STEM-, SPRAAK- EN TAALPATHOLOGIE

Vol. 27, 2022, Special Issue, pp. 177-205
Special Issue for Steven Gillis

Gepubliceerd online: 5 november 2022

University of Groningen Press
https://doi.org/10.21827/32.8310/2022-SG-177

Dutch Norming Study For 208 Color Drawings Depicting Transitive Events

Edwige Sijyeniyo[1], Robert J. Hartsuiker[2], & Sarah Bernolet[1]

[1] *Department of Linguistics, University of Antwerp*

[2] *Department of Experimental Psychology, Ghent University*

Samenvatting

We hebben een normeringsstudie uitgevoerd onder Nederlandse kinderen (6-12 jaar) om te onderzoeken of zij onze set van 208 kleurentekeningen, ontwikkeld voor longitudinaal onderzoek naar de productie en het begrip van transitieve zinnen, correct zouden interpreteren als de transitieve actie op elke tekening werd beschreven met een actieve of een passieve zin.

De kinderen gaven in 93.02% van de gevallen een correct antwoord voor onze tekeningen, wat aangeeft dat de tekeningen duidelijk zijn in termen van hoe de transitieve acties worden weergegeven. Er waren verschillende factoren die bijdroegen tot een incorrecte interpretatie van een overgankelijke actie in een plaatje. (1) Afbeeldingen waren moeilijker te interpreteren wanneer een passieve zin werd gebruikt om een overgankelijke gebeurtenis te beschrijven dan wanneer een actieve zin werd gebruikt; (2) kinderen maakten meer fouten bij het interpreteren van afbeeldingen wanneer abstracte werkwoorden, zoals "inhalen" en "vervangen", werden gebruikt om een overgankelijke gebeurtenis te beschrijven; (3) kinderen hadden moeite met het interpreteren van onze afbeeldingen wanneer bepaalde werkwoorden ("roepen" en "volgen") werden gepresenteerd in een passieve zin, terwijl andere werkwoorden ("blokkeren" en "inhalen") *meer* correcte antwoorden opleverden als de overgankelijke gebeurtenis werd beschreven met een passieve zin; ten slotte, (4) gaven jongere kinderen meer incorrecte antwoorden op onze afbeeldingen dan oudere kinderen, ongeacht of de overgankelijke handeling in een afbeelding werd beschreven met een actieve of een passieve zin. De jongere kinderen gaven vooral significant meer incorrecte antwoorden op plaatjes met het werkwoord "groeten" in vergelijking met de oudere kinderen.

Vanwege de vele werkwoord-naamwoord combinaties in onze set is deze erg nuttig voor onderzoekers die de verwerving van transitieve structuren bij kinderen willen onderzoeken, niet alleen voor studies met typisch ontwikkelende kinderen, maar ook voor studies met kinderen die een achterstand in hun taalontwikkeling hebben.

Trefwoorden: Normeringsstudie, set van kleurentekeningen, transitieve structuren, L1 syntaxverwerving

Correspondentieadres:
Edwige Sijyeniyo
Department of Linguistics
University of Antwerp
Prinsstraat 13, 2000 Antwerp
Belgium
E-mail: Edwige.Sijyeniyo@uantwerpen.be

Abstract

We conducted a norming study amongst Dutch children (6-12 years) to investigate whether they would correctly interpret our set of 208 color drawings, developed for longitudinal research on the production and comprehension of transitive sentences, if the transitive action on each drawing was described with an active or a passive sentence.

The children provided 93.02% correct answers to our pictures, which indicates that the pictures are clear in terms of how the transitive actions are displayed. There were several factors that contributed to an incorrect interpretation of a transitive action in a picture. (1) Pictures were more difficult to interpret when a passive sentence was used to describe a transitive event compared to an active sentence; (2) children made more errors interpreting pictures when abstract verbs, such as "inhalen" (to overtake) and "vervangen" (to replace), were used to describe a transitive event; (3) children had difficulty interpreting our pictures when particular verbs ("roepen" [to shout] and "volgen" [to follow]) were presented in a passive sentence, while other verbs ("blokkeren" [to block] and "inhalen") rendered more *correct* answers when the transitive event was described with a passive sentence; finally, (4) younger children provided more incorrect answers to our pictures than older children, regardless of whether the transitive action in the picture was described with an active or a passive sentence. In particular, the younger children provided significantly more incorrect answers to pictures with the verb "groeten" (to greet) compared to the older children.

Because our set contains many verb-noun combinations, the set is very useful to researchers who intend to investigate the acquisition of transitive structures in children in a longitudinal design, not only for studies with typically developing children, but also in studies with children who have a delay in their language development.
Keywords: Norming study, picture set, transitive structures, L1 syntax acquisition

Introduction

A crucial step in carrying out meaningful experiments is to choose appropriate stimuli. For instance, researchers in psychology and psycholinguistics often use pictures as stimuli in their experiments to investigate different aspects of language such as sentence production and comprehension. These pictures need to be explicit and clear enough such that they elicit specific syntactic structures from language users. A research domain in which pictures are often used to investigate the production and comprehension of syntax is the field of structural priming (Bock, 1986). Structural priming is the tendency of speakers to repeat structures that they have been recently exposed to; this is often investigated with picture descriptions or sentence completions (Mahowald et al., 2016). Most structural priming studies use pictures where transitive or ditransitive actions are displayed, and, often, these studies have used just two picture sets: the set by Bock (1986) and the one by Branigan et al. (2000). These sets have in common that they consist of black-and-white line drawings with combinations of animate and inanimate nouns performing or undergoing an action, and that they have been mostly used in studies with native speakers or highly proficient second language (L2) speakers (see e.g., Bernolet et al., 2009; Favier et al., 2019).

A difference between both sets is that the pictures[1] in Bock are rather outdated and may not necessarily appeal to, for instance, children, whereas the picture set of Branigan et al. seems more suitable to use in studies that investigate language production in special populations. However, both picture sets include verbs and nouns that are somewhat *infrequent*[2] and may be familiar to native language users or highly proficient L2 speakers, but not to other groups such as late L2 learners or children. Some infrequent verbs and nouns in Branigan et al. include, for instance, "to polish"[3], "to scold"[4], and "a monk"[5]. This important characteristic impedes researchers to confidently use the two picture sets when testing non-native or young language users.

Using frequent verbs (and nouns) in stimuli is especially important in studies involving children, since several studies have suggested that children are sensitive to lexical frequency information (Theakston et al., 2004); and that this frequency effect on the acquisition of lexical items is crucial in their development of syntactic knowledge (Matthews et al., 2005; see Ambridge et al., 2015 for an extensive review). If young children are exposed to an infrequent, and perhaps unfamiliar, lexical item and are asked to formulate a syntactic structure with that verb, they might have difficulty formulating the elicited syntactic structure since lexical representations play a role during the selection of syntactic structures (Ambridge et al., 2015; Pickering & Branigan, 1998). This highlights the importance of using frequent lexical items in stimuli that are aimed at investigating the production of syntax in young L1 users (but also in late L2 learners).

To keep the attention span of children to the task, many studies involving children use color drawings in their stimuli (see e.g., Huttenlocher et al., 2003; Goldwater et al., 2011). However, it is noticeable that the number of items in these studies is often limited. For instance, Huttenlocher et al. used 40 color drawings as their stimuli to investigate the priming of transitive and dative structures in children (two sets of 20 drawings for each structure). Moreover, Goldwater et al. used nine critical items (color drawings) in their structural priming experiment with 4- and 5-years old. Using a limited number of items usually requires more participants in the test sample to reach sufficient statistical power for observing priming effects (see Mahowald et al., 2016). It seems that, ideally, in studies with less experienced language users (i.e., children and late L2 speakers), there should be a balance between sufficient experimental items that use frequent verbs and enough participants.

Note that Muylle et al. (2020) designed a stimulus set of 423 animated movie clips depicting transitive and dative actions (prepositional object datives ["The woman gives a flower to the man"] vs. direct object datives ["The woman gives the man a flower"]) that can be used

[1] See OSF directory (https://tinyurl.com/DutchNormingStudy) for an example of pictures from both picture sets.

[2] In this paper, we denote frequent words based on Zipf-frequencies (Van Heuven et al., 2014; Keuleers et al., 2010). Zipf-frequencies range from 1 to 7, the values 1-3 indicate low frequent words (i.e., frequencies of 1 per million words and lower) and the values 4-7 indicate high frequent words (i.e., words with frequencies of 10 per million words or higher).

