

Processing of positive-causal and negative-causal coherence relations in primary school children and adults: a test of the cumulative cognitive complexity approach in German*

JULIA KNOEPKE, TOBIAS RICHTER AND
MAJ-BRITT ISBERNER

University of Kassel

JOHANNES NAUMANN

Goethe University Frankfurt

YVONNE NEEB

German Institute for International Educational Research

AND

SABINE WEINERT

University of Bamberg

*(Received 29 October 2014 – Revised 17 July 2015 – Accepted 16 December 2015 –
First published online 5 February 2016)*

ABSTRACT

Establishing local coherence relations is central to text comprehension. Positive-causal coherence relations link a cause and its consequence, whereas negative-causal coherence relations add a contrastive meaning (negation) to the causal link. According to the cumulative cognitive complexity approach, negative-causal coherence relations are cognitively

* The research presented in this paper was supported by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF, grants 01 GJ 0985, 01 GJ 0986, 01 GJ 1206A, 01 GJ 1206B). We would like to thank David Nitz for programming assistance and the numerous student research assistants who were involved in collecting data for this project. Researchers who are interested in the material used in the visual and auditory semantic verification tasks are invited to send an e-mail to the first or the second author. Address for correspondence: Julia Knoepke, University of Kassel, Department of Psychology, Holländische Straße 36–38, 34127 Kassel, Germany. tel: +49 (0) 561 804-7166; fax: +49 (0) 561 804-3586; e-mail: julia.knoepke@uni-kassel.de

more complex than positive-causal ones. Therefore, they require greater cognitive effort during text comprehension and are acquired later in language development. The present cross-sectional study tested these predictions for German primary school children from Grades 1 to 4 and adults in reading and listening comprehension. Accuracy data in a semantic verification task support the predictions of the cumulative cognitive complexity approach. Negative-causal coherence relations are cognitively more demanding than positive-causal ones. Moreover, our findings indicate that children's comprehension of negative-causal coherence relations continues to develop throughout the course of primary school. Findings are discussed with respect to the generalizability of the cumulative cognitive complexity approach to German.

INTRODUCTION

Text comprehension may be regarded as building a coherent mental representation of the text and its contents. Accordingly, the ability to establish local and global coherence, that is, linking the content of adjacent and distant statements of a text, is central to text comprehension (Graesser Millis & Zwaan, 1997; Van Dijk & Kintsch, 1983). Two of the most intensely investigated coherence relations are positive-causal coherence relations (1), which link a cause and its consequence, and negative-causal coherence relations (2), which add a contrastive meaning or a negation to the causal link (Sanders, Spooren & Noordman, 1992):

- (1) Tom ate too many cherries. Therefore, he now has a stomachache.
- (2) Anna is ill. Nevertheless, she goes to school.

According to the cumulative cognitive complexity approach (Evers-Vermeul & Sanders, 2009; Spooren & Sanders, 2008), positive-causal and negative-causal relations differ in internal complexity, and therefore vary in the cognitive effort they require during text comprehension (a detailed explanation is provided in the next sections). Furthermore, the order of acquisition of coherence relations is also assumed to depend on their internal complexity. Less complex relations, such as positive-causal coherence relations, are assumed to be acquired earlier than negative-causal coherence relations. Studies investigating the processing of coherence relations on English- and Dutch-speaking children and adults report findings that are consistent with the cumulative cognitive complexity approach (e.g. Bloom, Lahey, Hood, Lifter & Fiess, 1980; Evers-Vermeul & Sanders, 2009; Goldman & Murray, 1992; Spooren & Sanders, 2008).

The present study addressed three research questions about the scope and explanatory power of the cumulative cognitive complexity approach. First,

most of the existing studies have been conducted with English- and Dutch-speaking participants, but little is known about the processing of these coherence relations in German. Second, the extent that the processing of coherence relations and its development run in parallel in listening and reading comprehension (as predicted, for example, by the simple view of reading: Gough & Tunmer, 1986; Hoover & Gough, 1990) is still unknown. To our knowledge, no study to date has employed the method of parallel tasks and materials to investigate the processing of positive-causal and negative-causal coherence relations in both written and spoken text comprehension. Finally, although several studies demonstrated early usage of positive-additive, positive-causal, and negative-additive coherence relations and connectives in preschoolers' speech production (e.g. Bloom *et al.*, 1980; Van Veen, Evers-Vermeul, Sanders & Van den Bergh, 2009, 2013), findings by Cain, Patson, and Andrews (2005) suggest that comprehension of connectives and coherence relations still develops during primary school. According to the cumulative cognitive complexity approach, this might hold, in particular, for the most complex type of coherence relations, namely negative-causal coherence relations, which rarely occur in preschoolers' spontaneous speech production.

Against this background, the present study (a) investigated the processing of positive-causal and negative-causal coherence relations in German primary school children and adults, (b) directly compared the processing of positive-causal and negative-causal coherence relations in written and spoken language, and (c) examined developmental changes as indicated by age differences in the processing of these coherence relations during the first four years of primary school. In the following section, we will outline the core assumptions of the cumulative cognitive complexity approach and discuss previous studies testing its assumptions. The cumulative cognitive complexity approach and questions raised by existing research motivated the predictions tested in the present study, whose rationale and findings will be presented afterwards.

Positive-causal and negative-causal coherence relations

Coherence relations are cognitive meaning relations that link the mental representations of adjacent and distant statements of a text by adding a distinct aspect of meaning that is "more than the sum of the parts" (Sanders *et al.*, 1992, p. 2; see also Sanders & Noordman, 2000). According to Sanders *et al.*, coherence relations can be characterized along four dimensions. The first dimension, termed BASIC OPERATION, refers to the strength with which two statements are connected. They can be connected either strongly (causal) or weakly (additive). The second dimension distinguishes between a semantic and a pragmatic SOURCE OF

COHERENCE, i.e. statements can be coherent with respect to their propositional content (semantic) or to the illocutionary meaning of the speech act (pragmatic). The third dimension refers to the ORDER OF THE TWO STATEMENTS for causal relations, which are the basic order (cause–consequence) or the non-basic order (consequence–cause). No such distinction is made for additive relations. The final dimension is POLARITY, characterizing coherence relations as either positive (containing no negation or contrast) or negative (adding a negative or contrastive meaning to the additivity or causality of the coherence relation).

One of the most intensely investigated types of coherence relations are causal links between statements. According to the taxonomy of Sanders *et al.* (1992), these coherence relations can be further distinguished along the polarity dimension into positive-causal and negative-causal relations. Positive-causal relations, such as expressed in (1), link two statements, one of which designates the cause of the effect described in the other one. In contrast, negative-causal relations, such as the relation expressed in (2), link two statements, one of which describes a state of affairs that would normally prevent the event described in the other statement from happening.

These coherence relations can be (but need not be) signalled by means of positive-causal connectives such as *because* or *therefore* or negative-causal connectives such as *although* or *nevertheless*. Connectives can function as ‘processing instructions’ (Kamalski, Sanders & Lentz, 2008, p. 324), which guide the reader’s expectations on how to integrate the upcoming statement into a coherent mental representation (Murray, 1995). Several studies underline the essential role connectives play in text comprehension. Connectives were found to decrease processing times, enhance text memory, and facilitate on-line integration of information and inference generation (e.g. Degand, Lefèvre & Bestgen, 1999; Maury & Teisseranc, 2005; Millis, Golding & Barker, 1995; Millis & Just, 1994; Sanders, Land & Mulder, 2007; Sanders & Noordman, 2000; Van Silfhout, Evers-Vermeul, Mak & Sanders, 2014), especially in readers with little prior knowledge of the text content (e.g. Kamalski *et al.*, 2008).

The cumulative cognitive complexity approach

Connectives and the coherence relations signalled by them are assumed to differ in the cognitive processing effort they require during text comprehension. One prominent account of the cognitive processing of coherence relations is the cumulative cognitive complexity approach (Evers-Vermeul & Sanders, 2009; Spooren & Sanders, 2008). Based on the taxonomy of Sanders *et al.* (1992), the cumulative cognitive complexity approach assumes that coherence relations differ in their internal cognitive complexity depending on the dimensions of their characterization. Along

the basic operation dimension, causal relations can be regarded as cognitively more complex than additive relations, because they involve a cause–consequence link in addition to the additive combination of the statements (Spooren & Sanders, 2008). In other words, causal relations are specified with respect to causality, whereas additive relations are underspecified with regard to this feature (Evers-Vermeul & Sanders, 2009). Likewise, along the dimension of polarity, negative relations are assumed to be cognitively more complex than positive relations, because they add a contrastive component to the relation (Spooren & Sanders, 2008). This characterization can also be extended to the combination of the two dimensions, resulting in a complexity hierarchy of coherence relations (Evers-Vermeul & Sanders, 2009; Spooren & Sanders, 2008): Positive-additive coherence relations should be cognitively the least complex, because they do not involve a causal link or a contrast. Positive-causal and negative-additive coherence relations are assumed to be more complex than positive-additive coherence relations, because they involve either a causal link (positive-causal relations) or a contrastive link (negative-additive relations). Finally, negative-causal relations should be the most complex, because they involve both a causal link and a contrast. Based on this hierarchy, Evers-Vermeul and Sanders (2009) and Spooren and Sanders (2008) assumed that the cognitive processing of coherence relations is more demanding for complex compared to less complex relations and that complex relations are acquired later than less complex ones. Accordingly, positive-additive coherence relations should be easier to process than positive-causal and negative-additive coherence relations, and these, in turn, should be easier to process than negative-causal coherence relations. In addition, positive-additive coherence relations should be acquired first and negative-causal coherence relations last, with positive-causal and negative-additive coherence relations between the two.

