>
- Schulz, P.** (2003). *Factivity: Its nature and acquisition*. Tübingen: Max Niemeyer Verlag.
- Schulz, P.** (2018). Telicity in typical and impaired acquisition. In K. Syrett & S. Arunachalam (Eds.), *Semantics in language acquisition* (pp. 123–150). Amsterdam, Philadelphia: John Benjamins.
- Schulz, P., & Grimm, A.** (2019). The age factor revisited: Timing in acquisition interacts with age of onset in bilingual acquisition. *Frontiers in Psychology*, *9*: 2732. doi: <https://doi.org/10.3389/fpsyg.2018.02732>
- Stephany, U.** (1997). The acquisition of Greek. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Volume 4* (pp. 183–333). Mahwah, NJ: Lawrence Erlbaum.
- Stevenson, R. J., & Pollitt, C.** (1987). The acquisition of temporal terms. *Journal of Child Language*, *14*(3), 533–545. <https://doi.org/10.1017/S0305000900010278>
- Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M.** (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology*, *92*, 87–100. <https://doi.org/10.1016/j.cogpsych.2016.10.003>

- Trosborg, A.** (1982). Children's comprehension of 'before' and 'after' reinvestigated. *Journal of Child Language*, **9**(2), 381–402. <https://doi.org/10.1017/S0305000900004773>
- 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](https://doi.org/10.1016/s0010-0277(99)00032-3)
- Tsakali, V., & Vamvouka, K.** (2021). The development of *before* and *after* in Greek. In T. Markopoulos, C. Vlachos, A. Arxakis, D. Papazachariou, G. Xydopoulos, & A. Roussou (Eds.), *Proceedings of the 14th International Conference on Greek Linguistics (ICGL14)* (pp. 1319–1332). Patras: University of Patras. <https://doi.org/10.26220/icgl.v1i1.3767>
- Tsimpli, I. M.** (2014). Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, **4**(3), 283–313. <https://doi.org/10.1075/lab.4.3.01tsi>
- Tsimpli, I. M., Papadopoulou, D., & Mylonaki, A.** (2010). Temporal modification in Greek adverbial clauses: The role of aspect and negation. *Lingua*, **120**(3), 649–672. <https://doi.org/10.1016/j.lingua.2008.10.008>
- van Hout, A.** (2018). On the acquisition of event culmination. In K. Syrett & S. Arunachalam (Eds.), *Semantics in language acquisition* (pp. 95–121). Amsterdam, Philadelphia: John Benjamins.
- Wagner, L., & Holt, R. F.** (2023). Time after time: Factors influencing children's comprehension of *before* and *after*. *Journal of Child Language*, 1–9. <https://doi.org/10.1017/S0305000923000612>
- Zhang, M., & Hudson, J. A.** (2018). The development of temporal concepts: Linguistic factors and cognitive processes. *Frontiers in Psychology*, **9**: 2451. <https://doi.org/10.3389/fpsyg.2018.02451>

Appendix

Table 1. Test sentences used in List A of the picture-selection task

Test block	Sentence number	
A	1	Afu epline to podhilato tu, eghrapse mia karta. 'After he cleaned his bicycle, he wrote a postcard.'
	2	Fuskose ena kokino baloni prin anapsi to keru. 'He blew a red balloon before he lit the candle.'
	3	Prin zoghrافisi mia ikona, endise tin kukla tu. 'Before he drew a picture, he dressed his doll.'
	4	Evale to maksilari sti thiki afu sidherose ena pukamiso. 'He put the pillow into the pillowcase after he ironed a shirt.'
	5	Afu skupise to xali, epline to aftokinito. 'After he vacuumed the carpet, he washed the car.'
	6	Prin ftiaksi mia xartini saita, epline ena potiri. 'Before he folded a paper airplane, he washed a glass.'
B	1	Tilikse tin afisa prin sikosi tis kaltses apo to patoma. 'He rolled up the poster before he lifted the socks from the floor.'
	2	Zoghrافise ena karavi afu potise ton kakto tu. 'He drew a ship after he watered his cactus.'
	3	Afu kolise mia afisa ston tixo, fuskose tin bala tu. 'After he hung up a poster, he pumped up his ball.'
	4	Prin etimasi ti valitsa tu, katharise mia banana. 'Before he packed his suitcase, he peeled a banana.'
	5	Efaje ena milo afu dhjavase to ghrama. 'He ate an apple after he read the letter.'
	6	Petakse to xarti sta skupidhja prin metrisi to xartziliki tu. 'He threw away the paper before he counted his pocket money.'
C	1	Ipie ena anapsiktiko afu ksetilikse tin kukla tis. 'She drank a soft drink after she unpacked her doll.'
	2	Afu jemise to paghouri tis, evghale mia fotoghrافia. 'After she filled her drinking bottle, she took a picture.'
	3	Prin foresi ta paputsia tis, eghrapse ena simioma. 'Before she put on her shoes, she wrote a note.'
	4	Afu anikse to parathiro, erapse ena kumbi. 'After she opened the window, she sewed on a button.'
	5	Ipie ena potiri ximo prin kseskonisi to ghrافio tis. 'She drank a glass of juice before she wiped her desk.'
	6	Prin ftiaksi ti vivliothiki tis, evale ton kathrefti tis sti thesi tu. 'Before she put up her shelf, she put up her mirror.'

