Published in final edited form as:
Child Dev. 2014 May ; 85(3): 1168–1180. doi:10.1111/cdev.12164.

Toddlers Default to Canonical Surface-to-Meaning Mapping When Learning Verbs

Isabelle Dautriche[*],
LSCP, Ecole Normale Supérieure / CNRS, Paris Maternité Port-Royal (AP-HP), Université Paris Descartes

Alejandrina Cristia[*],
Max Planck Institute for Psycholinguistics, Netherlands

Perrine Brusini,
LSCP, Ecole Normale Supérieure / CNRS, Paris Maternité Port-Royal (AP-HP), Université Paris Descartes

Sylvia Yuan,
University of California, Berkeley

Cynthia Fisher, and
University of Illinois, Urbana-Champaign

Anne Christophe
LSCP, Ecole Normale Supérieure / CNRS, Paris Maternité Port-Royal (AP-HP), Université Paris Descartes

Abstract

Previous work has shown that toddlers readily encode each noun in the sentence as a distinct argument of the verb. However, languages allow multiple mappings between form and meaning which do not fit this canonical format. Two experiments examined French 28-month-olds' interpretation of right-dislocated sentences ($noun_i$-verb, $noun_i$) where the presence of clear, language-specific cues should block such a canonical mapping. Toddlers (N = 96) interpreted novel verbs embedded in these sentences as transitive, disregarding prosodic cues to dislocation (Experiment 1) but correctly interpreted right-dislocated sentences containing well-known verbs (Experiment 2). These results suggest that toddlers can integrate multiple cues in ideal conditions, but default to canonical surface-to-meaning mapping when extracting structural information about novel verbs in semantically impoverished conditions.

One of the most remarkable properties of language is that it is deeply productive, as a result of listeners' remarkable resilience to novelty. For example, even a sentence containing unknown words is fairly interpretable. Faced with "The glop mooped the nuck", we readily assume that there was an agent ('the glop') who performed some causative action

Correspondence regarding this article should be addressed to Isabelle Dautriche, Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France. isabelle.dautriche@gmail.com.
[*]Joint authorship

('mooping') on a patient ('the nuck'). In contrast, had we read "The glop mooped", we would not assume that 'mooping' involved an action performed on a patient. Our ability to abstract such rich structural information is likely informed by multiple mechanisms. In this article, we will show that some of them prevail when toddlers are discovering the meaning of novel verbs in semantically impoverished conditions.

Previous research has uncovered one mechanism that is readily used from early on. The *structure-mapping account of early verb learning* (Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003) proposes that children are biased to use the number of nouns (or Determiner Phrases, DPs) as a cue to verb meaning. Hence, verbs occurring with two DPs (e.g., "He chases it") are automatically associated with a two-participant interpretation, whereas verbs occurring with one DP (e.g., "He smiles") are associated with a one-participant interpretation. A rich body of literature documents a predictable relationship between the syntactic realization of verb arguments and the mental representation of events across languages (for a review, see Levin & Rappaport, 2005). Indeed, even in languages where DPs can be omitted (e.g., saying "chase" for "he chases it"), and thus the alignment between the surface number of DPs and the number of participant-roles in the verb's semantic representation is less regular, children interpreted sentences with familiar verbs in accordance with the number of overt DPs in the sentence (Kannada: Lidz et al. 2003; Turkish: Göksun, Küntay & Naigles, 2008; Mandarin Chinese: Lee & Naigles, 2008). Moreover, this mechanism is at play even in novel verb learning, and may be evident in toddlers as soon as they can process multi-word utterances (e.g., Yuan, Fisher & Snedeker, 2012). For example, two-year-olds who hear a novel verb preceded by one DP and followed by another one readily recover a causative event, in which the first referent is the agent and the second the patient (e.g., Naigles, 1990). A representation of sentences based on the number of DPs may help young children to detect how DPs in different positions map onto thematic roles in their language. For example, in English, it may allow them to establish that a sequence of the form DP-V-DP maps onto the abstract representation DP_{agent}-V-$DP_{patient}$, by learning from sentence-interpretations selected via structure-mapping, or by simply assuming that each DP maps onto a distinct participant role following the canonical Agent-Patient ordering present in 97% of all world languages (Dryers, 2005; Gertner & Fisher, 2012).

Nonetheless, a canonical mapping between the number of DPs and the number of arguments of the verb will not be sufficient to capture the variety of structures present in one's native language. Most languages also allow additional surface-to-meaning mappings. A well-known case is short passive sentences (e.g., "the boy has been selected"), which are fairly common in English (Gordon & Chafetz, 1990) and which denote a two-participant event although only one DP is mentioned ($DP_{patient}$-V). In English as in most languages, passives are marked through the addition of morphemes or whole words, markers that should block a canonical mapping between the number of DPs and the number of participant roles in the semantic representation of the verb. Passives are often late-learned, presumably in part because they are rarer than active sentences (Pinker, Lebeaux & Frost, 1987). Some recent work suggests that only three-year-olds with high vocabularies succeeded in linking a short passive with a novel verb to a relational event involving two participants (Messenger &

Fisher, 2011). Even in full passives where the number of DPs is congruent with the number of arguments of the verb (e.g., “the mouse was eaten by the cat”) but where the thematic roles are not found in their canonical positions ($DP_{patient}$-V-DP_{agent}), some developmental evidence shows that children have difficulty parsing these constructions through age 5 (for English: Maratsos, Fox, Becker, & Chalkley, 1985), though they have some ability to understand passives much earlier (see Bencini & Valian, 2008, and Demuth, Moloi, & Machobane, 2010 for Sesotho). Before this age, morphological cues are not always correctly integrated during sentence interpretation, since English-speaking children tend to provide a canonical interpretation DP_{agent}-V-$DP_{patient}$ to these sentences (Maratsos et al., 1985).