Several studies with English- and Dutch-speaking children and adults report findings that appear to support the cumulative cognitive complexity approach. In a self-paced reading study, Millis and Just (1994, Exp. 4) found that English readers provided less accurate responses to comprehension questions and responded more slowly when reading negative-causal sentence pairs linked by *although* compared to positive-causal sentence pairs linked by *because*. Goldman and Murray (1992) found that English students exhibited more difficulties in appropriately filling in negative connectives into blank slots between sentences compared to positive-additive and positive-causal connectives (but see also Sanders & Noordman, 2000, who found faster processing times for causal compared to additive relations).

With respect to the acquisition order of coherence relations, several studies (most of them focusing on language production) provided evidence that

English-speaking preschool children (e.g. Bloom *et al.*, 1980) and primary school children (e.g. Katz & Brent, 1968; Wing & Scholnick, 1981) use additive relations and connectives before causal ones, and positive relations and connectives before negative ones. Furthermore, their knowledge of coherence relations (negative coherence relations in particular) appears to be still developing during the primary school years (e.g. Cain & Nash, 2011; Katz & Brent, 1968; Wing & Scholnick, 1981). In Dutch, Evers-Vermeul and Sanders (2009) analyzed spontaneous language-production data of twelve children aged 1;5 to 5;6 to explicitly test the cumulative cognitive complexity approach. As expected, they found that additive relations are acquired before causal relations and that positive relations are acquired before negative relations. Spooren and Sanders (2008) obtained comparable results when investigating Dutch primary school children's production and comprehension of coherence relations.

Open question 1: cross-linguistic generalization of cumulative cognitive complexity to German

Although the studies discussed in the preceding section lend support to the cumulative cognitive complexity approach, open questions remain regarding its generalizability and its explanatory power. First, the cumulative cognitive complexity approach appears to explain processing differences and acquisition orders of coherence relations and connectives in English and Dutch speakers, but it still lacks cross-linguistic validation beyond these two languages. Sanders *et al.*'s (1992) taxonomy of cognitive coherence relations can be applied to German, and the semantics of the connectives signalling these coherence relations in German are similar to those in English and Dutch. Thus, the cumulative cognitive complexity approach is also expected to apply to German-speaking children. However, the existing studies with this population provide only surprisingly limited support for the approach.

Kail and Weissenborn (1984), for example, investigated the acquisition order of the substitutive and the contrastive use of the adversative *aber/sondern* 'but' in German eight- to ten-year-olds, and *mais* 'but' in French children of the same age. They found that the substitutive use of these connectives (3) is acquired earlier than the contrastive use (4).

- (3) Kathy's father didn't sell the old car but he gave it to Kathy.
- (4) Joe is an Indian but he didn't win the rodeo.

(examples from Kail & Weissenborn, 1984, pp. 147–148)

They argued that the contrastive use is acquired later, because the comprehension of contrastive links requires more complex inferences. To comprehend a sentence containing a contrastive adversative (4), one must

infer that Indians usually win rodeos, whereas no such inference is necessary to comprehend sentences containing a substitutive adversative. The substitutive use of adversatives is associated with a negative-additive relation, whereas the contrastive use is associated with a negative-causal relation. These findings by Kail and Weissenborn support the cumulative cognitive complexity approach, but only for effects within the category of negative coherence relations.

Additional but rather weak support for the cumulative cognitive complexity approach can be found in an eye-tracking study by Köhne and Demberg (2013), who investigated the function of positive-causal and negative-causal connectives to restrict possible discourse referents during on-line text processing in German adults. In Experiment 1, based on the visual world paradigm, participants listened to short spoken texts that expressed either a positive-causal or a negative-causal coherence relation signalled by their respective connectives. At the same time, participants viewed a visual scene with one target object matching the final word in the text and three distractor objects. Köhne and Demberg found that as soon as the connective occurred, participants focused significantly less often on those objects that could be excluded based on the meaning of the connective. However, they recognized the target objects more slowly in negative-causal compared to positive-causal sentence pairs and responded less accurately to comprehension questions following negative-causal sentence pairs. In line with the assumptions of the cumulative cognitive complexity approach, the authors suggested that processing negative-causal coherence relations is cognitively more demanding than processing positive-causal coherence relations. However, in contrast to Experiment 1, the results of a second eye-tracking experiment with written texts (Köhne & Demberg, Exp. 2) failed to support the cumulative cognitive complexity approach. No significant reading time differences were found, no increase in response accuracy to comprehension questions, and no decrease in reaction times for positive-causal compared to negative-causal coherence relations.

Finally, Dragon, Berendes, Weinert, Heppt, and Stanat (2015) found that German second- and third-graders showed considerable difficulties in comprehending negative-causal sentence pairs compared to positive-causal sentence pairs. They presented the children with coherent sentence pairs containing appropriately used positive-causal and negative-causal connectives, or with sentence pairs that had been made incoherent by exchanging positive-causal and negative-causal connectives. When asked to indicate whether the sentence pairs made sense, children systematically rejected coherent negative-causal sentence pairs and accepted incoherent negative-causal sentence pairs. Dragon *et al.* suggested that most children ignored the negative-causal connectives (or even all connectives) and

judged the sentence pairs only on the basis of their semantic and situational compatibility independent of the connective meaning. These findings indicate that German second- and third-graders have problems comprehending negative-causal connectives. This interpretation is in line with the assumption implied by the cumulative cognitive complexity approach that an adequate understanding of the cognitively complex negative-causal coherence relations and connectives is mastered comparatively late in language development. However, the findings by Dragon *et al.* are restricted to oral language processing and thus cannot be generalized to written language processing. Furthermore, only accuracy of the sensibility judgments was recorded but not response latency. Longer response latencies for negative-causal compared to positive-causal sentence pairs might lend further support to the assumptions of the cumulative cognitive complexity approach.

Open question 2: processing of coherence relations in reading vs. listening comprehension

The existing relevant literature also lacks evidence on the extent that coherence relations and their respective connectives are processed similarly in listening and reading comprehension, and on whether the ability to comprehend positive-causal and negative-causal coherence relations develops in parallel or at a different pace during elementary school years. According to the simple view of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), reading comprehension (R) can be defined as the product of two cognitive component skills, namely decoding (D) and general linguistic (or listening) comprehension (C). The formula $R = D \times C$ implies that the cognitive processes of text comprehension should be basically the same for listening and reading comprehension after written word forms have been decoded. On the basis of this account, similar findings should be expected for the processing of positive-causal and negative-causal coherence relations in written and spoken language processing. However, children might conceivably be able to comprehend complex coherence relations (such as negative-causal ones) at a later age when they read a text compared to when they listen to it. Young readers in the primary school years need to spend a considerable amount of cognitive resources on lower-level reading processes such as decoding (e.g. Knoepke & Richter, *in press*; Perfetti, 1985). Consequently, they might lack the cognitive resources required to comprehend cognitively complex coherence relations.

To our knowledge, no study to date has examined the comprehension of positive-causal and negative-causal coherence relations in both spoken and written language processing using the method of parallel tasks and

materials. The existing studies used either written (e.g. Cain & Nash, 2011; Cain *et al.*, 2005) or spoken text materials (e.g. Kail & Weissenborn, 1984; Katz & Brent, 1968). Few studies used both written and spoken text materials, but none directly compared performance in both modalities. For example, Köhne and Demberg (2013) used spoken text materials in their first, and written materials in their second, eye-tracking experiment. However, the tasks were different in the two experiments. In Experiment 1, participants were required to identify a target object in a visual scene, whereas in Experiment 2, they only read texts. Thus, the results in the two experiments are not directly comparable.

Open question 3: developmental changes in the processing of coherence relations during the primary school years

A third question that has been relatively neglected by previous research is how the processing of positive-causal and negative-causal coherence relations and connectives develops throughout the years of primary school. Many studies that have investigated the acquisition order of coherence relations and connectives focused on the first time that specific connectives occur in preschooler's speech production (e.g. Bloom *et al.*, 1980; Evers-Vermeul & Sanders, 2009; Van Veen *et al.*, 2009, 2013). Although they found adequate production of positive-additive, positive-causal, and negative-additive connectives and coherence relations, Cain *et al.* (2005) argued that comprehension of connectives and coherence relations might still be incomplete and developing during the years of primary school. According to the cumulative cognitive complexity approach, this might hold, in particular, for the most complex coherence relations, namely negative-causal coherence relations, which occur rarely in preschoolers' spontaneous speech production (e.g. Katz & Brent, 1968; Spooren & Sanders, 2008).

THE PRESENT STUDY

The present study pursued three related aims. The first aim was to replicate the findings from previous studies that the processing of negative-causal coherence relations requires more cognitive effort than the processing of positive-causal coherence relations in German primary school children and adults. The second aim was to directly compare via parallel tasks and experimental materials the processing of these coherence relations in written and spoken language comprehension. The third aim was to investigate age-related differences in the processing of positive-causal and negative-causal coherence relations during the first four years of primary school.

To this end, we used a computerized semantic verification task in a visual and auditory version that was similar to the task used by Cain and Nash

(2011, Exp. 2 and 3; see also the paradigm used in Murray, 1997, Exp. 2). German primary school children and adults were presented with spoken (all grade levels and adults) and written sentence pairs (Grades 3 and 4 and adults) that contained positive-causal or negative-causal coherence relations signalled by either an appropriate connective resulting in a coherent sentence pair such as (5) or an inappropriate connective resulting in an incoherent sentence pair such as (6).