[3] "To polish" has a Zipf-frequency of 3.03

[4] "To scold" has a Zipf-frequency of 3.12

[5] "monk" has a Zipf-frequency of 3.86

to investigate transitive and ditransitive comprehension and production in groups such as adult L2 learners or children. Indeed, as the researchers propose, using clips that depict motion events may be easier for non-native or young language users since the action does not have to be inferred, and therefore, motion clips may be more ecologically valid than static pictures. However, some movie clips in Muylle et al. are rather violent (e.g., "A boxer punches a dancer"). Because of this, some of their clips may still not be very appropriate to use in studies with children.

In addition to animated movie clips, some studies have used live action objects, such as bunnies and ducks (Gertner et al., 2006) to elicit syntactic structures from young children in a language production task. Though these types of stimuli may strongly appeal to children, the development of such stimuli is effortful and can often only be used for the aforementioned group and not for other groups of language users (e.g., late L2 learners, people with aphasia...).

Here, we propose a set of colored pictures that was initially designed to test transitive comprehension and production in adult learners of Dutch. However, the purpose of the current study is to investigate whether our picture set is suitable to use in studies with young children too. For this reason, we conducted a norming study amongst young monolingual Dutch children.

Current study

In this study, we used 208 color drawings that were initially developed for longitudinal research on the production and comprehension of transitive sentences in adult learners of Dutch. Prior to constructing the color drawings for the late learners, we consulted their language learning materials to make sure that we used familiar verbs and nouns in our stimuli set[6]. Based on the learning materials, we chose transitive verbs (e.g., to call, to carry, to follow...), names of human professions (e.g., a cook, a baker, a teacher...), and names of vehicles (e.g., a car, a truck, a motorbike...) (see Appendix S1 for the verbs and nouns). We used a restricted number of verbs (12 verbs) and nouns (8 animate nouns and 8 inanimate nouns) for the color drawings since we wanted to avoid that the learners would be exposed to many different words. However, despite the small number of verbs and nouns that we used, our picture set was designed such that different verb-noun combinations are possible, which allows the same verb to be presented with different nouns while avoiding lexical overlap between items. This makes our picture set particularly suitable for longitudinal designs, since researchers can, for instance, test the acquisition of specific transitive verbs and both transitive structures (active and passive sentences) in adult and young language users, while avoiding a learning effect caused by repetition of the pictures.

A previous study showed that our picture set is indeed suitable to test the production of transitive structures in a group of late learners of Dutch. In a structural priming experi-

[6]To our knowledge, there exist no frequency data bases on the acquisition of Dutch verbs in late adult learners. For this reason, using the learners' learning materials was reliable for our study as it ensured us that we were using familiar words to these learners.

ment, Van Lieburg et al. (2022) observed that the learners produced 49.5% active sentence responses, 29.9% passive sentence responses and 20.7% "Other" responses to the transitive pictures. This suggests that the transitive events in our drawings largely elicit active as well as passive responses.

In the current study, however, we investigated whether the same picture set is suitable to use in studies involving young children. For this purpose, monolingual Dutch children (a group of 6-7 years old and a group of 11-12 years old) indicated whether our drawings can be interpreted equally well with an active and a passive sentence. We predict that these children will have difficulty interpreting our pictures when a transitive event is described with a passive sentence, since passives are complex structures and are learned relatively late (around the age of 3) compared to active sentences (Messenger & Fisher, 2018; Verrips, 1996). Passives are particularly complex to acquire by children due to their noncanonical word order (the agent of the sentence does not occur in the grammatical subject position compared to conventional word order rules [particularly in SVO languages]). Moreover, language production studies have suggested that children are more likely to first produce short passives (passive structures in which the by-phrase is not overtly realized: "The bear is hugged") than full passives (Fox & Grodzinsky, 1998). However, language comprehension studies have not found a consensus as to whether short passives are comprehended earlier than full passives (Hirsch & Wexler, 2006). Although young children can produce passive sentences spontaneously by the age of 3, they seem to have persistent trouble with understanding some type of passives. A common difficulty in young children's understanding of passives are reversed passives (Messenger et al., 2012; Armon-Lotem et al., 2016). Reversed passives are passives in which full grammatical knowledge is required for the child to interpret who does what to whom (Armon-Lotem et al., 2016). For instance, in the sentence "The woman was kissed by the man", the animate noun in the grammatical subject position and the one in the by-phrase can be exchanged since both nouns can do the kissing. Thus, here, the grammatical subject can be replaced with the agent in the by-phrase, while the sentence remains grammatically correct. If a child would be asked who does the kissing, they would most likely answer that the woman is the one doing the kissing (based on conventional word order rules). Nonreversible passives, however, are less challenging for young children to understand since children can rely on world knowledge to interpret who does what to whom (e.g., "The cheese was eaten by the mouse") (Armon-Lotem et al., 2016). Here, the child knows that only the mouse can do the act of 'eating'. Next to this, children have difficulty understanding passives with mental state verbs (e.g., *remembered, forget, know, believe* – "Joanna was forgotten by Olivia"). This type of passives is challenging for young children because the agent is experiencing the mental state (here, Olivia is doing the forgetting), while nothing "happens" to the patient. In contrast, passive sentences with actional verbs (e.g., *to push, to hug, to carry...*) are less challenging for children since the transitive action is carried out by the agent, while the patient undergoes the transitive action (Armon-Lotem et al., 2016). These features of the passive structure explain why the production and comprehension of (full) passives remains challenging throughout the first years of a child's education (Messenger et al., 2012).

Next to the difficulty in passive sentence comprehension, we also investigated to what extent the transitive verbs on our pictures would affect the children's comprehension of each

transitive action, when an active or a passive was used to describe this action. We focused on verbs specifically (and not on the nouns in our pictures) because verbs are the most important building block in constructing syntactic structures (see Ambridge et al., 2015). Before conducting our experiment, we did not test whether the children knew the verbs that we used on our pictures, but we assumed that some of our pictures contain verbs that are abstract and may not have been acquired yet by the children. For example, one of our pictures depicts an event in which a car overtakes a truck ("De auto haalt de vrachtwagen in" – "The car overtakes the truck"), and the picture has been drawn from a bird's eye view (see Figure 1). Because adult learners have experience in traffic, understanding who passes whom is not particularly difficult. However, "inhalen" may be a challenging word for children because they do not have that same experience yet, especially because the action verb has been drawn from a bird's eye view[7]. What is more, young children might also have not acquired this verb yet ("inhalen" is acquired around the age of 8;61 years by Dutch children, see Brysbaert et al., 2014). In addition to this, "inhalen" has a Zipf-frequency of 2.72 (Keuleers et al., 2010); this means that it is considered a low frequent word[8], which may cause even more difficulty for children to interpret pictures with this verb correctly. Moreover, we expect that the combination of some abstract verbs (such as "inhalen") with a passive voice will further restrict the children's comprehension of the transitive event in our pictures.

Finally, we anticipate that, in general, the older children will be better at interpreting our pictures, either described with an active or a passive sentence, compared to the younger children. If this turns out to be the case, then our picture set will be suitable to test how children make progress in the comprehension of transitive structures.

Method

Participants

Two hundred participants (86 participants in the first grade (41 males and 45 females) and 114 participants in the fifth and sixth grade, 56 males and 58 females) took part in this study. Participants in the first grade were 6-7 years old and participants in the higher grades were 11 or 12 years old. Five classes in the first grade participated and five classes in the higher grades took part in this study (two classes in the fifth grade and three classes in the sixth grade). The participants were recruited at the same primary school in Antwerp, Belgium. All parents provided their consent for their child's participation in this study. The classes received a monetary reward for their participation. This study was approved by the ethical committee of the University of Antwerp (SHW_1877).

[7] Of course, if "inhalen" depicted a transitive event that is more familiar to children (a bike race or a running contest) or if pictures with this verb were drawn from a side view, the verb may be easier for children to comprehend. However, here, we specifically predict that the combination of the abstract verb "inhalen" together with the bird's eye view will probably require the children to make more inferences, which will probably lead to difficulty in comprehension.

[8] It is important to note that the verb Zipf-frequencies in Keuleers et al. (2010) are based on Dutch film subtitles.