In addition to morphology, languages can recruit prosody to allow additional interpretations beyond those predicted by the structure-mapping account, as is the case with right dislocations such as “He_i's great, Tom_i”. In French, right dislocations serve a wide range of pragmatic functions such as topicalizing or highlighting the referent, shifting topic, and clarifying a referent by making explicit a pronoun the speaker suddenly realizes the audience is unsure of (with the last function enumerated seen as a secondary function of right dislocations in French: Ashby, 1988; Reinhart, 1981). In right-dislocated sentences, a DP surfaces in a prosodically marked constituent placed at the right periphery (the dislocated DP), and is typically resumed within the sentence by a pronoun or a clitic (DeCat, 2007). Hence, one argument of the verb is copied such that the sentence has an extra DP but does not hold an additional participant-role, thus forcing non-canonical interpretations where the number of DPs is different from the number of participant roles in the sentence. For example, in French both "il_i mange, le $lapin_i$" *he_i eats, the rabbit$_i$* (where "le lapin" *the rabbit* is coreferential with the pronoun "il" *he* and hence is doing the eating) and "il_i mange le $lapin_j$" *he_i eats the rabbit$_j$* (where the rabbit is being eaten) are perfectly grammatical, and have exactly the same superficial structure in terms of morphemes and word order. However, the underlying structure is very different: whereas the former contains only one participant and can therefore be used with intransitive verbs (e.g., "il_i rigole, le $lapin_i$" *he_i laughs, the rabbit$_i$*), the latter contains two participants and can only be used with transitive verbs (i.e., *"il_j rigole le $lapin_j$" **he_i laughs the rabbit$_j$*).

French right-dislocations are particularly interesting because they are present even in the speech of two-year-olds (Dautriche, 2012). Since these structures are present in children's output, it is expected that they will be able to process the prosodic cues that uniquely distinguish them. It is not surprising that toddlers pick up on this surface pattern early on, since it is very frequent (DeCat, 2007; about 5% of *all* sentences uttered by parents are dislocated, Dautriche, 2012), and it is unambiguously cued by prosody, which is salient even for infants (Christophe, Gout, Peperkamp & Morgan, 2003; Christophe, Millotte, Bernal & Lidz, 2008; Gerken, Jusczyk & Mandel, 1994). In right dislocations, the verb is followed by an intonational phrase boundary, and the prosodic contour of the dislocated DP copies the prosodic contour of the rest of the sentence (see Ashby, 1994 and Rossi, 1999). This is particularly obvious in interrogative sentences. For example, in “il a mangé, le chien?” *it ate, the dog?* (Did the dog eat?) there would be two pitch rises, one on ‘eat’ and another on the dislocated element ‘dog.’ In contrast, in the transitive version of this sentence, “il a mangé le chien?” *it ate the dog?* (Did something eat the dog?) only a single pitch rise at the

end of the sentence would be observed. However, to understand these structures, the listener must overcome her bias to map each DP onto an argument of the verb, since the post-verbal DP of right dislocated sentences cannot be interpreted as having a participant role separate from that assigned to the pronominal DP.

The question that arises, then, is: How might toddlers interpret a novel verb in right-dislocated sentences where multiple surface-to-meaning mappings are possible? One possibility is that toddlers exploit the prosodic marking of these sentences, and readily use this language-specific cue to discard a one-to-one mapping between the number of DPs and the number of verb arguments. Alternatively, inhibiting this basic strategy may be costly, and might fail in a demanding situation where word meanings are uncertain. We distinguish between these two hypotheses by using right-dislocations in a verb-learning task requiring toddlers to go beyond a canonical surface-to-meaning mapping interpretation. Previous studies have shown that English-speaking 25-month-olds can already use language-specific knowledge such as word order and morphology to distinguish between intransitive sentences with two DPs conjoined in a single argument position (“the duck and the bunny are gorping”) and transitive sentences (“the duck is gorping the bunny”) (Naigles, 1990). Hence, we predicted that, if two-year-olds can readily abandon a canonical surface-to-meaning mapping procedure when the situation calls for it, then French-learning toddlers exposed to a novel verb in a right-dislocated context should be able to resist the canonical interpretation, identifying the pronominal DP and the right-dislocated DP as a single verb argument, and thus treat this verb as intransitive. In contrast, should toddlers default to a canonical surface-to-meaning mapping, they might (incorrectly) assume that both DPs map onto distinct arguments of the verb and hence entertain a transitive interpretation of the novel verb.