- (5) Sandra war nicht müde. Trotzdem ging sie ins Bett.
 ‘Sandra was not tired. Nevertheless, she went to bed.’
- (6) *Das Wetter war gut. Trotzdem setzte Laura eine Sonnenbrille auf.
 *‘The weather was good. Nevertheless, Laura put on her sunglasses.’

The participants’ task was to judge the coherence of the sentence pairs. Mastery of this task requires the comprehension of the coherence relation between the two statements, and mastery of negative-causal sentence pairs requires the comprehension of the connective. Response latencies and accuracy were recorded. This paradigm has already been shown to be sensitive to differences in the cognitive processing of coherence relations and to detect developmental changes in children’s ability to establish local coherence (Cain & Nash, 2011, Exp. 2 and 3).

According to the cumulative cognitive complexity approach, we expected that children at all grade levels and adults need more time and make more errors when judging the coherence of the cognitively more complex coherent and incoherent negative-causal sentence pairs compared to the less complex positive-causal sentence pairs.

Two different patterns of outcomes seem possible from the analysis of differences between reading and listening comprehension. The simple view of reading implies that processing of positive-causal and negative-causal coherence relations should not differ between written and auditory text presentation and that the ability to comprehend these relations should develop at the same pace in both modalities. In contrast, theories and findings emphasizing that reading comprehension may be hampered by inefficient word-level processes (e.g. Perfetti, 1985) suggest that younger readers might be able to comprehend negative-causal relations presented auditorily while struggling with the same relations when they are presented in writing. Thus, younger readers might have more difficulties processing the cognitively more complex negative-causal coherence relations in the visual version of the task compared to the auditory version.

Furthermore, based on Cain *et al.*’s (2005) suggestion that children’s comprehension of connectives and coherence relations might still be incomplete and developing during the primary school years, we expected to find age-related differences in the comprehension of both positive-causal

and negative-causal coherence relations in primary school children. Because of their higher cognitive complexity, we particularly expected developmental trends in the processing of negative-causal coherence relations to lag behind developmental trends in the processing of positive-causal coherence relations.

Finally, Cain and Nash (2011, Exp. 3) found longer reading times for sentences preceded by an inappropriate connective (incoherent sentence pair) compared to sentences preceded by an appropriate connective (coherent sentence pair). They suggested that the longer processing times for sentences following an inappropriate connective reflect a laborious attempt to repair the current mental model of the text when inconsistent information cannot be integrated. Based on their findings, we also expected to find shorter response latencies for coherent (containing an appropriate connective) compared to incoherent sentence pairs (containing an inappropriate connective) in our study.

METHOD

Participants

Participants were 422 German primary school children from Grades 1 to 4, recruited from ten schools (32 classes) in the Frankfurt am Main area, Cologne, and Kassel, and 78 undergraduate students of psychology from the University of Kassel (Germany), who received course credit for their participation.

Children. Two hundred and seventy-seven primary school children participated in the auditory version of the task and 269 children in the visual version of the task. Of these children, 124 from Grades 3 and 4 took part in both the auditory and the visual version. Because of limitations on the time available for testing, 96 children failed to complete the auditory version of the task and 107 children could not complete the visual version of the task. By mischance, the untimely aborting of the programme led to errors preventing the adequate saving of their data. However, the errors did not result in a systematic exclusion of data. In addition, the data of 38 (auditory version of the task) and 37 (visual version of the task) non-native German speakers were excluded from the analysis. Overall, 143 native German speakers (66 boys, 70 girls, 7 missing gender information) from Grades 1 to 4 completed the auditory version of the task, and 125 (72 boys, 53 girls) completed the visual version of the task. Sample characteristics are provided in Table 1. Sociodemographic data were collected via a questionnaire completed by the parents. This questionnaire was supplemented by a teacher questionnaire that was used in the event of missing or incomplete parents' questionnaires. Children only participated in the study when parents had given their written consent.

TABLE 1. *Sample characteristics in the auditory and the visual semantic verification task*

	Auditory version				Visual version			
	N	Age			N	Age		
		M(SD)	Min	Max		M(SD)	Min	Max
Grade 1	35	7;5 (0;5)	6;9	8;3	–	–	–	–
Grade 2	18	8;2 (0;5)	7;3	8;11	–	–	–	–
Grade 3	31	9;5 (0;5)	8;5	10;7	57	9;5 (0;5)	8;5	10;6
Grade 4	59	10;5 (0;4)	9;10	11;4	68	10;4 (0;5)	9;0	11;3
Adults	60	23;0 (5;3)	18;5	41;8	60	23;0 (5;3)	18;5	41;8

Adults. All 78 psychology students provided their written consent and completed both the auditory and the visual version of the task. Sociodemographic data were collected via a questionnaire. Seven participants (9.0%) were removed because of missing sociodemographic data and an additional 11 non-native German speakers (14.1%) were removed. The data from the remaining 60 (47 female, 13 male) participants were analyzed.

Materials

We used the visual semantic verification task from the computerized German-speaking test battery ProDi-L (ProDi-L: Prozessbezogene Diagnostik des Leseverstehens bei Grundschulkindern [Process-based assessment of reading skills in primary school children], Richter, Naumann, Isberner, Neeb & Knoepke, *in press*) and its parallel auditory counterpart ProDi-H.

Visual semantic verification task. The visual semantic verification task of ProDi-L contained seventy written experimental sentence pairs and two additional practice sentence pairs of varying length (length in characters for the sentence pair: $M = 58.53$, $SD = 10.68$, $Min = 35$, $Max = 96$; length in characters for the second sentence: $M = 30.57$, $SD = 6.90$, $Min = 18$, $Max = 52$). Half of the sentence pairs were coherent, the other half incoherent. The participants' task was to judge the coherence of the sentence pairs ("Does the second sentence go with the first one?") as fast and as accurately as possible by pressing a green button for coherent sentence pairs ("yes") and a red button for incoherent sentence pairs ("no"). Of the total of seventy sentence pairs, thirty (29 for the adults) were chosen as critical items according to the aims of the present study. These critical sentence pairs contained a positive-causal or a negative-causal coherence relation signalled by either an appropriate

connective (coherent sentence pair) or an inappropriate connective (incoherent sentence pair): six coherent positive-causal sentence pairs (7), six (5 for the adults) incoherent positive-causal sentence pairs (8), twelve coherent negative-causal sentence pairs (5) (repeated here as (9)), and six incoherent negative-causal sentence pairs (6) (repeated here as (10)).

- (7) Lena war zu lange in der Sonne. Darum bekam sie einen Sonnenbrand.
 ‘Lena stayed in the sun for too long. Therefore, she got a sunburn.’
- (8) *Roland hat verschlafen. Darum kommt er pünktlich zur Schule.
 *‘Roland overslept. Therefore, he arrived at school on time.’
- (9) Sandra war nicht müde. Trotzdem ging sie ins Bett.
 ‘Sandra was not tired. Nevertheless, she went to bed.’
- (10) *Das Wetter war gut. Trotzdem setzte Laura eine Sonnenbrille auf.
 *‘The weather was good. Nevertheless, Laura put on her sunglasses.’

Note that the incoherence in the incoherent negative-causal sentence pairs as illustrated in (10) was created by linking two actually compatible sentences with a negative-causal connective.

In addition to these critical items, the semantic verification task contained forty (41 for adults) sentence pairs, which were coherent and positive-additive (11), coherent and positive-causal without a connective (12), or incoherent sentence pairs containing semantic violations of varying degree (13–15).

- (11) Felix baut ein Haus mit Legosteinen. Dabei hört er Musik.
 ‘Felix is building a house of Lego bricks. While doing this, he is listening to music.’
- (12) Klara möchte ein Buch lesen. Sie geht in die Bücherei.
 ‘Klara wants to read a book. She goes to the library.’
- (13) *Katie geht gerne Schlittschuhlaufen. Dabei trägt er blaue Handschuhe.
 *‘Katie loves to go skating. While doing this, he wears blue gloves.’
- (14) *Sina isst ein Stück Kuchen. Trotzdem fallen die Blätter im Herbst.
 *‘Sina is eating a piece of cake. Nevertheless, the leaves fall in autumn.’
- (15) *Die Löwen brüllen laut. Dann wiegt es weniger.
 *‘The lions roar loudly. Then it weighs less.’

The critical sentence pairs contained the positive-causal connectives *deshalb*, *daher*, *darum*, or *denn* ‘therefore’, and the negative-causal connective *trotzdem* ‘nevertheless’. Based on their frequencies in the childLex database (Schroeder, Würzner, Heister, Geyken & Kliegl, 2015), primary school children are likely to be familiar with these connectives in

Grades 1 and 2 (frequency per 1 million words for *deshalb*: 300.5; *daher*: 34.8; *darum*: 99.2; *denn*: 889.4; *trotzdem*: 257.6) and Grades 3 and 4 (*deshalb*: 309.1; *daher*: 45.0; *darum*: 150.8; *denn*: 704.9; *trotzdem*: 303.1).

Some of the filler sentence pairs also contained positive-causal and negative-causal connectives. However, these items differed from the critical sentence pairs, because little or no referential overlap exists between the first and the second statements, as illustrated in (14). Most of them could easily be rejected as incoherent based on the lack of semantic overlap. In contrast, participants needed to consider the local coherence relations to verify the critical sentence pairs and consider the connective to make the correct judgement on negative-causal sentence pairs. Stimulus material and instructions were the same for children and adults.