Stimuli & Design

Stimuli

We used 208[9] pictures for the present study. The 208 pictures were pairs of target and competitor pictures (thus, there were 104 pairs). We used 12 different verbs to depict transitive events in the set of 208 pictures. The target and competitor pictures always used the same verb. The verbs were more or less equally divided across the 208 pictures: four verbs (*aanrijden* [to hit], *bellen* [to call], *dragen* [to carry], and *volgen* [to follow]) each occurred 20 times in the set and eight other verbs (see Appendix S1 for all the verbs; their Zipf-frequency and the age of acquisition of these verbs) occurred 16 times in the pictures. Importantly, because we used the verbs based on the L2 learning materials of late Dutch learners (see above), we did not control for how many verbs were abstract or concrete (i.e., we did not equally divide the verbs into an abstract and concrete category). We believe that our materials consist of 3 abstract verbs (*inhalen* [to overtake], *vervangen* [to replace]; *blokkeren* [to block]), while the remaining 9 verbs may be considered concrete verbs (see Appendix S2 for examples of pictures with the abstract verbs "vervangen" and "blokkeren"; see Figure 1 for an example with the verb "inhalen"). Next to this, the agents and the patients of the transitive actions could be one of eight animate entities referring to a profession (e.g., a baker, a construction worker, a cook) or one of eight inanimate entities referring to a vehicle (e.g., an ambulance, a firetruck, a car [see Appendix S1 for all the nouns]). Thus, our sentences contained only animate subject – animate object (e.g., "De bakker helpt de zangeres" – "The baker helps the singer") or inanimate subject – inanimate object ("De vrachtwagen volgt de auto" – "The truck follows the car") combinations. It is important to note that all sentences in our picture set were reversible (i.e., the agent and patient can be exchanged while our sentences remain grammatically correct) and that all our verbs were actional verbs (our materials did not have any mental state verbs).

Design

The target-competitor pairs were divided in five different lists consisting of 42 or 41 items. After doing this, we randomly assigned the five classes in the lower and higher grade to one of the five lists (these lists were later converted into separate booklets, see below). The first three lists had 42 target-competitor pairs and the last two lists consisted of 41 pairs. In every list, we manipulated the position of the target pictures within each target-competitor pair. This means that, within a list, half of the target pictures (21 or 20 pictures) appeared on the left and the accompanying competitor pictures appeared on the right (see Figure 1). Similarly, the other half of the target pictures (also 21 or 20 pictures) were on the right and the accompanying competitors were on the left. Next to this, we aimed to equally present an active and passive interpretation of a target picture within each target-competitor pair.

[9]There were 194 unique pictures, and 14 pictures were used twice, but elicited two different states of affairs (e.g., the picture for an ambulance replacing a bus elicited the sentence "the ambulance replaces the bus" in one trial, while it elicited the sentence "the bus is being replaced by the ambulance" in another trial). The pictures that occurred twice in the set are marked in the OSF directory.

For this reason, within a list, half of the sentences (i.e., the sentences that participants would hear during the experiment) were actives and the other half were passives. For example, if the children would be presented with an active sentence "The car overtakes the truck", they would see two pictures (see Figure 1). In the correct target picture, the agent is the car, and the truck is the patient. In the competitor picture, the agent is the truck, and the patient is the car. The competitor picture could either be described with an active sentence "The truck overtakes the car" or with a passive sentence "The car is being overtaken by the truck". We included the competitor picture to test whether the children would interpret the first noun phrase as an agent (active interpretation) or a patient (passive interpretation).

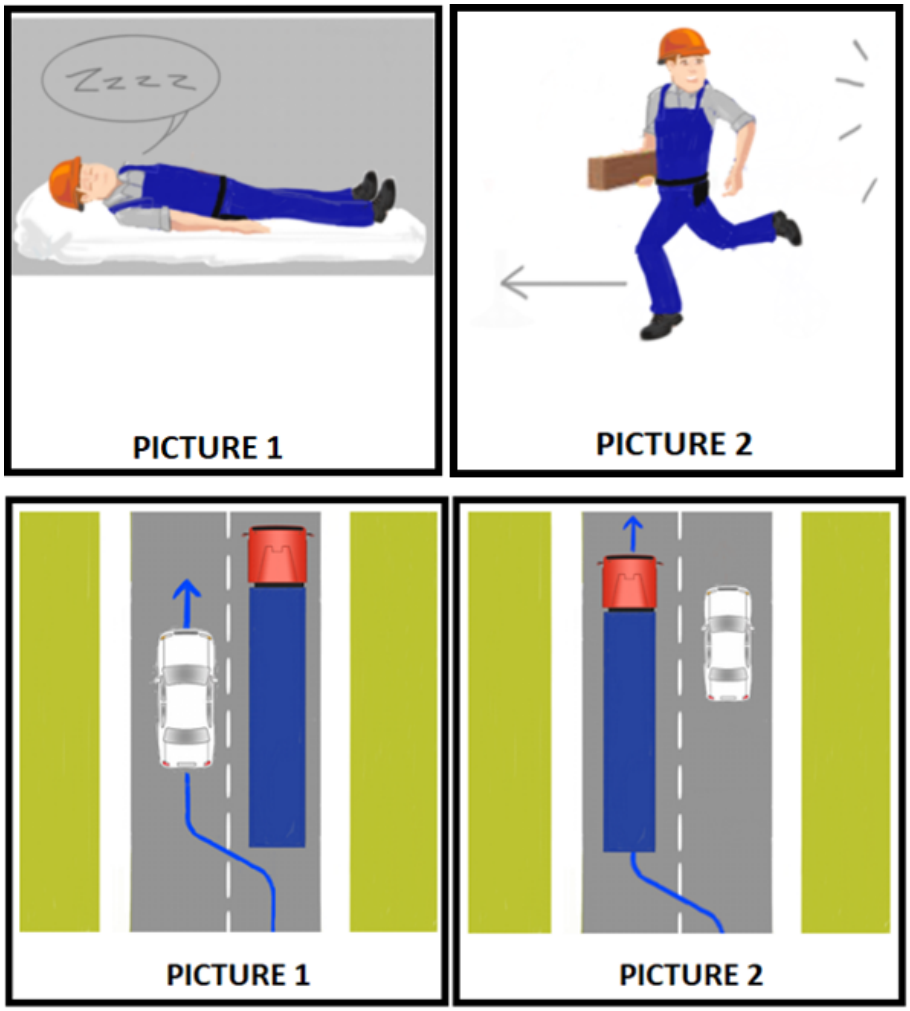


Figure 1: An example of a target-competitor pair in the practice trial for the sentence "De bouwvakker slaapt" – The construction man sleeps (1) and in an experimental trial (2) for the sentence "De auto haalt de vrachtwagen in" – The car overtakes the truck. For both trials, the correct picture is on the left and the competitor picture is on the right.

Because each target-competitor item appeared once within a list, we could only present the active or passive interpretation of that target picture. Since our goal was to test the children's comprehension of a target picture in both transitive sentence structures, we manipulated the condition of each target picture across lists. That is, each target picture was presented with an active sentence in one list and with a passive sentence in another list. In this way, we collected data on whether children interpreted the target picture correctly when it was described with an active sentence and its passive counterpart. Furthermore, the target-competitor items were constructed such that each transitive verb occurred at least once in each list; thus, every participant was exposed to all twelve verbs.

Our list construction deviates from conventional experimental lists in psycholinguistics due to several reasons. (1) Because we tested young children, we could not expose them to all 104 target-competitor pair items due to their attention span and thus, we reduced the number of items that children were exposed to within a list (42 or 41 items). (2) Our aim was to test whether a target picture could be interpreted equally well with an active and passive sentence, and because of this, across lists, there were two conditions, but we had five experimental lists (the first 3 lists had 42 items, and the last two list had 41 items). For this reason, we selected a different set of items for each list.

Procedure

Because of the COVID-19 pandemic, the experiment could not be administered by the researchers themselves. For this reason, the children were tested in class and the experiment was administered by their teachers, whom we provided with detailed instructions. Each class was assigned a separate list (there were five lists in total). We presented the five lists with the target-competitor items in five separate Microsoft PowerPoint presentations. The teachers were also provided with a Word Document containing the pictures of the (in)animate nouns that would appear in the experiment. The teachers first showed this document on a digital white board to the children to familiarize them with the pictures of the nouns and their accompanying names. After this, the teachers could start with the experiment, which was presented in a PowerPoint Presentation on a digital white board. Each list (i.e., each presentation) started with the same practice trial consisting of an audio file of the intransitive sentence "De bouwvakker slaapt" – "The builder sleeps". We used an intransitive sentence in the practice trial so that the goal of the experiment (how children interpret transitive events when a passive or an active sentence is used) would not be revealed. After participants had listened to the audio file, the teacher showed the target and competitor pictures describing the intransitive event on a separate page in the PowerPoint presentation (see Figure 1). The two pictures were framed with a black line. The phrase "picture 1" appeared at the bottom of the left picture, and "picture 2" appeared at the bottom of the right picture (Calibri bold, 18-point font, see Figure 1). At this point, the teacher asked the class which picture depicted the sentence "The builder sleeps" and the children had to say the answer out loud. After the practice trial, teachers could start with the experimental trials. For each trial, the children listened to an audio file of a sentence describing either an active or passive transitive event. Subsequently, a picture of the target and competitor item,

that depicted the transitive event, was shown on the PowerPoint-presentation. Crucially, the same target-competitor items that were shown in the PowerPoint presentation were also presented to each participant in their answer booklet. It is important to make clear that, unlike the practice trial that did not appear in participants' answer booklet, the experimental trials all appeared in the children's answer booklet. In each child's answer booklet, every target and competitor item appeared on one page next to each other (see Figure 1 for an example of the verb 'inhalen' – to overtake). The presentation of the target-competitor items in the PowerPoint presentation and in participants' answer booklet was done to enable group testing. Participants had to indicate which picture (either picture 1 or 2) matched the sentence they had heard; they marked this in their booklets.