Experiment 1: Using prosodic cues with novel verbs

As in previous work focusing on the role of syntactic structure in verb learning (Arunachalam & Waxman, 2010; Yuan & Fisher, 2009), we presented toddlers with dialogues introducing a novel verb *’daser’* in either transitive sentences (“Il a dasé le bébé” *He dased the baby*), or intransitive sentences (“Il a dasé” *He dased*) (See Figure 1 for sample dialogues). The dialogue video presented two women conversing such that no visual information about the verb’s meaning was provided. During the subsequent test phase, toddlers watched two videos presented side-by-side: a two-participant event (one girl swinging another girl’s leg) and a one-participant event (a girl making circles with her arm) and were prompted to look at the video that matched ‘daser’, using intransitive constructions that are adequate both for intransitive and transitive verbs (“Regarde celle qui dase!” *Look at the one who dases*!). In addition, there was a right-dislocated condition, in which the novel verb was embedded in right-dislocated sentences in the dialogue (“Il a dasé, le bébé” *He dased, the baby*). Finally, there was a control condition in which toddlers were not exposed to the novel verb before the test phase.

Following Yuan and Fisher (2009), we predicted that 28-month-olds, well-seasoned in these tasks (cf. Naigles, 1990), who heard the novel verb in *transitive* sentences should look more at the two-participant event than toddlers who heard the novel verb in intransitive sentences.

As for the *intransitive* dialogue condition, evidence from past studies suggests that toddlers may not have any preference for one or the other event, because an intransitive sentence could also refer to the role of one participant in a two-participant event, in which case both visual alternatives would be correct (Arunachalam & Waxman, 2010; Noble, Rowland & Pine, 2011; Yuan, et al., 2012). The critical case concerned the right-dislocated sentences, where clear prosodic cues to a dislocation should lead 28-month-old toddlers, expected to have a command of the right-dislocation (cf. Dautriche, 2012), to abandon the canonical surface-to-meaning mapping. If toddlers in this condition behave like those in the intransitive condition (e.g., they look less at the two-participant event than toddlers in the transitive condition), this would indicate that toddlers can readily use the prosodic cues marking a right-dislocation to adopt an interpretation that conflicts with a canonical mapping. In contrast, if toddlers in this condition behave like those in the transitive condition (e.g., they look more at the two-participant event than toddlers in the intransitive condition), this would be an indication that a canonical mapping of the number of DPs onto distinct arguments of the verb is strong enough to overcome a dramatic prosodic boundary at this age.

Method

Participants—Sixty-four 28-month-olds (mean = 28.15; min = 27.13; max = 29.13; females = 30) participated in the study. Sixteen toddlers were randomly assigned to each of the four experimental conditions (transitive, dislocated, intransitive and control). Six additional toddlers were not included because of fussiness ($n = 4$), experimenter error ($n = 1$) or because they knew one of the actresses in the video ($n = 1$). All toddlers were recruited from Paris and were acquiring French as their first language with less than 30% exposure to another language. Parents completed the French version of the MacArthur Communicative Development Inventory (Kern S, 2007). Vocabulary ranged from 71 to 614 words (mean = 392) and there were no differences in vocabulary between conditions ($F < 1$).

Apparatus, procedure, and stimuli—Toddlers sat on their parent's lap 70cm away from a 42" television screen. Two videos of 30×30cm were displayed simultaneously on the right and left side of the screen and separated by 30cm. Toddlers' eye movements were recorded by a hidden camera above the screen. The caregiver and the experimenter listened to acoustic masking throughout the test.

The procedure was similar to that of Yuan and Fisher (2009). The experiment began with two practice trials with familiar verbs, one involving transitive verbs ("pousser" *to push* or "porter" *to carry*) and one involving intransitive verbs ("danser" *to dance* or "marcher" *to walk*). The practice trials consisted of two 8s test events in which a synchronized pair of videos was presented on both sides of the screen along with a soundtrack that encouraged them to look at one of the videos. In order to give toddlers sufficient time to inspect the action of the videos, a preview period preceded each trial. Each video of the test event was seen alone for 5s before the test event started as shown in Figure 2. These practice trials familiarized toddlers with the procedure, "teaching" them that they would pre-view two events on the screen, and then would see the events paired, while hearing an audio track that matched only one of them. The practice trials did not include a dialogue phase.

After the two practice trials, toddlers in the transitive, intransitive, and right-dislocated conditions saw the dialogue phase appropriate for their condition. Two four-sentence dialogue video clips of 28s separated by a 3s interval were presented simultaneously on both sides of the screen. Thus, each toddler was exposed to 8 transitive sentences (e.g., "il dase le bébé" *he is dasing the baby)* or 8 dislocated sentences (e.g., “il dase, le bébé" *he is dasing, the baby*) or 8 plain intransitive sentences (e.g. “il dase*" he is dasing*). Three types of dialogue videos (transitive, dislocated and intransitive) were recorded, all of them showing a conversation between two women speaking in child-directed speech. The actors (the third and the last authors) are highly trained in prosody, and were able to produce the typical prosodic patterns associated with each of these interpretations in a very natural fashion. Acoustic analyses confirmed that there was a clear prosodic boundary between verb and final DP in the dislocated but not the transitive sentences (preboundary syllable lengthening: M_{trans}= 225ms vs. M_{disloc}= 633ms, $t(14)= -10.31$, $p < .001$; pitch drop: M_{trans}=−39Hz vs. M_{disloc}=−119Hz, $t(14) = 3.05$, $p < .05$). Note that there was no pause between the verb and the post-verbal DP; thus to differentiate the transitive from the dislocated sentences, infants had to interpret the prosodic structure of the sentence rather than simply discarding any material occurring after a substantial pause. In a pilot experiment, we asked twenty naive adults to decide who was performing the action in each of these test sentences. French adults interpreted the transitive and dislocated sentences correctly over 90% of the time.