Auditory semantic verification task. The items of the auditory semantic verification task of ProDi-H differed from the items in the visual version but were strictly parallel with respect to construction criteria and length (character length of the whole sentence pair: $M = 58.44$, $SD = 10.59$, $Min = 31$, $Max = 99$). Half of the sentence pairs were recorded by a professional male speaker (a radio journalist), the other half by a trained female speaker (an actress).

Procedure

Both semantic verification tasks were administered in the context of a larger cross-sectional study. The study investigated processes of listening and reading comprehension with various measures on the word, the sentence, and the text level (Richter, Isberner, Naumann & Kutzner, 2012; Richter, Isberner, Naumann & Neeb, 2013). Children were tested together in classrooms of the participating schools at the end of the school year. Adults were tested in groups of one to six in a laboratory at the University of Kassel Psychology Department. Tasks embedded in a story of an extraterrestrial named Reli who came to earth to learn the earthlings' language were presented on notebook computers. Reli walked the participants through the tasks in a manner suitable for children. He introduced the instructions via headphones and gave feedback during the two practice trials preceding the presentation of each task. When participants provided an incorrect answer to one of the practice trials, the practice trials were repeated until the response was correct. In the visual version of the task, the first sentence of the sentence pair appeared in the middle of the notebook screen. When the participants finished reading, they pressed the space bar and then the second sentence appeared below the first one (see Figure 1). In the auditory version of the task, participants were presented with the spoken sentence pairs via headphones. Items were presented in a random order. Response accuracy

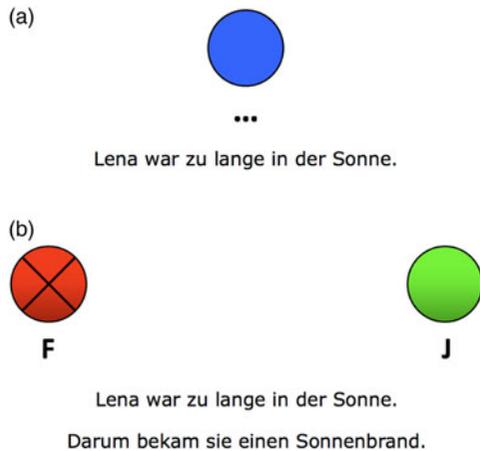


Fig. 1. Example of the presentation of sentence pairs in the visual version of the semantic verification task. Initially, the first sentence appeared on the notebook screen (a), and after pressing the spacebar, the second sentence of the sentence pair appeared below the first one (b). The sentence pair remained on the screen until one of the response keys was pressed.

and log-transformed response latencies from presentation onset (of the second sentence in the visual version and of the whole sentence pair in the auditory version) to the pressing of the response button were recorded. The visual semantic verification task was presented first for children and adults who participated in both the visual and the auditory version.

RESULTS

Only critical items were included in the analysis. Responses that were unusually slow or fast (3 *SD* or more below the item-specific mean and 2 *SD* or more below or above the person-specific mean after standardizing each item by its item-specific mean) were excluded from the analyses, which resulted in a loss of 4.8% of the children's data in the visual version of the task and 6.8% in the auditory version of the task for the critical sentence pairs. In the sample of adults, 4.8% of the data were excluded from the analysis in the visual version of the task and 5.2% in the auditory version. Log-transformed response latencies were analyzed using Linear Mixed Models (LMM; Baayen, Davidson & Bates, 2008) with crossed random effects for items nested within participants and participants nested within items. For the accuracy data, we estimated Generalized Linear Mixed Models (GLMM) with a logit link function (Dixon, 2008). Given that the presentation order of the items was randomized between participants, the results of the mixed-models analyses were assumed to be unbiased despite the small proportion of missing data.

All models were estimated and tested with the software packages *lme4* (Bates, Maechler, Bolker, Walker, Christensen & Sigmann, 2014) and *lmerTest* for R (Kuznetsova, Brockhoff & Christensen, 2014). All significance tests were based on a Type I error probability of .05. Separate models were estimated for the auditory and the visual version of the task and for accuracy and log-transformed response latencies as dependent variables. In addition, separate models were estimated for children and adults to model linear trends in the processing of coherence relations throughout primary school years by including grade level for children as a discrete but interval-scaled predictor. Because participants and items were sampled from a larger population, intercepts for persons and items were allowed to vary randomly. To control for main effects of item length and position of item presentation throughout the course of the experiment, the number of characters and the position of each item for each participant were included as grand mean-centred predictor variables. Main effects (fixed effects) were estimated for variables of theoretical interest. Polarity of the coherence relation and coherence were included as contrast-coded predictor variables (polarity: 1 = positive-causal, -1 = negative-causal; coherence: 1 = coherent, -1 = incoherent). Grade level was included as grand mean-centred predictor variable to model linear developmental trends from Grades 1 to 4 in children. The age of adult participants was included as grand mean-centred predictor variable to control for a possible increase in response latency because of increasing age. In addition, interaction effects were estimated for polarity and coherence and for grade level and age with polarity and coherence to account for possible developmental differences in the processing of the various types of sentence pairs, although we expected to find no interaction effects for adults' age with polarity and coherence. Finally, log-transformed response latency was included as grand mean-centred predictor variable in the GLMMs and accuracy as dummy-coded predictor variable (0 = correct, 1 = incorrect) in the LMMs to control for a potential speed-accuracy trade-off. In this way, the intercepts (and all main and interaction effects) were estimated for correct responses in the LMMs and for mean response latency in the GLMMs. The parameter estimates for the fixed and random effects are provided in Table 2 for the GLMMs and in Table 3 for the LMMs. In the following sections, we focus on the variables of theoretical interest, namely polarity, coherence, grade level (for children), and their interactions.

Visual semantic verification task

Response accuracy (children). The GLMM for the accuracy data revealed significant main effects of polarity ($\beta = 1.05$; $z = 10.64$; $p < .001$), coherence

TABLE 2. *Fixed effects and variance components in the GLMMs for response accuracy in children and adults*

Parameter	Children		Adults	
	Visual task β (SE)	Auditory task β (SE)	Visual task β (SE)	Auditory task β (SE)
	Fixed effects			
Intercept	1.52 (0.13)*	1.38 (0.19)*	3.90 (0.35)*	3.99 (0.36)*
Characters ^a	0.01 (0.02)	-0.00 (0.02)	0.04 (0.03)	0.03 (0.03)
Polarity ^b	1.05 (0.10)*	1.36 (0.18)*	1.08 (0.27)*	0.84 (0.29)*
Coherence ^b	0.34 (0.10)*	-0.25 (0.18)	0.10 (0.27)	0.28 (0.29)
Grade level ^a /Age ^a	0.52 (0.20)*	0.41 (0.07)*	0.09 (0.06)	0.10 (0.06)
Response latency ^a	0.40 (0.08)*	0.65 (0.19)*	0.03 (0.53)	-1.62 (1.45)
Polarity \times Coherence	0.28 (0.10)*	-0.17 (0.18)	-0.07 (0.26)	0.40 (0.28)
Polarity \times Grade level/Age	-0.21 (0.11)	-0.09 (0.04)	0.01 (0.05)	0.08 (0.05)
Coherence \times Grade level/ Age	-0.14 (0.11)	-0.00 (0.04)	-0.09 (0.05)	-0.00 (0.05)
Polarity \times Response latency	0.39 (0.07)*	0.53 (0.17)*	1.00 (0.51)	1.00 (1.32)
Coherence \times Response latency	-0.26 (0.07)*	-1.04 (0.17)*	-0.08 (0.51)	-0.55 (1.28)
Polarity \times Coherence \times Grade level/Age	0.04 (0.11)	-0.00 (0.04)	-0.09 (0.05)	-0.02 (0.05)
Polarity \times Coherence \times Response latency	0.18 (0.07)*	-0.38 (0.17)*	0.31 (0.51)	-0.24 (1.29)
	Variance components			
Subjects	0.86 (0.93)	0.59 (0.77)	1.31 (1.14)	1.15 (1.07)
Items	0.15 (0.39)	0.72 (0.85)	0.46 (0.68)	0.83 (0.91)

NOTES: ^a grand mean-centred; ^b contrast-coded; Characters: number of characters for the sentence pair (auditory version) and the second sentence (visual version); Polarity: positive = 1, negative = -1; Coherence: coherent = 1, incoherent = -1; Grade level/Age: Grade level (children) and age (adults); * $p < .05$ (two-tailed).

($\beta = 0.34$; $z = 3.52$; $p < .001$), log-transformed response latencies ($\beta = 0.40$; $z = 4.87$; $p < .001$), and grade level ($\beta = 0.52$; $z = 2.59$; $p = .01$).

An overall increase was found in response accuracy from Grades 3 to 4. Furthermore, the model revealed three significant two-way interactions of polarity and coherence ($\beta = 0.28$; $z = 2.84$; $p = .005$), polarity and log-transformed response latency ($\beta = 0.39$; $z = 5.83$; $p < .001$), and coherence and log-transformed response latency ($\beta = -0.26$; $z = -3.96$; $p < .001$). However, these interactions were qualified by a significant three-way interaction of polarity, coherence, and log-transformed response latency ($\beta = 0.18$; $z = 2.78$; $p = .01$). The three-way interaction is depicted in Figure 2a.