We told teachers that each audio file could be listened to twice. Because we wanted to provide teachers with some flexibility in how they played the audio files to the children, teachers could either choose to let the children listen to an audio file the first time, after which they marked the picture in their booklet, and listened to the same audio file again, or teachers could let the children listen to each audio file twice before they provided an answer in their booklet. In this way, teachers could freely choose which method worked best for their class. We emphasized that once teachers had chosen a method, they had to stick to it throughout the experiment. We also provided teachers with an Excel sheet in which the audio sentences were written down. Teachers could consult the sheet in case they wanted to know which sentence the children would hear next. We instructed teachers to allow the children to ask questions during the experiment or to further explain what was expected from them. Also, we told teachers that if any technical difficulties would occur during the experiment (i.e., if an audio file could not be played), they were allowed to read the sentence out loud to the children; they could use the Excel sheet for this. Sessions took approximately 30 to 45 minutes.

Coding

The answers to the pictures were coded as correct or incorrect. If the target picture was marked, the trial was coded as correct and if the competitor was marked instead, the trial was coded as incorrect. Moreover, if it was clear that the pupils first gave one answer and then changed it (e.g., by erasing it), the first given answer was coded. We did this because, in this study, we were specifically focused on the children's initial comprehension of the sentence they had heard. We believe that including the self-corrected trials in the analyses would not provide us with the most accurate comprehension of a given picture. On trials where children did not provide an answer, the trial was marked as a missing value (i.e., NA). We did not include the missing values in our analyses reported below, since our dependent variable was a binary response, and the number of missing values was very low (see below).

Analysis

We fitted the answers to the pictures (correct answer vs. incorrect answer, correct answer as the reference level) to generalized linear mixed effect models as implemented in the *lme4*

package (Bates et al., 2015) in R (R Core Team, 2020). Our full model consisted of a three-way interaction between *Condition* (active, passive, active as reference level) * *Verb* (aanrijden [to hit] as reference level) * *Grade* (high grade vs. low grade, high grade as reference level); we used a Bobyqa optimizer to increase convergeability (Powell, 2009). We accounted for the non-independence between observations from the same participant and from the same item, by entering the random intercepts for *participant* and *item*. Conform the maximal random effects structure as proposed by Barr et al., (2013), we added *Condition* and *Verb* as random slopes over *participant* and *Condition* as a random slope over *item*.

Due to convergence issues and singular fit warnings, the maximal model was simplified in a stepwise way. We first simplified the random effects structure by testing if the random slope terms could be omitted without decreasing the fit of the model. Second, the fixed effects part was simplified by testing which interactions were not significant; for this, we used the *drop1* function from the basic stats package in *R* (version 4.1.2) Based on the output that we received from the drop1 function, we decided which interactions could be dropped from the model (*Condition*Verb*Grade*, and *Condition*Grade* were not significant). Lastly, the conditional and marginal R^2 values of the final model, which are measures of effect sizes, were calculated using the *rsquared* function from the piecewiseSEM package (version 2.1.0., Lefcheck, 2016).[10]

Results

In total, we collected 8,319 answers to the active and passive descriptions of our pictures. There were 7,700 correct answers (92.60%), 578 incorrect answers (6.95%) and 41 missing values (0.493%).

For the active condition, the *higher* grade provided 2321 correct answers (97.8%), 39 incorrect answers (1.64%) and there were 11 missing values (0.46%). For the passive condition, the higher grade provided 2272 correct answers (95.8%), 93 incorrect answers (3.92%) and they did not give an answer on 6 trials (0.25%).

For active sentences, the *lower* grade gave 1649 correct answers (92.2%), 136 incorrect answers (7.61%) and there were 3 missing trials (0.17%). For passive sentences, the lower grade provided 1458 correct answers (81.5%), 310 incorrect answers (17.3%) and there was no answer for 21 trials (1.17%). Figure 2 shows the number of correct and incorrect answers per sentence condition per grade. However, note that the percentages are slightly different from the percentages listed above since we did not include the missing values in the plot and in further analyses.

Our final model consisted of a significant two-way interaction between *Condition*Verb* ($p < .001$) and a two-way interaction between *Verb*Grade* ($p < .001$) and *participant* and *item*

[10]We performed exploratory analyses to investigate whether there was an interaction between Zipf-frequency * Verb and an interaction between AoA (age of acquisition) * Verb, but we found no interaction between these variables. For this reason, we decided to exclude the Zipf-frequencies and AoA values from our main analysis (see Appendix S1 for these values per verb). Throughout the rest of the paper, further comments on verb frequencies and AoA are thus rather descriptive than drawn from inferential statistics.

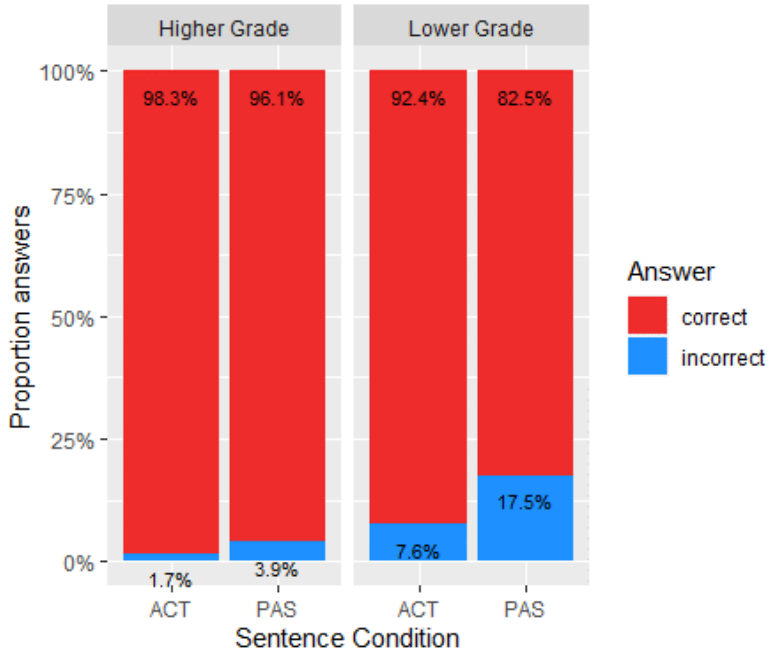


Figure 2: Proportion of correct and incorrect answers per sentence condition and per grade.

as random intercepts. The fixed effects of the final model explained 14.84% of the variance (marginal R^2; Nakagawa & Schielzeth, 2013) and conditional on the random effects, they explained 21.10% of the variance (conditional R^2; Nakagawa & Schielzeth, 2013). Table 1 only shows the significant results in the output for the final model.

The negative estimate for the intercept indicates that children provided significantly more correct answers (β = -5.58, $p < .001$) to active sentences than to passive sentences. There was thus a significant increase in the number of incorrect answers for passive sentences compared to active sentences, which is noticeable in the strong effect for the predictor *Condition* ($\beta = 1.55$, $p < .001$).

For the verbs, we found a significant effect for "inhalen" (to overtake): the positive estimate suggests that this verb significantly rendered more incorrect answers to the target pictures in comparison to the reference verb "aanrijden" (to hit) in active sentences (β = 2.37, $p < .001$). We found the same effect for the verb "vervangen" (to replace) ($\beta = 2.69$, $p <$ .001).

The positive estimate for the predictor *Grade* (with high grade as reference level) suggests that children in the lower grade significantly provided more incorrect answers in the reference level than children in the higher grade (β 1.80, $p < .001$) (see also Figure 3 for the percentages).