All toddlers in the three experimental conditions (transitive, right-dislocated and intransitive) then saw the same test phase, which followed the same procedure as the practice trials (Figure 2). Specifically, toddlers saw two novel actions side by side, one with a single participant and another with two participants (the same videos as in Yuan & Fisher, 2009). They heard a series of sentences featuring the novel verb in an intransitive construction in all conditions: (“Hey elle dase! Tu la vois qui dase? Regarde celle qui dase! Elle est où celle qui dase?" *Hey she is dasing! Do you see the one who is dasing? Look at the one who is dasing! Where is the one who is dasing?*). The auditory stimuli in the test trial were identical across all conditions, and had been recorded by the last author. The order of the practice trials and the left/right position of the target video were counterbalanced across participants, within each of the three conditions.

To these three experimental conditions, we added a fourth control condition where toddlers saw the two practice trials directly followed by a test trial with **no** dialogue exposure. Thus during the test trial, toddlers had no opportunity to infer aspects of the verb’s meaning from the structure of dialogue sentences, but underwent the same testing procedure: they had to look for the novel verb ’daser’ while watching the one-participant and the two-participant videos. This condition gives toddlers the same urging that toddlers get in the three experimental conditions to choose one of the videos, hence revealing their baseline preference.

Analyses—Videos were digitized at 25 frames per second and looking times to the left, the right, and away from the screen were coded from muted video frame by frame using SuperCoder (Hollich, 2005). Reliability was assessed for 9% of the data (6 toddlers), the two coders agreed on 97% of the frames (Cohen's Kappa: 0.94). Looking times were averaged across the two test events. Data from test events where toddlers looked away more than half

of the event duration were considered as missing (4 test events overall, 2 in the control condition and 2 in the transitive condition). Preliminary analysis of the proportion of time spent looking away revealed no significant effect of condition ($F < 1$) indicating that toddlers tended to look away equally in the four experimental conditions. Since the looking time to the two-participant and the one-participant events are complementary, we based our analysis on the proportion of looking time to the two-participant event over the total looking time (i.e., including the time spent looking away). For the sake of completeness, Figure 3 shows the proportion of looking time to the one- and the two-participant events as well as the time spent looking away.

Results

As Figure 3 shows, looking times were affected by dialogue condition. An ANOVA with a between-subject factor Condition (Transitive, Intransitive, Dislocated or Control) revealed a significant effect of Condition on proportion of looking time to the two-participant event ($F(3,60) = 2.98$, $p < .05$), showing that toddlers in different conditions behaved differently.

To probe this effect, we conducted a set of restricted comparisons between pairs of conditions with a between-subjects factor Condition. For the comparison between transitive and intransitive conditions, there was a significant effect of Condition ($F(1, 30) = 5.25$ $p < .05$) reflecting the fact that toddlers who heard the novel verb in the transitive dialogue looked longer at the two-participant event relative to toddlers in the intransitive condition. This replicates prior results with English-speaking two-year-olds (Yuan & Fisher, 2009). Turning to the control condition, comparisons between the transitive and control conditions as well as the intransitive and control conditions did not result in a significant effect (both $F < 1$) suggesting that both the transitive and intransitive structures contributed to making the experimental conditions different from each other (in the control condition, toddlers showed no intrinsic preference for one event over the other, as the difference in looking time between the two- and one-participant events did not differ from zero: $t(15) = 0.81$, $p > .4$).

Regarding our main question, the comparison between dislocated and transitive conditions revealed no effect of Condition ($F(1,30) = 0.45$ $p > .5$): toddlers in the right-dislocated condition looked at the two-participant video just as did toddlers in the transitive condition. Finally, the comparison between dislocated and intransitive conditions revealed a significant effect of Condition ($F(1, 30) = 7.73$, $p < .01$) confirming that toddlers did not interpret dislocated sentences as being intransitive.

Discussion

In a verb-learning task, two-year-olds interpreted a novel verb appearing in right-dislocated sentences as referring to a two-participant event, similarly to how they encoded a novel verb in a transitive context. They behaved as if they ignored the prosodic information and interpreted each DP as a separate argument of the verb, even in the presence of clear cues indicating that this interpretation should not be adopted.

To understand these results, we need to consider two potential interpretations. The first is that toddlers simply do not know the prosody of right-dislocation. The adults in our pilot

study correctly interpreted the right-dislocated structures used in this experiment, but it is possible that *two-year-olds* have not, in fact, learned to use these prosodic cues at all. We address this question in Experiment 2.