Figure 2a shows that there was a positive relationship between response latency and accuracy for positive-causal coherent, positive-causal incoherent, and negative-causal incoherent sentence pairs. In contrast, longer response

TABLE 3. *Fixed effects and variance components in the LMMs for response latency in children and adults*

Parameter	Children		Adults	
	Visual task β (SE)	Auditory task β (SE)	Visual task β (SE)	Auditory task β (SE)
Fixed Effects				
Intercept	7.749 (0.05)*	8.54 (0.03)*	7.44 (0.03)*	8.39 (0.02)*
Characters ^a	0.01 (0.00)*	0.01 (0.00)*	0.01 (0.00)*	0.01 (0.00)*
Polarity ^b	-0.03 (0.02)	-0.05 (0.02)*	-0.08 (0.02)*	-0.03 (0.02)
Coherence ^b	-0.09 (0.02)*	-0.04 (0.02)*	-0.08 (0.02)*	-0.01 (0.02)
Grade level ^a /Age ^a	-0.08 (0.09)	-0.02 (0.02)	0.01 (0.00)*	0.00 (0.00)*
Response accuracy ^c	-0.08 (0.03)*	-0.03 (0.01)	-0.00 (0.04)	0.01 (0.01)
Polarity \times Coherence	0.01 (0.02)	0.01 (0.02)	0.00 (0.02)	0.00 (0.02)
Polarity \times Grade level/Age	-0.01 (0.02)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Coherence \times Grade level/Age	0.01 (0.02)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
Polarity \times Response accuracy	-0.04 (0.03)	0.02 (0.01)	-0.09 (0.04)*	-0.00 (0.01)
Coherence \times Response accuracy	0.07 (0.03)*	0.05 (0.01)*	-0.03 (0.04)	0.00 (0.01)
Polarity \times Coherence \times Grade level/Age	-0.01 (0.02)	-0.01 (0.00)*	0.00 (0.00)	-0.00 (0.00)
Polarity \times Coherence \times Response accuracy	-0.07 (0.03)*	0.01 (0.01)	-0.05 (0.04)	0.01 (0.01)
Variance components				
Subjects	0.25 (0.50)	0.06 (0.25)	0.03 (0.17)	0.00 (0.03)
Items	0.01 (0.09)	0.01 (0.08)	0.01 (0.11)	0.01 (0.08)

NOTES: ^a grand mean-centred; ^b contrast-coded; ^c dummy-coded; Characters: number of characters for the sentence pair (auditory version) and the second sentence (visual version); Polarity: positive = 1, negative = -1; Coherence: coherent = 1, incoherent = -1; Grade level/Age: grade level (children) and age (adults); Response accuracy: correct response = 0, incorrect response = 1; * $p < .05$ (two-tailed).

latencies were associated with less accurate responses in negative-causal coherent sentence pairs.

Response latency (children). The LMM for response latency revealed significant main effects of coherence ($\beta = -0.09$; $t(112) = -4.10$; $p < .001$) and accuracy ($\beta = -0.08$; $t(112) = -2.62$; $p = .01$). Furthermore, the significant two-way interaction of coherence and accuracy ($\beta = 0.07$; $t(112) = 2.28$; $p = .02$) and the three-way interaction of polarity, coherence, and accuracy ($\beta = -0.07$; $t(112) = -2.31$; $p = .02$) indicate a positive relationship between response latency and accuracy (i.e. a speed-accuracy trade-off) for all sentence pairs except for negative-causal coherent sentence pairs

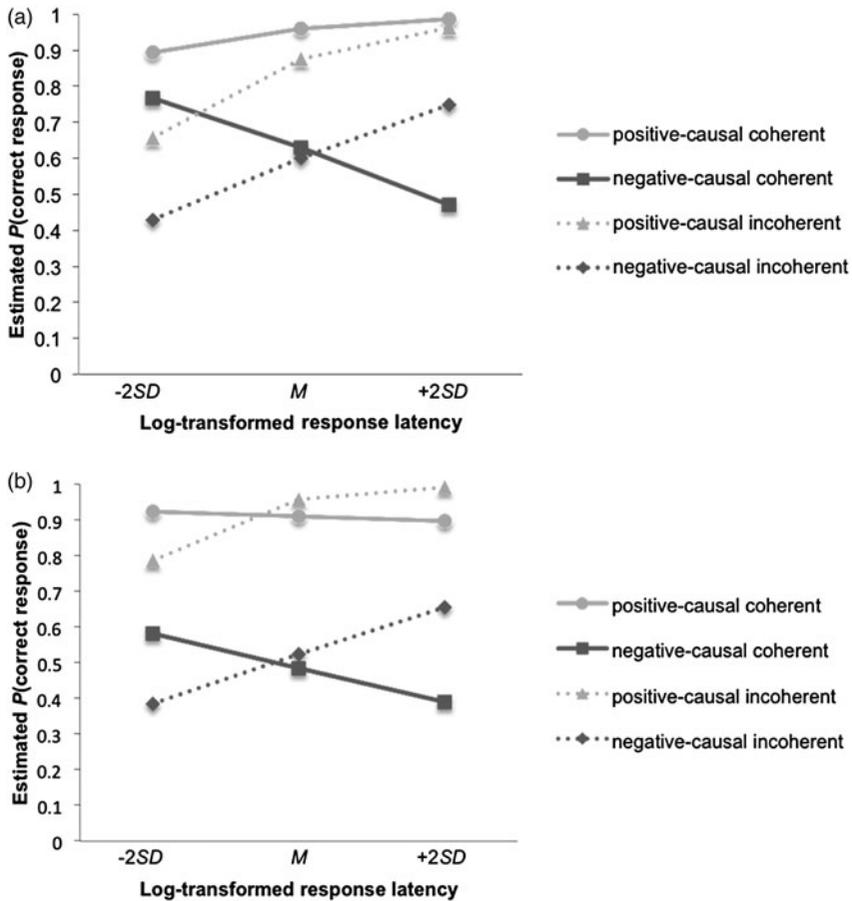


Fig. 2. Interaction of log-transformed response latency, polarity, and coherence with model-based estimated probability of correct responses as the dependent variable in the visual semantic verification task (top) and in the auditory semantic verification task (bottom) for primary school children.

(Figure 3). Inaccurate responses of the latter pairs were associated with longer response latencies than accurate responses.

Response accuracy (adults). The GLMM for response accuracy revealed a significant main effect of polarity ($\beta = 1.08$; $z = 3.98$; $p < .001$), indicating that adults responded with higher accuracy to positive-causal sentence pairs (99%) compared to negative-causal sentence pairs (93–95%). None of the other main effects or interaction effects were significant.

Response latency (adults). The LMM for response latency revealed significant main effects of polarity ($\beta = -0.08$; $t(47) = -3.54$; $p = .002$),

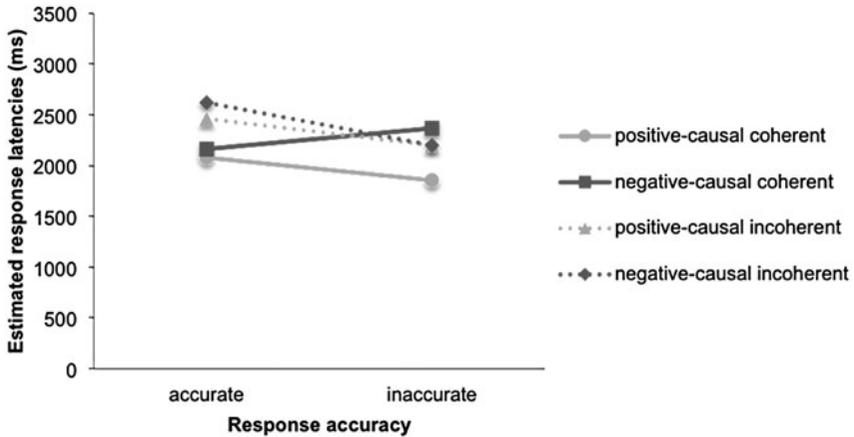


Fig. 3. Interaction of response accuracy, polarity, and coherence with model-based estimated response latency as the dependent variable in the visual semantic verification task for primary school children.

coherence ($\beta = -0.08$; $t(47) = -3.44$; $p = .002$), and age ($\beta = 0.01$; $t(47) = 2.85$; $p = .01$). Adults provided faster responses for coherent compared to incoherent sentence pairs, and younger participants provided slightly faster responses than older participants. Furthermore, a significant two-way interaction of polarity and response accuracy was found ($\beta = -0.09$; $t(47) = -2.21$; $p = .03$), indicating a speed-accuracy trade-off for positive-causal sentence pairs, whereas longer response latencies were associated with less accurate responses for negative-causal sentence pairs.

The results of the visual semantic verification task can be summarized as follows: the analysis revealed an overall increase in children's response accuracy from Grades 3 to 4. As predicted by the cumulative cognitive complexity approach, children provided more accurate responses to coherent and incoherent positive-causal sentence pairs compared to incoherent negative-causal sentence pairs. They also responded less accurately to coherent negative-causal sentence pairs than to both types of positive-causal sentence pairs but only when response latencies were long. Somewhat unexpectedly, response accuracy was negatively associated with response latency for negative-causal coherent sentence pairs. A similar three-way interaction between polarity, coherence, and accuracy was found for response latency as dependent variable. In contrast to response accuracy, children's response latencies cannot be clearly interpreted as being consistent with the cumulative cognitive complexity approach. Only inaccurate responses to coherent negative-causal sentence pairs were slower than coherent positive-causal sentence pairs.

Adults provided less accurate and slower responses for negative-causal compared to positive-causal sentence pairs. This pattern of results is consistent with the cumulative cognitive complexity approach. Furthermore, there was a speed–accuracy trade-off for positive-causal sentence pairs, whereas the reverse pattern was observed for negative-causal sentence pairs. Finally, response speed decreased with increasing age.