Table 1. (Continued)

Test block	Sentence number	
D	1	Prin dhiplosi to pulover, eksise ti miti tu molivju. 'Before she folded the jumper, she sharpened the pencil.'
	2	Eklise to parathiro afu evale to piato sto trapezi. 'She closed the window after she put the plate on the table.'
	3	Anikse to radhjofo prin tighanisi ena avgho. 'She turned on the radio before she fried an egg.'
	4	Afu epsise to ghliko, evale ena vivlio sto rafi. 'After she baked the cake, she put a book on the shelf.'
	5	Eklise tin porta afu anakalipse enan isto araxnis. 'She shut the door after she spotted a cobweb.'
	6	Eskise ti lista me ta psonia prin vali ksana to ghala sto psijio. 'She tore up the shopping list before she put the milk back.'

Table 2. Test sentences used in List B of the picture-selection task

Test block	Sentence number	
A	1	Prin plini to podhilato tu, eghrapse mia karta. 'Before he cleaned his bicycle, he wrote a postcard.'
	2	Fuskose ena kokino baloni afu anapse to keru. 'He blew a red balloon after he lit the candle.'
	3	Afu zoghrfise mia ikona, endise tin kukla tu. 'After he drew a picture, he dressed his doll.'
	4	Evale to maksilari sti thiki prin sidherosi ena pukamiso. 'He put the pillow into the pillowcase before he ironed a shirt.'
	5	Prin skupisi to xali, epline to aftokineto. 'Before he vacuumed the carpet, he washed the car.'
	6	Afu eftiakse mia xartini saita, epline ena potiri. 'After he folded a paper airplane, he washed a glass.'

Table 2. (Continued)

Test block	Sentence number	
B	1	Tilikse tin afisa afu sikose tis kaltses apo to patoma. 'He rolled up the poster after he lifted the socks from the floor.'
	2	Zoghrafise ena karavi prin potisi ton kaktu tu. 'He drew a ship before he watered his cactus.'
	3	Prin kolisi mia afisa ston tixo, fuskose tin bala tu. 'Before he hung up a poster, he pumped up his ball.'
	4	Afu etimase ti valitsa tu, katharise mia banana. 'After he packed his suitcase, he peeled a banana.'
	5	Efaje ena milo prin dhjavasi to ghrama. 'He ate an apple before he read the letter.'
	6	Petakse to xarti sta skupidhja afu metrise to xartziliki tu. 'He threw away the paper after he counted his pocket money.'
C	1	Ipie ena anapsiktiko prin ksetiliki tin kukla tis. 'She drank a soft drink before she unpacked her doll.'
	2	Prin jemisi to paghuri tis, evghale mia fotoghrafia. 'Before she filled her drinking bottle, she took a picture.'
	3	Afu forese ta paputsia tis, eghrapse ena simioma. 'After she put on her shoes, she wrote a note.'
	4	Prin aniksi to parathiro, erapse ena kumbi. 'Before she opened the window, she sewed on a button.'
	5	Ipie ena potiri ximo afu kseskonise to ghrafio tis. 'She drank a glass of juice after she wiped her desk.'
	6	Afu eftiakse ti vivliothiki tis, evale ton kathrefti tis sti thesi tu. 'After she put up her shelf, she put up her mirror.'
D	1	Afu dhipse to pulover, eksise ti miti tu molivju. 'After she folded the jumper, she sharpened the pencil.'
	2	Eklise to parathiro prin vali to piato sto trapezi. 'She closed the window before she put the plate on the table.'
	3	Anikse to radhjofono afu tighanise ena avggho. 'She turned on the radio after she fried an egg.'
	4	Prin psisi to ghliko, evale ena vivlio sto rafi. 'Before she baked the cake, she put a book on the shelf.'
	5	Eklise tin porta prin anakalipsi enan isto araxnis. 'She shut the door before she spotted a cobweb.'
	6	Eskise ti lista me ta psonia afu evale ksana to ghala sto psijio. 'She tore up the shopping list after she put the milk back.'

Table 3. Output of the baseline model for the younger subgroup (age range: 6;1–8;7)

	Estimate	Standard error	z-value	Pr(> z)
(Intercept)	2.217	0.218	10.161	< 0.001
Iconicity	1.377	0.487	2.828	0.005
Connective	0.943	0.346	2.729	0.006
Iconicity*Connective	-1.635	0.688	-2.376	0.018

Model formula: `glmer(Result ~ Iconicity*Connective + (1 + Iconicity + Connective| IDNumber) + (1| Item), control = glmerControl(optimizer = "bobyqa"), family = binomial(link = "logit"), data)`.

Table 4. Output of the baseline model for the older subgroup (age range: 8;10–11;11)

	Estimate	Standard error	z-value	Pr(> z)
(Intercept)	4.022	0.632	6.360	< 0.001
Iconicity	-0.215	0.910	-0.236	0.813
Connective	0.308	0.699	0.441	0.659
Iconicity*Connective	-6.460	2.339	-2.761	0.006

Model formula: `glmer(Result ~ Iconicity*Connective + (1 + Iconicity + Connective| IDNumber) + (1| Item), control = glmerControl(optimizer = "bobyqa"), family = binomial(link = "logit"), data)`.

Cite this article: Makrodimitris, C., & Schulz, P. (2024). Comprehension of complex sentences containing temporal connectives: How children are led down the event-semantic kindergarten-path. *Journal of Child Language* 1–33, <https://doi.org/10.1017/S0305000924000205>