A second, radically different possibility, is that toddlers do know the prosody of right-dislocation, but failed to use prosody in this experimental setting because the verb was unknown, and perhaps because right-dislocation needs to be pragmatically licensed. In natural exchanges between children and their parents recorded in CHILDES, right-dislocations seemed to occur when the event being described was visible; for example, the sentence ”il croque, le crocodile” *it munches, the crocodile* was uttered while looking at a picture book showing a crocodile eating some animal (Lyon Corpus, Marie, Session 31, Demuth and Tremblay, 2008). Additionally, the dislocated component is often the discourse topic; therefore, right-dislocations are more likely to occur in the presence of a clear sustained topic. For instance, parents can focus the attention of the child on the topic before using it in a right-dislocated sentence: ”Et qu’est ce qu’il sort du four? Un gros gâteau. T’as vu? Il a gonflé, le gâteau” *What does he take out from the oven? A big cake! Do you see that? It puffed up, the cake!* (Lyon Corpus, Marie, Session 27a, Demuth and Tremblay, 2008). In Experiment 1, we followed previous research investigating toddlers’ extraction of syntactic properties from dialogues because this procedure isolates the impact of the sentence syntactic structure on verb interpretation. However, this could pose a problem for right-dislocations specifically. Notice that in these dialogues the action, and not the dislocated DP, is the topic; and no visual referent matches the dislocated DP. Hence, one could argue that right-dislocation is not pragmatically licensed in these dialogues (however see Ashby, 1988, for evidence that the pragmatic context of right-dislocation is variable). It might be that toddlers are sensitive to the pragmatic context, and their systematic transitive interpretation or the right-dislocated sentences (ignoring the prosody) followed from their perception that dislocation is unlikely in that context. In order to entertain this explanation, however, one must show that toddlers can indeed process the prosodic cues of right dislocation.

Thus, before interpreting the results of Experiment 1, we need to assess whether toddlers know about right-dislocated prosody, rather than failing to understand right-dislocated structures altogether. While corpora suggests that toddlers can incorporate dislocations in their own productions, perhaps those samples were not representative of toddlers in general; or perhaps those sentences were imitations of adults’ sentences and did not capture toddlers’ ability to *understand* dislocated sentences. We addressed this potential confound in Experiment 2, where toddlers were tested on their comprehension of right-dislocated sentences featuring familiar verbs.

Experiment 2: Using prosodic cues with familiar verbs

In this experiment, we sought to assess the possibility that toddlers are simply unable to comprehend right-dislocated sentences. For this, we selected common verbs which are transitive but tolerate object omission. This allowed us to form pairs of transitive and right-dislocated sentences featuring the same words in the same order, but in which the post-

verbal DP was either the patient (transitive sentence) or the agent (right-dislocated sentence). To illustrate, here is an example with the verb ‘manger’ (to eat):

1. *Transitive*: Il_i mange le $canard_j$

   (*It_i is eating the $duck_j$* = something is eating a duck)

2. *Dislocated*: Il_i mange, le $canard_i$

   (*It_i is eating, the $duck_i$* = the duck is eating)

Hence, a simple way to investigate toddlers’ understanding of right-dislocated sentences is to test their comprehension of the agent-patient relation in these sentences. To do so, we presented toddlers with familiar verbs introduced in either transitive or right-dislocated sentences while they saw two videos on the screen. For example, a sample toddler may have heard “il_i mange, le $canard_i$” *it_i eats, the $duck_i$* (where the duck is the agent) while seeing a pair of videos (Figure 4): one in which the referent of the post-verbal DP (“le canard” *the duck*) is the patient of the action (such as a tiger eating a duck) and one in which the post-verbal DP is the agent (a duck eating bread). If toddlers cannot process the prosodic cues signaling dislocation, they should interpret “le canard” (*the duck*) as the patient of the action, as should toddlers in the transitive condition (e.g., hearing “il_i mange le $canard_j$” *it_i eats the $duck_j$*), and hence look more at the corresponding event where the duck is being eaten. On the contrary, if toddlers comprehend right-dislocated sentences and can use the prosody of the sentence to conclude that “le canard” (*the duck*) shares the agentive role of the preceding pronoun, then they should look more toward the event where the duck is eating, thus systematically differing from toddlers in the transitive condition.

Method

Participants—Thirty-two 28-month-old toddlers participated in the study (mean = 28.5, min = 27.17, max = 28.27; females = 19). Sixteen toddlers were assigned to each of the two experimental conditions (transitive, dislocated). Nine additional toddlers were excluded from the final analysis because of fussiness ($n = 5$), experimental error ($n = 2$) or bilingualism ($n = 2$).

Apparatus, procedure, stimuli and analyses—The same apparatus as in Experiment 1 was used. This experiment started with 4 videos that introduced each of the puppet characters: a rabbit, a duck, a tiger and a monkey. Each puppet was presented in a 5s video waving its arm and was labeled once. After this preview, toddlers saw one practice trial with the same structure as the test trials (see below), involving the intransitive verb ’sauter’ *to jump*. As in Experiment 1, this trial familiarized toddlers with the procedure and showed them that the soundtrack matched one of the videos.