As expected, adults also needed more time to respond to incoherent compared to coherent sentence pairs. For children, positive-causal incoherent sentence pairs were associated with longer response latencies than positive-causal coherent sentence pairs, whereas negative-causal incoherent sentence pairs were associated with longer response latencies than negative-causal sentence pairs only when the responses were accurate.

Auditory semantic verification task

Response accuracy (children). The GLMM for response accuracy in the auditory semantic verification task revealed significant main effects of polarity ($\beta = 1.36$; $z = 7.60$; $p < .001$), grade level ($\beta = 0.41$; $z = 5.90$; $p < .001$), and log-transformed response latency ($\beta = 0.65$; $z = 3.49$; $p < .001$). The significant effect of grade level indicates that older children provided more accurate responses than younger children.

Furthermore, the analysis revealed three significant interaction effects of polarity and log-transformed response latency ($\beta = 0.53$; $z = 3.14$; $p = .002$), coherence and log-transformed response latency ($\beta = -1.04$; $z = -6.23$; $p < .001$), and a three-way interaction of polarity, coherence, and log-transformed response latency ($\beta = -0.38$; $z = -2.28$; $p = .02$). The three-way interaction of polarity, coherence, and log-transformed response latency is depicted in [Figure 2b](#). A positive relationship between response latency and accuracy (i.e. a speed–accuracy trade-off) was obtained only for incoherent sentence pairs. In contrast, a more thorough processing resulted in no response accuracy increase for the coherent sentence pairs. As in the visual semantic verification task, slower responses were associated with lower accuracy for negative-causal coherent sentence pairs. [Figure 2b](#) illustrates that children responded with overall higher accuracy to positive-causal sentence pairs compared to negative-causal sentence pairs. As depicted in [Figure 4](#), children's responses to negative-causal sentence pairs were strikingly less accurate than their responses to positive-causal sentence pairs at all grade levels. At Grade 1, their response accuracy to negative-causal sentence pairs was far below chance level (28–31% correct). In other words, they systematically rejected coherent negative-causal sentence pairs as incoherent and accepted incoherent negative-causal sentence pairs as coherent. At Grade 2, they responded to negative-causal sentence pairs slightly below chance level (39–43% correct), at Grade 3 at chance level

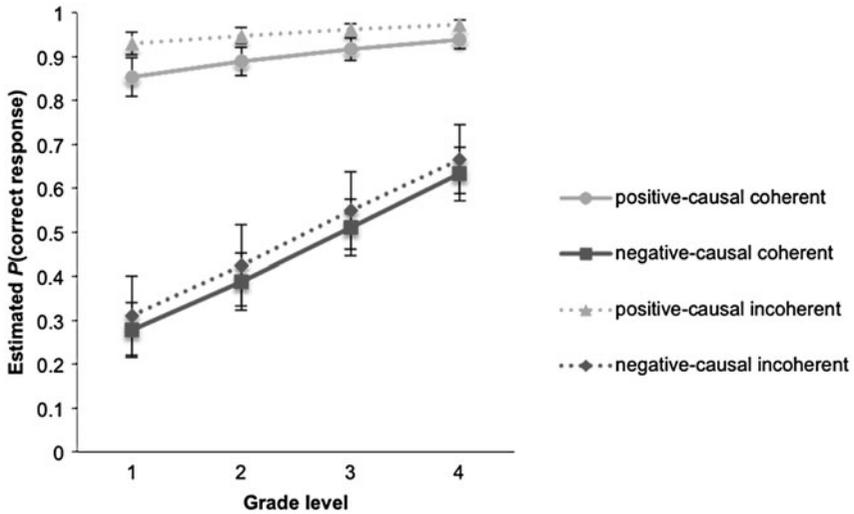


Fig. 4. Model-based estimated probability of correct responses with standard error in the auditory semantic verification task by grade level, polarity, and coherence in primary school children.

(51–55% correct), and only at Grade 4 at a level above chance (63–67% correct). Nevertheless, their responses to these items were still remarkably less accurate compared to positive-causal sentence pairs.

Response latency (children). The LMM for response latency revealed significant main effects of polarity ($\beta = -0.05$; $t(130) = -3.37$; $p = .002$) and coherence ($\beta = -0.04$; $t(130) = -2.29$; $p = .03$). Furthermore, the two-way interaction of coherence and accuracy ($\beta = 0.05$; $t(130) = 3.63$; $p < .001$) was significant, indicating a speed–accuracy trade-off only for incoherent sentence pairs. Finally, the three-way interaction of polarity, coherence, and grade level ($\beta = -0.01$; $t(130) = -3.08$; $p = .002$) reached significance. Older children provided faster responses than younger children to coherent and incoherent positive-causal sentence pairs and to incoherent negative-causal sentence pairs. In contrast, no increase was found in the response speed of coherent negative-causal sentence pairs from Grades 1 to 4 (Figure 5).

Response accuracy (adults). The GLMM for response accuracy revealed a significant main effect of polarity ($\beta = 0.84$; $z = 2.89$; $p = .004$) indicating higher response accuracy for positive-causal compared to negative-causal sentence pairs. None of the other main or interaction effects reached significance.

Response latency (adults). The LMM for response latency revealed a significant main effect of age. Older participants responded slightly more

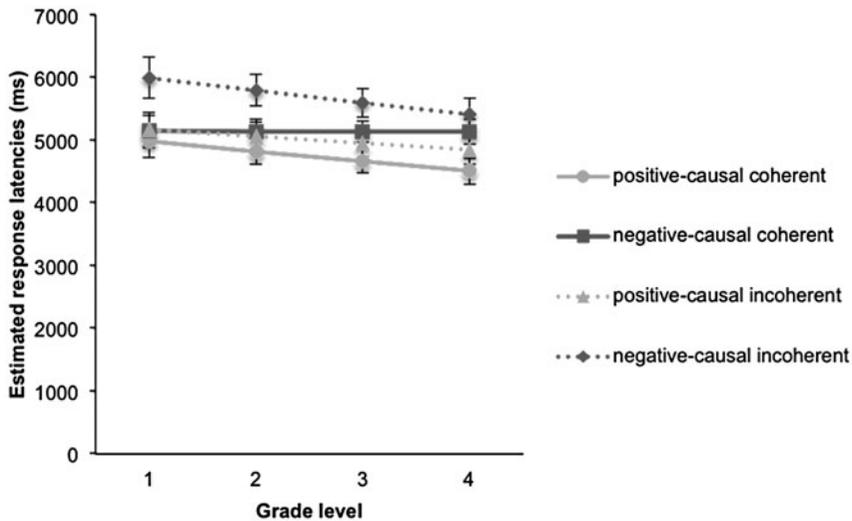


Fig. 5. Model-based estimated log-transformed response latency with standard error in the auditory semantic verification task by grade level, polarity, and coherence in primary school children.

slowly compared to younger participants ($\beta = 0.004$; $t(47) = 4.24$; $p < .001$). None of the other main or interaction effects reached significance.

The results of the auditory semantic verification task can be summarized as follows: the analysis revealed an overall increase of children's response accuracy from Grades 1 to 4. In line with the cumulative cognitive complexity approach, children's responses to positive-causal sentence pairs were more accurate than to negative-causal sentence pairs at all grade levels (Figure 2b). Their response accuracy in negative-causal sentence pairs was far below chance level at Grade 1, slightly below chance level at Grade 2, and at chance level at Grade 3. Only fourth-grade pupils responded with above-chance level accuracy to negative-causal sentence pairs (Figure 4). Moreover, the three-way interaction of polarity, coherence, and response latency indicated a speed-accuracy trade-off for the incoherent sentence pairs and the reverse pattern, i.e. longer latencies associated with less accurate responses, for the coherent negative-causal sentence pairs. Children needed more time to respond to negative-causal incoherent compared to positive-causal incoherent sentence pairs. The pattern was different for coherent sentence pairs. Negative-causal sentence pairs were processed slower than positive-causal sentence pairs only in the higher grade levels. This result was due to the fact that children's response speed increased with increasing grade level for all types of sentence pairs except for coherent negative-causal sentence pairs (see Figure 5).

Adults responded more accurately to positive-causal sentence pairs than to negative-causal sentence pairs, as expected by the cumulative cognitive complexity approach. However, no effect of polarity was found in the response latency data. Finally, response speed decreased with increasing age.

Children but not adults responded faster to coherent than to incoherent sentence pairs, but this effect held only for accurate responses. The same interaction described differently, children responded more accurately to coherent than to incoherent sentence pairs but only when response latency was low. For long response latencies, the pattern was reversed.

DISCUSSION

The present study pursued three related aims. The first aim was to replicate the finding that the processing of negative-causal coherence relations requires more cognitive effort than the processing of positive-causal coherence relations. We expected children's and adults' responses to the cognitively more complex negative-causal sentence pairs to be slower and less accurate than their responses to the cognitively less complex positive-causal sentence pairs. Our hypothesis was supported by the analyses of response accuracy for children and adults in the visual and the auditory modality. The more complex negative-causal coherence relations resulted in more errors in both the visual and auditory semantic verification judgement task compared to the less complex positive-causal coherence relations in children and adults. The only exception was that children providing slow responses (-2 *SD*) in the visual semantic verification task made fewer errors when responding to coherent negative-causal sentence pairs compared to incoherent positive-causal sentence pairs. However, as expected, their responses to coherent negative-causal sentence pairs were less accurate compared to coherent positive-causal sentence pairs (Figure 2a). In sum, the response accuracy results of the younger children in our study were in line with the findings for German second- and third-graders in Dragon *et al.*'s (2015) study.