Though there was a significant interaction between *Condition* * *Verb* ($p < .001$) (obtained by the *drop1* function), this was not the same for all levels of the predictor *Verb*. That is, we found a significant interaction between *Condition* and the verb "blokkeren" (to block");

Table 1: Final model output with the predictors *Condition* (active condition as reference level), *Verb* (aanrijden as reference level) and *Grade* (high grade as reference level). We only report the significant results in this table (see Appendix S3 for the full model output).

Summary of the fixed effects in the multilevel logit model (N = 8278; **log-likelihood** = -1594.9) **Fixed effects**	β-coefficient	SE	Z-value	p-value
(Intercept)	-5.58	0.51	-10.89	<.001***
Condition (passive)	1.55	0.40	3.86	<.001 ***
Verb (inhalen)	2.37	0.58	4.06	<.001 ***
Verb (vervangen)	2.69	0.58	4.62	<.001 ***
Grade (low)	1.80	0.41	4.37	<.001 ***
Condition (passive) * Verb (blokkeren)	-1.89	0.56	-3.37	<.001 ***
Condition (passive) * Verb (inhalen)	-1.12	0.46	-2.41	<.05 *
Condition (passive) * Verb (roepen)	2.11	1.06	1.98	<.05 *
Condition (passive) * Verb (volgen)	2.30	0.81	2.82	<.01 **
Verb (groeten) * Grade (low)	1.52	0.70	2.18	<.05 *

the negative estimate suggests that, in contrast with the active condition, the number of incorrect answers for this verb in the passive condition was lower than for the reference verb "aanrijden" (to hit) (β = -1.89, $p < .001$) (see Figure 3). Moreover, the negative estimate of the interaction between *Condition* and the verb "inhalen" (to overtake) suggests that the effect of condition was smaller for this verb than for the reference verb (β = -1.12, $p < .05$). We also found a significant interaction between *Condition* and the verb "roepen" (to call) ($\beta = 2.11$, $p < .05$). As the estimate is positive, the effect of Condition is larger for this verb than for the reference verb "aanrijden". We observed the same effect for the significant interaction between *Condition* and the verb 'volgen' (to follow) ($\beta = 2.30$, $p < .01$): here, the increase in the number of incorrect responses in the passive condition was also larger in comparison with the reference verb (see Figure 3).

There was also a significant interaction between the verb "groeten" (to greet) and *Grade*; the positive estimate indicates that children in the lower grade provided more incorrect answers for this verb compared to the reference level than children in the higher grade (β = 1.52, $p < .05$).

Note that we did not find a significant interaction between *Condition*Grade*, though, descriptively, it seems that children in the lower grade provided more incorrect answers in passive sentences than in active sentences compared to children in the higher grade (see Figure 3). Due to the variance in the many words that we used, the interaction between our condition and the two grades might have not been statistically significant.

Because the effect for each verb on the outcome variable was compared to only one verb

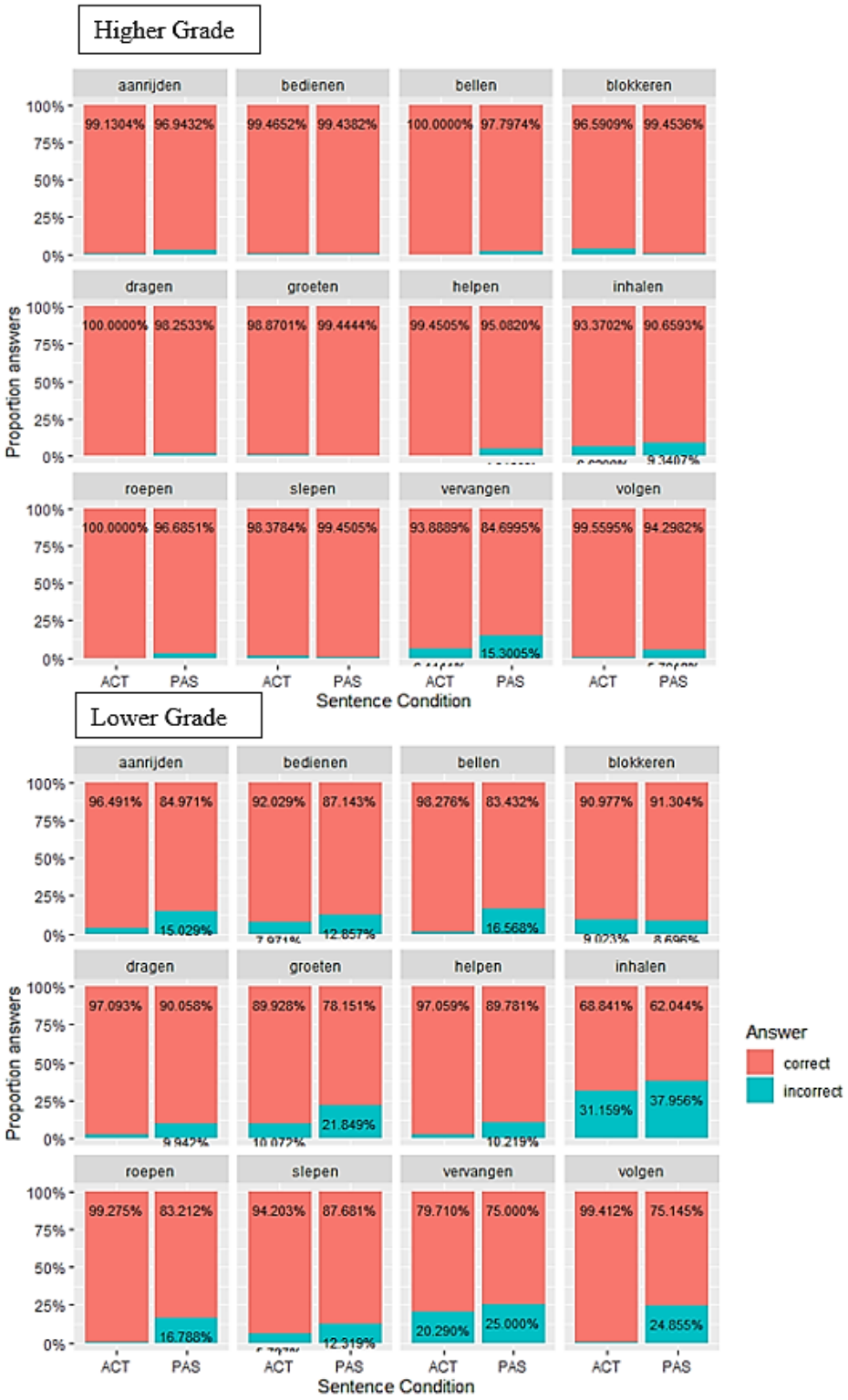


Figure 3: The proportion of correct and incorrect answers to each verb per sentence condition per grade. The upper bar plot shows the higher grade and the lower bar plot the lower grade. The pink bars indicate the correct answers and the blue bars the incorrect answers.

(the reference level "aanrijden" [to hit]), we performed pairwise comparisons using the *emmeans* function (from the package 'emmeans', Lenth 2019) to investigate whether the verbs significantly differed from each other (see Table 2). From Table 2, it seems that the verbs "inhalen" (to overtake) and "vervangen" (to replace) were indeed more difficult to comprehend for the children compared to the other ten verbs (see also Figure 3).

Table 2: Pairwise comparisons of the verbs. In the first column, the first verb significantly rendered more correct answers than the second verb. In the second column, the second verb produced more correct answers than the first verb.

More **correct** answers for the **first** verb than for the **second** verb		More **correct** answers for the **second** verb than for the **first** verb	
aanrijden – inhalen:	Z= -5.47, p <.001	inhalen – slepen:	Z= 5.22, p <.001
aanrijden – vervangen:	Z= -5.19, p <.001	inhalen – roepen:	Z= 4.92, p <.001
bedienen – inhalen:	Z= -5.07, p <.001	vervangen – volgen:	Z= 4.92, p <.001
bedienen – vervangen:	Z= -4.85, p <.001		
bellen – inhalen:	Z= -5.87, p <.001		
bellen – vervangen:	Z= -5.65, p <.001		
blokkeren – inhalen:	Z= -4.72, p <.001		
blokkeren – vervangen:	Z= -4.46, p <.001		
dragen – inhalen:	Z= -6.27, p <.001		
dragen – vervangen:	Z= -6.02, p <.001		
groeten – inhalen:	Z= -4.41, p <.001		
groeten – vervangen:	Z= -4.18, p <.01		
helpen – inhalen:	Z= -5.29, p <.001		
helpen – vervangen:	Z= -5.04, p <.001		
roepen – vervangen:	Z= -4.74, p <.001		
slepen – vervangen:	Z= -4.96, p <.001		

Discussion

In this study, a group of 200 Dutch monolingual Dutch children indicated whether our 208 color drawings depicting transitive action events could be interpreted equally well with an active and a passive sentence. Generally, we found that the children provided 93.02% correct answers to the pictures when an active or a passive sentence was used to describe a transitive event in a picture. In total, there were only 6.98% incorrect answers to the pictures. These results imply that our pictures can be interpreted with both transitive structures and are suitable for language studies with (Dutch) children.