Then, 4 test trials involving familiar verbs were presented following the same procedure as the practice trial and the test events of Experiment 1 (see Figure 2). For each verb, a pair of videos was presented, both depicting the action named by the verb (see Figure 3). Toddlers in each condition heard 3 tokens of the test sentences during each test trial. For ’manger’ (to eat), the DP-patient video showed a tiger eating a duck and the DP-agent video showed a duck eating bread. These videos offered toddlers the option of interpreting the post-verbal

DP 'le canard' (the duck) either as the agent (appropriate for the dislocated condition, "Il_i va manger, le $canard_i$" *It_i will eat, the $duck_i$*) or as the patient of the action (appropriate for the transitive condition, "Il_i va manger le $canard_j$" *It_i will eat the $duck_j$*). Four pairs of videos were created for the following verbs: 'pousser', 'porter', 'manger' and 'taper' (push, carry, eat and hit). These verbs are likely to be known to two-year-olds, according to a vocabulary survey carried out in a previous experiment. As before, judgments from 10 adults confirmed that transitive and right-dislocated sentences were interpreted correctly more than 90% of the time. The same speaker recorded all soundtracks in child-directed speech. The order of the video pairs and the left-right position of the videos were counterbalanced across subjects, within each condition. Similarly to Experiment 1, acoustic analyses confirmed that there was a clear prosodic boundary between the verb and final DP in the dislocated but not the transitive sentences (preboundary syllable lengthening: M_{trans}= 285ms vs. M_{disloc}= 673ms, $t(22)= -6.83$, $p < .001$; pitch drop: M_{trans}=−35Hz vs. M_{disloc}=−113Hz, $t(22) = 10.65$, $p < .001$).

The data were coded as in Experiment 1. Data from test events in which the child looked away for more than half of the event's duration were treated as missing (5 test events overall). Looking times were averaged across the four test trials ("pousser" *push*, "taper" *hit*, "manger" *eat*, "porter" *carry*) for each condition. Preliminary analysis of the proportion of time spent looking away revealed no significant effect of condition ($F < 1$) indicating that toddlers tended to look away equally in the two experimental conditions. Since time spent looking away was consistent across conditions, we based our analysis on the proportion of looking time to the DP-patient video over total looking time. For the sake of completeness, Figure 5 shows the proportion of looking time to the DP-patient and the DP-agent videos, and time spent looking away.

Results

As Figure 5 shows, toddlers interpreted the dislocated sentences very differently from the transitive sentences. An ANOVA with a between-subjects factor Condition (Transitive, Dislocated) revealed a significant effect of Condition on the proportion of looking time to the DP-patient video ($F(1,30) = 8.08$, $p < .01$). Toddlers in the transitive condition looked more toward the DP-patient video than toddlers in the dislocated condition. In addition, toddlers in the transitive condition looked more toward the DP-patient video than the DP-agent video ($t(15) = 4.57$, $p < .001$) but this was not the case for toddlers in the dislocated condition ($t < 1$). Inspection of the results for each individual verb revealed that one of the verbs, "taper" (*to hit*), showed a heavy bias toward the DP-patient video. This video showed one puppet hitting another, and was deliberately made funny to avoid frightening toddlers (the animate hittee was cringing in between hits, while the patient on the other video was inanimate and not moving); unfortunately this item generated a strong baseline bias towards one of the events. Note that removing this item resulted in a marginal preference for the DP-agent video over the DP-patient video in the dislocated condition ($t(15) = -2.5$, $p = .05$) but did not change the main effect of Condition on the DP-patient video ($F(1,30) = 7.89$, $p < .01$).

Discussion

Experiment 2 confirmed the hypothesis, based on corpus analyses of toddlers' speech, that French-learning 28-month-olds understand right-dislocated sentences. Results clearly show that toddlers gave different interpretations to transitive and right-dislocated sentences, even though both sentences had identical surface word order. Thus, toddlers used the prosodic information in right-dislocated sentences to correctly identify the post-verbal DP as naming the agent, rather than the patient, of the event described by a familiar verb. More specifically, toddlers were clearly able use the salient prosodic cues of right-dislocations to block the tendency to map sentences with two DPs onto transitive interpretations. Furthermore, they were even able to employ different surface-to-meaning mapping procedures, those that canonically link each DP to a distinct participant role in the event, as in transitive sentences (First DP = Agent, Second DP= patient), and those that deviate from that assignment when prompted by specific prosodic cues (the first and the second DP share the agent role as in right dislocated sentences). These results rule out the first explanation put forward above, concerning toddlers' ignorance of prosodic cues to right dislocation, and suggest that toddlers fail to use their knowledge of these prosodic cues when interpreting a novel verb. The key question that remains, then, is: Why do toddlers succeed in using a different surface-to-meaning mapping procedure with familiar but not novel verbs?

In the case of Experiment 2, toddlers heard these dislocations with familiar verbs, that is, they already knew that these verbs could be used in both intransitive and transitive sentences (*eat*, *push*, etc.). Toddlers' knowledge of which constructions a given verb occurs in would make it easier for them to use all the information at their disposal to identify the intended structure in Experiment 2. Contrary to Experiment 1, they could rely on two sources of evidence pointing towards the intransitive interpretation: the prosody of right dislocations and some prior knowledge that the verb can appear in an intransitive construction. If prior syntactic knowledge retrieved from the lexicon about known verbs is a major reason why the familiar verbs are better processed in Experiment 2, then one would expect that with pre-exposure about the possible syntactic frames accepting the novel verb, toddlers might be able to go beyond a canonical mapping of each DP into a distinct verb argument while interpreting right-dislocated sentences featuring this verb.