In contrast to the response accuracy data, response latencies provided evidence in support of the cumulative cognitive complexity approach only in adults and only in the visual version of the task. As predicted by the cumulative cognitive complexity approach, adults needed more time to respond to negative-causal compared to positive-causal sentence pairs in the visual version of the task. However, no significant effect of polarity was found in the auditory modality. Children's responses to incoherent negative-causal sentence pairs were slower than their responses to incoherent positive-causal sentence pairs but only in the auditory version of the task. The expected difference in response latencies was obtained

only at higher grade levels for coherent sentence pairs. Data from the visual version of the task revealed no significant main effect of polarity.

In sum, the accuracy data provide evidence in favour of the cumulative cognitive complexity approach. However, the response latency data are more difficult to interpret. To understand the response latency data, closer scrutiny of their relationships with response accuracy in the children's data is helpful. We found positive associations of response latency with response accuracy for positive-causal coherent, positive-causal incoherent, and negative-causal incoherent sentence pairs in the visual version of the task, and for positive-causal incoherent and negative-causal incoherent sentence pairs in the auditory version of the task. These associations can be interpreted in terms of a speed-accuracy trade-off (Pachella, 1974): children were able to respond more accurately when they invested more cognitive effort in the task, as indicated by the longer response latencies. In contrast, longer response latencies in the negative-causal coherent sentence pairs were associated with LESS accurate responses. The unique feature of the negative-causal coherent sentence pairs is that they link together two statements that are inconsistent with common knowledge but nevertheless require a positive response in the verification task. Paradoxically, more accurate responses in the verification task result when children are unaware of the inconsistency between the two statements. When children detected the inconsistency (through validation processes; Isberner & Richter, 2013), they might have also engaged in more elaborative processes in an attempt to repair the inconsistency, which lead to longer response latencies. However, these repair processes seem to have failed more often than they were successful, i.e. children were not able to construct a meaningful mental model of the described situation in coherent negative-causal sentence pairs, as indicated by the lower response accuracy.

The second aim of the present study was to directly compare the processing of positive-causal and negative-causal coherence relations in written and spoken text comprehension. With response accuracy as the dependent variable, children's response patterns were highly similar for the visual and the auditory semantic verification task, i.e. children responded with overall higher response accuracy to positive-causal compared to negative-causal sentence pairs as expected by the cumulative cognitive complexity approach. The only exception was that children's responses to negative-causal coherent sentence pairs were not less accurate than their responses to positive-causal sentence pairs in the visual semantic verification task when children provided fast responses. Furthermore, longer response latencies were associated with less accurate responses in both tasks for negative-causal coherent sentence pairs.

Children's response latency patterns were somewhat different in the visual and the auditory semantic verification task. The expected effect of polarity in the visual semantic verification task was found only for coherent sentence pairs when inaccurate responses were provided. In the auditory semantic verification task, children's responses to incoherent positive-causal sentence pairs were faster than their responses to incoherent negative-causal sentence pairs. However, the expected effect of polarity for coherent sentence pairs was found only with increasing grade level, given that response speed for coherent positive-causal sentence pairs increased from Grades 1 to 4 but not for coherent negative-causal sentence pairs (Figure 5).

The pattern of adult response accuracy was highly similar for the visual and the auditory semantic verification task. They responded with higher accuracy to positive-causal compared to negative-causal sentence pairs in both tasks. However, the response latency of adults showed the expected effects of polarity and coherence only in the visual semantic verification task.

In sum, patterns of response accuracy were highly comparable for both modalities. This finding is consistent with the predictions based on the simple view of reading that processes of listening and reading comprehension are basically the same after written word forms have been decoded. However, the response latencies suggest another interpretation. The obtained differences in response latencies between the two modalities were neither consistent with the predictions implied by the simple view of reading nor with the prediction that negative-causal sentence pairs are more difficult to process in the visual than in the auditory version of the task. An alternative explanation for the obtained differences in the response latency patterns between the two modalities might be that the presentation of the sentence pairs differed slightly between the visual and the auditory version of the task. In the visual version of the task, the sentence pairs remained on the notebook screen until the participant pressed the response button, i.e. participants could reread the sentence pairs if necessary; but sentence pairs were only presented once in the auditory version of the task. The response latencies could thus reflect different sets of processes in the visual and the auditory versions of the task, which nullifies the possibility of comparing the latencies between the two tasks.

The third aim of the present study was to investigate age-related differences in the processing of negative-causal and positive-causal sentence pairs throughout the first four years of primary school. In line with the assumptions of the cumulative cognitive complexity approach, and with findings of previous studies with English and Dutch children, we found that developmental trends in the processing of negative-causal coherence relations lag behind developmental trends in the processing of positive-causal coherence relations. Children at all grade levels, even the

youngest children at Grades 1 and 2, responded with high accuracy to positive-causal sentence pairs and improved further with increasing grade level. At the end of Grade 4, children's response accuracy for positive-causal sentence pairs approached maximum scores in the auditory and the visual version of the task. In contrast, children's performance on negative-causal sentence pairs was fairly poor in both the visual and the auditory version of the task. First graders even responded with far below-chance accuracy to negative-causal sentence pairs. This finding indicates that children, at least at Grade 1, systematically accept incoherent negative-causal sentence pairs as coherent, and systematically reject coherent negative-causal sentence pairs as incoherent.

How can we explain these systematically incorrect answers of the youngest children in the auditory semantic verification task? A possible explanation might be that the younger children simply misunderstood the task. They might have judged the plausibility of the information conveyed by the sentence pair instead of judging its overall coherence, possibly because of a lack of metalinguistic knowledge. As a result, the current paradigm might not be appropriate for at least the youngest group of children in our study. To rule out this possibility, we analyzed a subset of eight incoherent filler items of the auditory semantic verification task. Similar to the critical items, four of the filler sentence pairs were incoherent because of being linked by an inappropriate positive-causal connective (16), and four of them were incoherent because of being linked by an inappropriate negative-causal connective (17).

- (16) *Der Vater schimpft viel. Darum hat der Vater einen Bart.
 *'The father grumbles a lot. Therefore, the father has a beard.'
- (17) *Marvin geht in die erste Klasse. Trotzdem mag er kein Gemüse.
 *'Marvin is in first grade. Nevertheless, he does not like vegetables.'

The difference between the incoherent critical items and these incoherent filler items is that the two statements in the filler items are related additively instead of causally, i.e. a positive-additive connective such as *and* would make the sentence pairs coherent. If children were to simply misunderstand the task and judge the plausibility or consistency of the events described in the two statements independent of the linking connective, we would expect them to systematically (and incorrectly) accept each of the eight incoherent filler items. The reason is that the contents of the sentences involved are not implausible or inconsistent (e.g. *Marvin is in first grade* and *he does not like vegetables*). However, if the youngest children's poor performance on negative-causal items in our auditory semantic verification task were due to difficulties in comprehending negative-causal connectives and coherence relations (in contrast to positive-causal ones), as predicted by the cumulative

cognitive complexity approach, children should correctly reject the four filler items linked by an inappropriate positive-causal connective and should struggle to correctly respond to the four filler items linked by a negative-causal connective. A GLMM revealed higher estimated probabilities for correct responses for filler sentence pairs linked by positive-causal connectives compared to filler sentence pairs linked by negative-causal connectives at all grade levels. The youngest children, in particular, performed strikingly poorly on filler items containing negative-causal connectives but not on filler items containing positive-causal connectives. Thus, children did not respond incorrectly to the eight filler items as would have been expected if they had simply misunderstood the task. In contrast, children at all grade levels responded less accurately to filler items containing negative-causal compared to positive-causal connectives. This result suggests that the poor performance on negative-causal items of the youngest children in our auditory semantic verification task was due to difficulties in processing negative-causal connectives and coherence relations as predicted by the cumulative cognitive complexity approach.

In sum, the accuracy results we obtained for children at Grades 1 to 4 (auditory task) and Grades 3 and 4 (visual task) indicate that, as expected, even the younger children performed fairly well on positive-causal coherence relations, whereas their performance on negative-causal coherence relations was poor. Moreover, despite the increase in response accuracy in negative-causal sentence pairs, children's response accuracy to negative-causal sentence pairs at the end of Grade 4 was still far below the accuracy obtained for positive-causal sentence pairs. These findings strongly suggest that the comprehension of negative-causal coherence relations is still developing throughout the course of primary school, whereas even the youngest children in our study mastered the comprehension of positive-causal coherence relations. These conclusions are consistent with previous findings on the acquisition order of coherence relations in primary school children (e.g. Cain & Nash, 2011; Katz & Brent, 1968; Spooren & Sanders, 2008; Wing & Scholnick, 1981) and with the assumptions of the cumulative cognitive complexity approach.

Finally, we expected shorter response latencies for coherent compared to incoherent sentence pairs. This pattern was observed in children in the visual version of the task for positive-causal sentence pairs but was observed only for accurate responses to negative-causal sentence pairs. In the auditory version of the task, younger children's responses to incoherent negative-causal sentence pairs were slower than their responses to coherent sentence pairs, whereas a similar pattern was found for positive-causal sentence pairs in older children. The expected pattern of response latencies was also observed in adults in the visual semantic

verification task. The longer response latencies for incoherent sentence pairs thus appear to reflect the cognitively laborious attempt to repair the mental model of the text when encountering and integrating incoherent information (Cain & Nash, 2011). Yet, the expected effect of coherence was not found for adults in the auditory version of the task, and only in interaction with polarity and response accuracy or grade level in the children's data. This null finding might be explained by the sensitivity of response latencies to several strategic or non-strategic on-line and off-line processes of text comprehension, which we could not control for (as elaborated below).