Though the children performed well on the task, we specifically investigated what factors contributed to incorrect answers. This is important to know as these factors may provide

valuable information for other researchers who want to use our picture set in their study. There were several factors that may have contributed to incorrect answers: (1) the use of an active or a passive sentence to describe a picture; (2) the verb that was used to describe the transitive action in a picture; (3) and the age of the children that we tested (we tested children in two different grades, a lower [6-7 years] and a higher grade [11-12 years]).

The strong main effect for our predictor *Condition* indicates that the children had more difficulty interpreting a picture when a passive sentence was used to describe a transitive event than when an active sentence was used. From early on, children receive little input to learn the passive structure. For instance, Gordon and Chafetz (1990) only found four full English passives in a corpus of 87,000 child-directed utterances. Moreover, it takes a while before school-age children (6 years and older) have acquired an adult-like interpretation of complex structures, such as passives (Montgomery et al., 2017). More importantly, all of our sentences were reversible passives. That is, the agent and patient of the passive structure can be exchanged while the sentence remains correct (e.g., "De bakker wordt gegroet door de leraar" - "The baker is being greeted by the teacher". As mentioned in the introduction, reversible passives are challenging for children because they have difficulty interpreting who does what to whom (see for a review Armon-Lotem et al., 2016). Non-reversible passives are easier to comprehend because children can rely on world knowledge to infer the meaning of the passive structure (e.g., The gras is being eaten by the cow). Researchers who intend to use our materials to investigate transitive comprehension in (Dutch) children should thus be aware of the fact that our pictures only consist of reversible passives. If our materials included non-reversible passives, it is possible that we might have observed a less strong effect for *Condition*.

We also tested to what extent the verb, describing the transitive event on our pictures, affected the number of incorrect interpretations of the transitive action. We found that the verbs "inhalen" (to overtake) and "vervangen" (to replace) produced significantly more incorrect answers to the pictures compared to our reference verb "aanrijden" (to hit) in active sentences. There could be several explanations for why these two verbs were significantly more difficult for children to comprehend than the other ten verbs that we used in our stimuli. We believe that the combination of these verbs together with how they were depicted on our pictures might have forced the children to make more inferences than on pictures that depicted concrete verbs (e.g., the verb "dragen" [to carry]). For instance, pictures with the verb "inhalen" were drawn from a bird's eye view perspective (see Figure 1 for an example). The bird's eye view may have hindered children's comprehension of "inhalen" because it might have been difficult to parse what is overtaking what. Even though we used arrows to indicate what object was performing the act of overtaking another object, it is likely that using arrows to indicate a motion might have been too abstract for the children. In this case, they do not only have to infer the meaning of the verb, but also the meaning of a "static" arrow that should depict a motion. The same argumentation can be made for the verb "vervangen" (see Appendix S2). Not only is "vervangen" a rather abstract verb, its depiction on our pictures may have also contributed to difficulties in processing the transitive event. More specifically, we depicted the transitive action of "replacing" with a red cross. For instance, for the sentence, "De ambulance vervangt de bus" ["The ambulance replaces

the bus"], we put a red cross through "the bus"; this was supposed to indicate that this object was replaced by the other object that did not have this cross (see Appendix S2). It is very likely that the combination of the picture and the verb may thus have caused processing difficulties for the children. Moreover, it could be possible that the children in our study might not have been familiar with both verbs, as these verbs are learned relatively late ("inhalen" is learned at 8;61 years and "vervangen" at 9;06 years, see Brysbaert et al., 2014). In our exploratory analyses, where we tested whether there was an interaction between *Verb* and AoA (age of acquisition), we did not find a significant interaction between these two variables. However, we still believe that some children (especially the younger ones) might not have acquired these two verbs yet at the time of testing. For this reason, this might have increased effortful processing of the pictures that contained these two verbs.

We also anticipated that comprehending some verbs could be even more difficult when they were used in passive sentences. Two verbs produced more incorrect answers to the pictures when a passive sentence was used compared to the reference verb "aanrijden" (to hit) in active sentences, namely, "roepen" (to shout) and "volgen" (to follow). Although "roepen" and "volgen" are learned somewhat early by Dutch children, ("roepen" is learned at the age of 4;96 years and "volgen" is learned at 5;84 years, Brysbaert et al., 2014), there could be other explanations for why these verbs produced significantly more incorrect answers to the pictures when they were used in a passive sentence. One could argue that the verb "roepen" (to shout) is a so-called *action verb*, which implies that this verb, ideally, appears in a sentence where the agent is in the subject position and the patient is in the direct object position of a sentence (Lempert, 1989). In other words, it may be that this verb is tightly linked to an argument structure that is compatible with how an active sentence is formed. Although "roepen" can equally occur in an active and a passive sentence, the verb's syntactic preference (i.e., verb bias, see Peter et al., 2015) may affect which syntactic structure children (and other speakers) expect the verb to occur in. What is more, it is likely that children are more frequently exposed to this verb in active sentences than in passive sentences. Thus, when "roepen" occurred in the unexpected passive sentence, the children might have had difficulty to parse who called whom in the picture (i.e., it is likely that the first noun was interpreted as an agent instead of a patient, leading to an incorrect answer for the passive interpretation of a picture). For "volgen" (to follow), the situation is a bit different. We believe that the pictures containing this verb might have caused some confusion for the children. In the pictures with the verb "volgen", we used an arrow to indicate what followed what. For instance, for the sentence "the car follows the truck", we used an arrow to display the direction of the transitive event (an arrow was placed between the agent and the patient, see Appendix S3). The arrow might have caused participants to automatically interpret the first noun as the agent ("the car") followed by the patient ("the truck"). In the case that the picture was described with a passive instead, children had to disregard the direction of the arrow, since the patient was mentioned first. For this reason, we assume that, since the arrow might have prompted an active interpretation of the transitive event, children might have had difficulty to comprehend what followed what when a passive sentence was used, leading to more incorrect answers for passives as opposed to actives.

We found the opposite effect for the verbs "blokkeren" (to block) and "inhalen" (to over-

take): it seems that these verbs produced significantly more correct answers in passive sentences compared to our reference verb "aanrijden" (to hit) in active sentences. The pictures in which these verbs were depicted always consisted of an inanimate agent and an inanimate patient (e.g., "De vrachtwagen haalt de auto in" – The truck overtakes the car; "De vrachtwagen blokkeert de auto" – The truck blocks the car). Several studies (see Bock, 1986; Gàmez & Vasilyeva, 2015) have shown that people have the tendency to produce passive sentences over active sentences when they encounter an inanimate actor and an animate undergoer. In our case, sentences with both transitive verbs ("blokkeren" and "inhalen") had an inanimate agent and patient. However, it could be the case that when children encountered the first inanimate noun, they might have activated a passive sentence over an active sentence, since the grammatical subject of the sentence was inanimate. Possibly, this could have facilitated the processing of pictures that contained these verbs.

Lastly, as expected, we found that the children in the lower grade provided more incorrect answers to the pictures (either described with active or passive sentences) than the children in the higher grade. This means that the older children were better at matching our pictures to the correct transitive structure (i.e., older children had a better comprehension of transitive structures than younger children). Additionally, we found a significant interaction between the verb "groeten" (to greet) and *Grade*: children in the lower grade significantly made more errors in interpreting pictures with this verb in active sentences compared to children in the higher grade. An explanation for this could be that "groeten" is a rather formal word that young children would not necessarily use themselves. More importantly, "groeten" is acquired at the age of 6;71 years (see Brysbaert et al., 2014), which implies that the children in the low grade might not have been familiar with this verb yet (or they might be familiar with this verb only in specific contexts), while the older children might have already known this verb. Another explanation could be that the transitive action of this verb was probably not clear on its pictures. The pictures with the verb "groeten" have two animate objects where the agent is holding one hand up (indicating the action of "to greet") while the patient is portrayed statically on the picture. It could be that the action event of the agent greeting the patient was not depicted explicitly enough, since the only indication for it is the agent's hand being held up. Children in the lower grade might have overlooked this important feature while trying to comprehend pictures with the verb "groeten", leading to incorrect answers.