A similar case can be made building on the second interpretation put forward for the results in Experiment 1, whereby toddlers ignored the prosodic information in the possible absence of a pragmatic context licensing a right dislocation. It might be that there are ordinarily both pragmatic and syntactic conditions on right dislocation – toddlers have to know that the verb is intransitive to permit the structure, and the structure has to be appropriate in the discourse context. From this viewpoint, the task set forth for toddlers in Experiment 1 is more ambiguous, as pragmatic criteria are not met and the verb is unknown so toddlers cannot tell if the syntactic criteria are met. This makes it very hard for toddlers to identify the intended right-dislocation structure. In sentences with known verbs as in Experiment 2, toddlers had more evidence about the sentence structure (e.g., the verb's syntactic knowledge) and hence better able to use the prosodic cue to dislocation to yield an intransitive interpretation. On this view, one might predict that satisfying the pragmatic requirements of the right-dislocated structure in Experiment 1 would increase toddler's ability to arrive at an

intransitive interpretation of right-dislocated sentences even with a novel verb. These interesting possibilities could be assessed in future work.

General discussion

The present paper set out to explore the mechanisms that allow young listeners to cope with the productive power of language. Previous work has shown that toddlers readily encode each DP in the sentence as having a distinct participant-role such that a novel verb embedded in a DP-V-DP transitive sentence would refer to a causal action where an agent acts on a patient. However, languages allow multiple mappings between form and meaning, some of which do not fit this canonical format (mapping of two DPs to different roles). We studied toddlers' reliance on a strict canonical surface-to-meaning mapping procedure in the presence of clear, language-specific cues that would normally block such an interpretation. French right-dislocation was the ideal phenomenon to study this question because it is common in the input and even produced by toddlers at about this age. Nonetheless, prosodic cues to dislocation did not suffice to prevent toddlers from interpreting each DP as a separate argument when learning a novel verb in Experiment 1. Prosody was not simply discarded during processing, as toddlers integrated it on the fly when interpreting sentences with familiar verbs in Experiment 2.

Experiment 2 provided novel evidence on toddlers' integration of multiple cues while interpreting sentences. In several prior studies, morpho-syntactic cues have been found to outweigh initial biases towards a unique DP-argument mapping. For instance, two-year-olds did not interpret the sentence "the duck and the bunny are gorping" as transitive despite the presence of two DPs (Arunachalam & Waxman, 2010; Naigles, 1990). Thus, toddlers can enrich their sentence representations using cues such as the word order (DP-DP-V), the morphological marking of the plural agreeing verb 'are', and the presence of the conjunction 'and', to interpret a non-canonical sentence. To our knowledge, the present study is the first to show that prosodic cues can also constrain interpretation in 28-month-olds. Indeed, previous studies have found that 4 to 6 year-old children are less adept than adults at using prosodic cues while interpreting sentences (Choi and Mazuka, 2003; Ito, Jincho, Minai, Mazuka, 2012, Sekerina & Trueswell, 2012; Snedeker & Yuan, 2008; but see Zhou, Crain & Zhan, 2012, for evidence that prosodic cues can be used robustly when children do not have to commit to a syntactic structure). While in these studies children are tested on prosodic cues that are not consistently represented in spontaneous speech (such as phrasal attachment ambiguity or contrast resolution, where prosodic disambiguation depends on the speaker's awareness of the ambiguity), our study uses a syntactic structure, right-dislocation, that is obligatorily accompanied by specific prosodic cues. Our results show that even very young children are able to use sophisticated surface-to-meaning mappings based on language-specific prosodic cues when interpreting sentences, provided that the prosodic cues are a systematic predictor of the syntactic structure involved, as is the case for right-dislocated sentences.

However, why did toddlers fail to use prosodic cues when using sentence structure to interpret novel verbs (Experiment 1), but succeed in using the same prosodic cues to interpret sentences with familiar verbs (Experiment 2)? We propose that the most

convincing explanation relates to optimal behavior in the presence of uncertainty. Adults comprehend sentences effortlessly despite errors that may occur in the production or perception stage, which may be viewed as *noise*. According to some researchers (Gibson, Bergen & Piantadosi, *in press*; Levy, 2008), people integrate in real time both their linguistic expectations about the sentence that was most likely uttered and their knowledge about the possible alterations (addition or deletion of linguistic items) the speaker is most likely to have made. Moreover, people warp their linguistic expectations depending on the level of uncertainty. In particular, adults may choose the most probable interpretation even when this entails altering the word sequence of words heard, provided that the alterations are minimal. For example, the syntactically correct but implausible sentence “the child gave the drawing her mother” is more likely to be interpreted as the more plausible sentence “the child gave the drawing *to* her mother” in contexts where there is a non-negligible probability that the function word ‘to’ is omitted (Gibson et al., *in press*). Other work suggests that adults and children may repair sentences in different ways. Naigles, Gleitman and Gleitman (1992) presented both adults and children with known verbs in ungrammatical sentence frames (e.g., “Noah comes the elephant to the ark”). Adults were more likely to alter the sentence to fit the verb-specific information (interpreting it as “Noah comes to the ark with the elephant”). In contrast, 2- to 3-year-old toddlers were more likely to alter the verb meaning to fit the sentence frame (interpreting it as “Noah brings the elephant to the ark”). Thus, toddlers and adults differ in what they consider the most plausible interpretation in conditions of uncertainty, presumably reflecting their different levels of knowledge about the behavior of particular verbs.