The results presented so far should be interpreted with possible limitations of the present study in mind. First, latencies were recorded only for responses at the end of the second sentence. More detailed analyses of on-line reading times might be more appropriate to reveal processing differences between positive-causal and negative-causal sentence pairs as expected by the cumulative cognitive complexity approach. For example, a recent self-paced reading study by Knoepke, Richter, and Diener (2015) investigated the timecourse of validating and integrating consistent and inconsistent information in positive-causal and negative-causal sentence pairs. Although they found no overall differences in reading times for the whole sentence pairs, their analyses revealed longer reading times on inconsistent information and at the end of the second statement in negative-causal sentence pairs compared to positive-causal sentence pairs, which contained only consistent information. In accordance with the assumptions of the cumulative cognitive complexity approach, their findings indicate increased cognitive effort in the processing of negative-causal coherence relations, in particular when detecting and integrating inconsistent information into the existing mental model of the text.

A general methodological limitation of the current study is its cross-sectional design. The more appropriate way to investigate developmental changes in language comprehension is to employ a longitudinal design. Given our design of investigating children from four different grade levels in parallel, we cannot be entirely sure whether the developmental changes are due to differences between the different groups. Nevertheless, the accuracy data are in accordance with previous findings on the acquisition order of positive-causal and negative-causal connectives and coherence relations. Thus, it is likely that our findings reflect real developmental trends instead of random group differences.

A third limitation addresses the external validity of our study. The material used to test the cumulative cognitive complexity approach included very short texts containing only one sentence pair each. Arguably, sentence pairs might not be sufficient to investigate processes of establishing coherence, because they are not sufficiently naturalistic. An

argument against this concern is that the comprehension and integration of two sentences already requires the establishment of local coherence relations, because local coherence per definition links the mental representations of two adjacent statements in a text. Thus, to study the establishment of local coherence in isolation, two adjacent sentences or statements represent a minimalistic but sufficient approach. Nonetheless, further research is needed to show whether the results we obtained generalize to longer texts.

Finally, we focused on only two types of coherence relations characterized along only two of the four dimensions of the coherence taxonomy proposed by Sanders *et al.* (1992). Further studies are needed to investigate whether the predictions of the cumulative cognitive complexity approach also hold for coherence relations other than positive-causal or negative-causal coherence relations. For example, research could focus on positive-additive and negative-additive relations, coherence relations based on pragmatic versus semantic sources, or coherence relations with basic versus non-basic order in German.

Although the response latency data are difficult to interpret, the response accuracy data in our study suggest that the cumulative cognitive complexity approach can be applied to German primary school children and adults. We demonstrated that the processing of negative-causal coherence relations is cognitively more demanding than the processing of positive-causal coherence relations in German children from Grades 1 to 4 and adults in reading and listening comprehension. We also showed that children's comprehension of negative-causal coherence relations is still developing throughout the course of primary school. These findings have important implications for classroom instruction and the composition of textbooks in primary school. Considering the crucial role the establishment of local coherence plays in written and spoken text comprehension (Graesser *et al.*, 1997; Van Dijk & Kintsch, 1983), the finding that primary school children even at Grades 3 and 4 exhibit great difficulties in processing negative-causal coherence relations is of high practical relevance. To ensure successful text comprehension as early as possible, a reasonable and promising, if not necessary, strategy should be to include systematic training concerning the establishment of coherence relations (e.g. in the form of strategy exercises) into primary school curricula.

REFERENCES

- Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B. & Sigmann, H. (2014). *Lme4: linear mixed-effects models using Eigen and S4* [Software]. R-package version 1.1-6. Online: <<http://cran.r-project.org/package=lme4>> (last accessed 1 May 2014).

- 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**, 235–61.
- Cain, K. & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology* **103**, 429–41.
- Cain, K., Patson, N. & Andrews, L. (2005). Age- and ability-related differences in young readers' use of conjunctions. *Journal of Child Language* **32**, 877–92.
- Degand, L., Lefèvre, N. & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design* **1**, 39–51.
- Dixon, P. (2008). Models of accuracy in repeated measures designs. *Journal of Memory and Language* **59**, 447–56.
- Dragon, N., Berendes, K., Weinert, S., Heppt, B. & Stanat, P. (2015). Ignorieren Grundschulkindern Konnektoren? – Untersuchung einer bildungssprachlichen Komponente [Do primary school children ignore connectives? – Investigation of a literary component]. *Zeitschrift für Erziehungswissenschaft* **18**, 803–25.
- Evers-Vermeul, J. & Sanders, T. (2009). The emergence of Dutch connectives: how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language* **36**, 829–54.
- Goldman, S. R. & Murray, J. D. (1992). Knowledge of connectors as cohesion devices in text: a comparative study of native-English and English-as-a-second-language speakers. *Journal of Educational Psychology* **84**, 504–19.
- Gough, P. B. & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education* **7**, 6–10.
- Graesser, A. C., Millis, K. K. & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology* **48**, 163–89.
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal* **2**, 127–60.
- Isberner, M.-B. & Richter, T. (2013). Can readers ignore implausibility? Evidence for nonstrategic monitoring of event-based plausibility in language comprehension. *Acta Psychologica* **142**, 15–22.
- Kail, M. & Weissenborn, J. (1984). A developmental cross-linguistic study of adversative connectives: French 'mais' and German 'aber/sondern'. *Journal of Child Language* **11**, 143–58.
- Kamalski, J., Sanders, T. & Lentz, L. (2008). Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: sorting things out. *Discourse Processes* **45**, 323–45.
- Katz, E. W. & Brent, S. B. (1968). Understanding connectives. *Journal of Verbal Learning and Verbal Behavior* **7**, 501–9.
- Knoepke, J., Richter, T., & Diener, P. (2015). Processing local coherence relations as a function of polarity and of position of the connective in German. Paper presented at the 25th Annual Meeting of the Society for Text & Discourse, July.
- Knoepke, J. & Richter, T. (in press). Reading comprehension: individual differences, disorders, and underlying cognitive processes. In A. Bar On & D. Ravid (eds), *Handbook of communication disorders: theoretical, empirical, and applied linguistic perspectives*. Berlin: De Gruyter Mouton.
- Köhne, J. & Demberg, V. (2013). The time-course of processing discourse connectives. In M. Knauff, M. Pauen, N. Sebanz & I. Wachsmuth (eds), *Proceedings of the 35th Annual Conference of the Cognitive Science Society (CogSci 2013, Berlin)*. Austin TX: Cognitive Science Society.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2014). *lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R-package version 2.06. Online: <<http://cran.r-project.org/web/packages/lmerTest/index.html>> (last accessed June 2014).
- Maury, P. & Teisseranc, A. (2005). The role of connectives in science text comprehension and memory. *Language and Cognitive Processes* **20**, 489–512.

- Millis, K. K., Golding, J. M. & Barker, G. (1995). Causal connectives increase inference generation. *Discourse Processes* **20**, 29–49.
- Millis, K. K. & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language* **33**, 128–47.
- Murray, J. D. (1995). Logical connectives and local coherence. In R. F. Lorch Jr. & E. J. O'Brien (eds), *Sources of coherence in reading*, 107–25. Hillsdale, NJ: Erlbaum.
- Murray, J. D. (1997). Connectives and narrative text: the role of continuity. *Memory and Cognition* **2**, 227–36.
- Pachella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (ed.), *Human information processing: tutorials in performance and cognition*, 41–82. Hillsdale, NJ: Erlbaum.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Richter, T., Isberner, M.-B., Naumann, J. & Kutzner, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern [Process-based assessment of reading skills in primary school children]. *Zeitschrift für Pädagogische Psychologie* **26**, 313–31.
- Richter, T., Isberner, M.-B., Naumann, J. & Neeb, Y. (2013). Lexical quality and reading comprehension in primary school children. *Scientific Studies of Reading* **17**, 415–34.
- Richter, T., Naumann, J., Isberner, M.-B., Neeb, Y. & Knoepke, J. (in press). *ProDi-L: Prozessbezogene Diagnostik des Leseverstehens bei Grundschulkindern* [Process-based assessment of reading skills in primary school children] [Computerized test]. Göttingen: Hogrefe.
- Sanders, T. J. M., Land, J. & Mulder, G. (2007). Linguistic markers of coherence improve text comprehension in functional contexts. *Information Design Journal* **15**, 219–35.
- Sanders, T. J. M. & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* **29**, 37–60.
- Sanders, T. J. M., Spooren, W. P. M. & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes* **15**, 1–35.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex: a lexical database for German read by children. *Behavior Research Methods*, **86**, 155–165.
- Spooren, W. P. M. & Sanders, T. J. M. (2008). The acquisition order of coherence relations: on cognitive complexity in discourse. *Journal of Pragmatics* **40**, 2003–26.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Silfhout, G., Evers-Vermeul, J., Mak, W. M. & Sanders, T. J. M. (2014). Connectives and layout as processing signals: how textual features affect students' processing and text representation. *Journal of Educational Psychology* **106**, 1036–48.
- Van Veen, R., Evers-Vermeul, J., Sanders, T. J. M. & Van den Bergh, H. (2009). Parental input and connective acquisition: a growth curve analysis. *First Language* **29**, 266–88.
- Van Veen, R., Evers-Vermeul, J., Sanders, T. J. M. & Van den Bergh, H. (2013). The influence of input on connective acquisition: a growth curve analysis of English *because* and German *weil*. *Journal of Child Language* **40**, 1003–31.
- Wing, C. S. & Scholnick, E. K. (1981). Children's comprehension of pragmatic concepts expressed in 'because', 'although', 'if' and 'unless'. *Journal of Child Language* **8**, 347–65.