Limitations and Future Directions

Though we chose to construct our picture set for Dutch language users, since, to our knowledge, no such set exists yet, we admit that this also comes with some disadvantages. As Dutch is only spoken in a few countries (the Netherlands, Belgium, and former Dutch colonies, such as Surinam, Curaçao, the country of Sint Maarten and Aruba), the use of this picture set may be limited. Of course, the transitive verbs can be translated to English (and other languages), which allows these pictures to be used in studies involving speakers of English. However, ideally, before using it in studies with, for instance, English speakers, one would first have to investigate whether the English variant of our picture set produces approximately the same number of correct interpretations of the transitive events in our pictures

(either described with an active or a passive sentence).

We are aware that the current study addresses verb frequency and the age of acquisition of the verbs descriptively. Because our materials were initially developed for late L2 learners of Dutch, we used verbs that were familiar to them. Prior to conducting this study, we did not test whether the children also knew these verbs and whether they had already acquired them or not (as we were mainly interested in whether children were able to interpret our pictures equally well with an active and a passive structure). For this reason, if researchers intend to use our materials to test, for instance, the online processing of transitive structures, it should be noted that lexical frequency may affect sentence processing (Huizeling et al., 2022). We emphasize, however, that the children only provided 7% incorrect answers, which means that we should not overlook the fact that our materials did largely elicit correct interpretations of both active and passive descriptions of transitive events (93%).

Another possible limitation is that we used our picture set to investigate the comprehension of active and passive events in a group of children who were already relatively far into their L1 syntax acquisition. As we know that children start understanding passives around the age of 3, we do not know whether this picture set is also suitable for the youngest group of language users. Because of the findings in the current study (it turns out that our picture set contains a few transitive verbs that are learned quite late by Dutch children), we believe that our picture set might probably be too complex for children younger than the age of 6. For this reason, we advise researchers to use our picture set for children in the same age group as the one we tested (6 – 12 years old).

Throughout this article, we used studies in structural priming involving children as a possible research domain for which our picture set can be used. However, we would like to stress that our picture set may be suitable for other research domains too, particularly in domains where sentences are elicited. Since our picture set was initially designed for longitudinal designs, our set can be useful to test the acquisition of transitive structures in children, especially the use of passives. Research has shown that passive sentences are quite difficult for children to such an extent that, even when they have already learned the passive structure, their passive sentence production remains effortful (Messenger et al., 2012). As our norming study showed that the younger children made more errors in *comprehending* pictures that were described with a passive sentence compared to the older children, future studies could investigate whether younger children also make more errors while *producing* passive sentences, and at what age the number of errors decreases. Apart from using our picture set to investigate the acquisition of transitives in typically developing children, our stimuli can also be used in studies that involve children with a delayed language development (e.g., children on the autism spectrum) as acquiring passives has been found to be even more effortful for them compared to children without such a delay (Ambridge et al., 2021). Moreover, because our picture set uses a limited vocabulary (verbs and nouns), we believe that it may not be too taxing on the language abilities of, for instance, people with a language impairment. We also think that, for the same reason, our picture set can be used in studies with healthy elderly people.

Lastly, we tested the comprehension of transitive structures, but our pictures could also be used in other modalities such as language production (spoken as well as written, see Van

Lieburg et al. [2022], who already used our picture set to investigate L2 syntax production). An interesting question could be whether language users spontaneously describe a transitive event with the more complex passive sentence structure or whether they will almost always produce active sentences because they are easier. This question is especially interesting for written language production, as research has shown that people use passives more spontaneously during writing than in spoken language (Hinkel, 2004).

Conclusion

This study set out to test whether our picture set, consisting of transitive events with different verbs, could be interpreted equally well with active and passive sentences amongst Dutch children who differed in age. Even though the children provided 93.02% correct answers, we found that they provided significantly more incorrect answers to pictures that used a passive sentence to describe a transitive event than to pictures that used an active sentence. Moreover, we found that some verbs ("inhalen" and "vervangen") significantly produced more incorrect answers regardless of whether they were presented in an active or a passive sentence compared to the other verbs that occurred in our pictures. Moreover, we observed that the children had difficulty in comprehending some verbs when they were presented in a passive sentence, while other verbs rendered more correct answers when the transitive event was described with a passive sentence. We also found that, in general, younger children had more difficulty in interpreting our pictures, either described with an active or a passive sentence. Only one verb seemed to be particularly difficult for the younger children, namely, the verb "groeten" (to greet). We believe that the relatively late age of acquisition of this verb (6;71 years) and the unclear depiction of the transitive action "groeten" on pictures could have hindered the young children's comprehension of this verb.

Altogether, our study shows that the colored pictures are mostly clear in terms of how the transitive events are displayed and how well they can be interpreted with either an active or a passive sentence by young children.

Transparency and Openness

The color drawings, all data, analysis code, and experimental lists are available at https://tinyurl.com/DutchNormingStudy. This study's design and the analyses were not pre-registered.

Acknowledgments

This research was funded by a BOF DOCPRO4 grant from the University of Antwerp (nr. 37003).
OSF repository: https://tinyurl.com/DutchNormingStudy.
We have no known conflict of interest to disclose.

Literature

Ambridge, B., Bidgood, A., & Thomas, K. (2021). Disentangling syntactic, semantic and pragmatic impairments in ASD: Elicited production of passives. *Journal of child language, 48*(1), 184-201. https://doi.org/10.1017/S0305000920000215

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of child language, 42*(2), 239 273. https://doi.org/10.1017/S030500091400049X

Armon-Lotem, S., Haman, E., Jensen de López, K., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., ... & van Der Lely, H. (2016). A large-scale cross-linguistic investigation of the acquisition of passive. *Language acquisition, 23*(1), 27-56.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

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

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition, 112*(2), 300- 317. https://doi.org/10.1016/j.cognition.2009.05.013

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), B13-B25. https://doi.org/10.1016/S0010-0277(99)000815

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150*, 80-84. https://doi.org/10.1016/j.actpsy.2014.04.010

Czypionka, A., Spalek, K., Wartenburger, I., & Krifka, M. (2017). On the interplay of object animacy and verb type during sentence comprehension in German: ERP evidence from the processing of transitive dative and accusative constructions. *Linguistics, 55*(6), 1383-1433. https://doi.org/10.1515/ling-2017-0031

Favier, S., Wright, A., Meyer, A., & Huettig, F. (2019). Proficiency modulates between- but not within-language structural priming. *Journal of Cultural Cognitive Science, 3*(1), 105–124. https://doi.org/10.1007/s41809-019-00029-1

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715-736. https://doi.org/10.1006/jmla.1994.1034

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied Longitudinal Analysis.* John Wiley & Sons.

Fox, D., & Grodzinsky, Y. (1998). Children's passive: A view from the by-phrase. *Linguistic inquiry, 29*(2), 311-332.

Gámez, P. B., & Vasilyeva, M. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology, 138*, 15-30. https://doi.org/10.1016/j.jecp.2015.04.009

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological science, 17*(8), 684 691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Goldwater, M. B., Tomlinson, M. T., Echols, C. H., & Love, B. C. (2011). Structural priming as structure-mapping: children use analogies from previous utterances to guide sentence production. *Cognitive science, 35*(1), 156-170. https://doi.org/10.1111/j.1551-6709.2010.01150.x

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*(3), 227-254. https://doi.org/10.1016/0010-0277(90)90058-R

Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language teaching research, 8*(1), 5-29. https://doi.org/10.1191/1362168804lr132oa

Hirsch, C., & Wexler, K. (2006). Children's passives and their resulting interpretation. In *The proceedings of the inaugural conference on generative approaches to language acquisition–North America, University of Connecticut Occasional Papers in Linguistics* (Vol. 4, pp. 125-136). Storrs, CT: University of Connecticut.

Huizeling, E., Arana, S., Hagoort, P., & Schoffelen, J. M. (2022). Lexical frequency and sentence context influence the brain's response to single words. *Neurobiology of Language, 3*(1), 149-179.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of memory and language, 50*(2), 182-195. https://doi.org/10.1016/j.jml.2003.09.003

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior research methods, 42*(3), 643-650.

Lefcheck, J.S. (2016). "piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics." *Methods in Ecology and Evolution, 7*(5), 573- 579. doi: 10.1111/2041-210X.12512

Lempert, H. (1989). Animacy constraints on preschool children's acquisition of syntax. *Child Development*, 237-245. https://doi.org/10.2307/1131088

Lenth, R. (2019). *Emmeans: estimated marginal means, aka least-squares means.* R package version, 1(2).

Lieburg, R. van, Sijyeniyo, E., Hartsuiker, R.J., & Bernolet, S. (2022). *The development of abstract syntactic representations in beginning learners of Dutch.* Manuscript in preparation.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language, 91.* https://doi.org/10.1016/j.jml.2016.03.009

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*(1), 121-136. https://doi.org/10.1016/j.cogdev.2004.08.001

Messenger, K., & Fisher, C. (2018). Mistakes weren't made: Three-year-olds' comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition, 178*, 118-132. https://doi.org/10.1016/j.cognition.2018.05.002

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language, 66*(4), 568-587. https://doi.org/10.1016/j.jml.2012.03.008

Montgomery, J. W., Gillam, R. B., Evans, J. L., & Sergeev, A. V. (2017). "Whatdunit?" Sentence comprehension abilities of children with SLI: Sensitivity to word order in canonical and noncanonical structures. *Journal of Speech, Language, and Hearing Research, 60*(9), 2603-2618. doi: 10.1044/2017_JSLHR-L-17-0025

Muylle, M., Wegner, T. G., Bernolet, S., & Hartsuiker, R. J. (2020). English norming data for 423 short animated action movie clips. *Acta Psychologica, 202*, 102957. https://doi.org/10.1016/j.actpsy.2019.102957

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution, 4*(2), 133 142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language, 81*, 1-15. https://doi.org/10.1016/j.jml.2014.12.002

Pickering, M., & Ferreira, V. (2008). Structural Priming: A Critical Review. *Psychological bulletin, 134*, 427–459. https://doi.org/10.1037/0033-2909.134.3.427

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and language, 39*(4), 633-651. https://doi.org/10.1006/jmla.1998.2592

Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge, 26*.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of child language, 31*(1), 61 99. https://doi.org/10.1017/S0305000903005956

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology, 67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521

Verrips, M. (1996). *Potatoes must Peel. The Acquisition of the Dutch Passive*. Ph.D. Dissertation, University of Amsterdam.

Appendix S1

Transitive Verbs Used in Picture Set.

Aanrijden, bellen, dragen and *volgen* occurred 20 times in the picture set. The rest of the verbs occurred 16 times in the picture set.

Dutch Verb	**English Translation**	**Zipf-frequency (Keuleers et al., 2010)**	**AoA (Brysbaert et al., 2014)**
aanrijden	to hit	3.25	9.28
bedienen	to serve	3.85	8.89
bellen	to call	5.35	6.03
blokkeren	to block	3.67	8.72
dragen	to carry	4.91	5.78
groeten	to greet	4.48	6.71
helpen	to help	5.76	5.40
inhalen	to overtake	2.72	8.61
roepen	to shout	4.53	4.96
slepen	to drag	3.95	8.03
vervangen	to replace	4.41	9.06
volgen	to follow	4.99	5.84

List of Human Professions and Vehicles

Human Professions	**English translation**	**Vehicles**	**English translation**
bakker	baker	ambulance	ambulance
bouwvakker	construction worker	auto	car
dokter	doctor	brandweer	fire truck
journalist	journalist	bus	bus
kok	cook	fiets	bike
leraar	teacher	motor	motor
slager	butcher	tram	tram
zangeres	singer	vrachtwagen	truck

Human professions

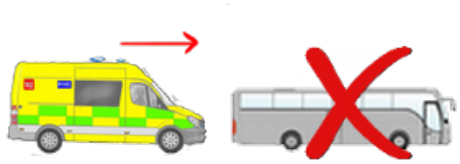
Vehicles

ambulance 	auto 
brandweer 	bus 

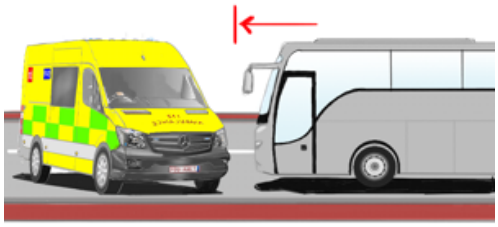
fiets 	motor 
tram 	vrachtwagen 

Appendix S2

Examples of pictures with the abstract verbs ‘vervangen’ [to replace]; ‘blokkeren’ [to block].



An example of a picture that depicts the sentence ‘De ambulance vervangt de bus’ – The ambulance replaces the bus



An example of a picture that depicts the sentence ‘De ambulance blokkeert de bus’ – The ambulance blocks the bus.

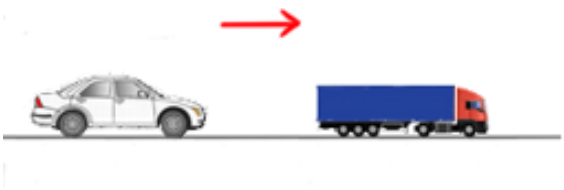
Appendix S3

Table 3: Full output of final model with the predictors Condition (active condition as reference level), Verb (aanrijden as reference level) and Grade (high grade as reference level).

Summary of the fixed effects in the multilevel logit model (N = 8278; **log-likelihood** = -1594.9) **Fixed effects**	β-coefficient	SE	Z-value	p-value
(Intercept)	-5.58	0.51	-10.89	<.001 ***
Condition (passive)	1.55	0.40	3.86	<.001 ***
Verb (bedienen)	-0.67	0.90	-0.74	0.45
Verb (bellen)	-1.57	0.88	-1.78	0.07 .
Verb (blokkeren)	1.21	0.67	1.80	0.07 .
Verb (dragen)	-0.79	0.82	-0.96	0.33
Verb (groeten)	-0.31	0.80	-0.38	0.69
Verb (helpen)	0.16	0.74	0.22	0.82
Verb (inhalen)	2.37	0.58	4.06	<.001 ***
Verb (roepen)	-2.06	1.16	-1.77	0.07 .
Verb (slepen)	0.17	0.75	0.22	0.81
Verb (vervangen)	2.69	0.58	4.62	<.001 ***
Verb (volgen)	-1.58	0.91	-1.73	0.08 .
Grade (low)	1.80	0.41	4.37	<.001 ***
Condition (passive) * Verb (bedienen)	-0.88	0.57	-1.54	0.12
Condition (passive) * Verb (bellen)	1.26	0.72	1.74	0.08 .
Condition (passive) * Verb (blokkeren)	-1.89	0.56	-3.37	<.001 ***
Condition (passive) * Verb (dragen)	-0.08	0.64	-0.12	0.90
Condition (passive) * Verb (groeten)	-0.62	0.54	-1.15	0.24
Condition (passive) * Verb (helpen)	0.21	0.64	0.32	0.74
Condition (passive) * Verb (inhalen)	-1.12	0.46	-2.41	<.05 *
Condition (passive) * Verb (roepen)	2.11	1.06	1.98	<.05 *
Condition (passive) * Verb (slepen)	-1.09	0.56	-1.93	0.05 .
Condition (passive) * Verb (vervangen)	-0.82	0.47	-1.74	0.08 .
Condition (passive) * Verb (volgen)	2.30	0.81	2.82	<.01 **
Verb (bedienen) * Grade (low)	-1.47	0.80	1.84	0.06 .
Verb (bellen) * Grade (low)	0.63	0.62	1.02	0.30
Verb (blokkeren) * Grade (low)	-0.07	0.58	-0.13	0.89
Verb (dragen) * Grade (low)	0.37	0.65	0.57	0.56
Verb (groeten) * Grade (low)	1.52	0.70	2.18	<.05 *
Verb (helpen) * Grade (low)	-0.75	0.55	-1.36	0.17
Verb (inhalen) * Grade (low)	0.38	0.45	0.85	0.39
Verb (roepen) * Grade (low)	0.19	0.59	0.31	0.75
Verb (slepen) * Grade (low)	0.53	0.64	0.83	0.40
Verb (vervangen) * Grade (low)	-0.75	0.45	-1.66	0.09 .
Verb (volgen) * Grade (low)	0.07	0.50	0.15	0.88

Appendix S4

The Transitive Picture for “The car follows the truck”. Note the use of the Arrow.



The arrow indicates the direction of the transitive event. It is likely that people will first process “the car” as the agent, leading to an active interpretation, before processing “the truck”. Confusion may arise when people hear “the truck” first, since the direction of the arrow should then be disregarded.