The question of how listeners repair unreliable input may be relevant to the current work. Recall that transitive sentences are more frequent than right-dislocated sentences in the input (since only 5% of sentences are right-dislocated). Therefore, in the presence of uncertainty (i.e., a novel verb extracted from a semantically, and possibly pragmatically, impoverished setting), toddlers might easily default to a more likely sentence structure candidate provided very few alterations need to be applied. This was the case in our study, where going from right-dislocated sentences to the canonical DP_{agent}-V-$DP_{patient}$ interpretation with two arguments required the deletion of a single element (the prosodic boundary, which is not even a lexical item). Compare this with previous studies where 25-month-olds succeeded in giving different interpretations to intransitive sentences with two DPs conjoined in subject position (”the duck and the bunny are gorping”) and transitive sentences (”the duck is gorping the bunny”) (Arunachalam & Waxman, 2010; Naigles, 1990). Contrary to our case where only one alteration is necessary to go from right-dislocated sentences to transitive sentences, in such conjoined intransitive sentences (”the duck and the bunny are gorping”), at least three alterations would be needed: one deletion (‘and’), one replacement (of ‘are’ by ‘is’) and one movement (‘the bunny’), making the transitive interpretation much harder to come up with. Interestingly, in Gertner & Fisher (2012), slightly younger toddlers of 21 months assigned similar interpretations to transitive sentences (“the girl is gorping the boy”) and conjoined-subject intransitive sentences (“the girl and the boy are gorping”). These toddlers linked both sentences to a causal event where a girl is acting on a boy, showing a bias to analyze the first of two nouns as the agent and the second as the patient (but see Arunachalam, Hansen, Escover, & Waxman, 2012 for success with these sentences at 21

months in a pragmatically supportive context). Such errors are a reflection of toddlers' current knowledge about the structure of their language in interaction with the current input sentence and the candidate scenes available. The alterations that toddlers are willing to overlook depend on their prior syntactic representation. If, at 21 months, toddlers' initial representation of the sentence is grounded in the set of DPs, where each DP gets a distinct participant role, then no accommodation might be required to allow any two-DP sentences to be interpreted as transitive. The highly similar sentences in Gertner and Fisher (2012) and Naigles (1990) led to drastically different results in 21- and 25-month-olds respectively. Due to incomplete knowledge of the native language, 21-month-olds may differ in the evidence they represent, or the 'cues' they use to identify sentence structure; as a result, their a priori syntactic expectation for a transitive sentence given the input sentence will differ from that of 25-month-olds or adults in this situation. This tells us that 25-month-olds' priors have changed and that toddlers are no longer able to accommodate such dramatic changes in sentence structure to entertain a transitive interpretation. Boosting reliance on typical structures is useful most of the time because it facilitates canonical (frequent) interpretations and thus prompts the adoption of the most rational interpretation given the listeners' current knowledge. However, when the distance is small between a candidate interpretation and an interpretation with a very high prior probability, then it can sometimes lead to incorrect representations, as we demonstrated here.

Even adults still rely on language-specific canonical sequences reflecting their structural priors to interpret sentences. For example, Ferreira (2003) documented that English-speaking adults experienced more difficulty in parsing sentences that did not follow the canonical transitive order (e.g., passive sentences) than sentences that did. Thus, under this account, it is reasonable for toddlers to interpret right-dislocated sentences as transitive in situations of uncertainty, given their accumulated linguistic knowledge.

Conclusion

In summary, we document that toddlers readily use prosodic cues to dislocation in order to interpret sentences featuring known verbs; however, they revert to a canonical surface-to-meaning mapping when learning a new verb. Hence, two-year-olds adopt a reasonable interpretation, given their linguistic knowledge, when in situations of uncertainty. These results suggest that, while toddlers readily posit interpretations for phrases and expressions they have never encountered before, these interpretations may not always be right, and sometimes reveal that young children can use cheaper tricks to perform outstanding feats.

Acknowledgments

This work was supported by grants from the French Agence Nationale de la Recherche (ANR-2010-BLAN-1901) and from French Fondation de France to Anne Christophe, from the National Institute of Child Health and Human Development (HD054448) to Cynthia Fisher, Fondation Fyssen and Ecole de Neurosciences de Paris to Alex Cristia, and a PhD fellowship from the Direction Générale de l'Armement (DGA, France) supported by the PhD program FdV (Frontières du Vivant) to Isabelle Dautriche.

We would like to thank Isabelle Brunet for the recruitment, Michel Dutat for the technical support and Hernan Anllo for his puppet mastery skill. We are grateful to the families that participated in this study. We also thank two anonymous reviewers for their comments on an earlier draft of this manuscript.

References

Ashby W. The syntax, pragmatics, and sociolinguistics of left- and right- dislocations in french. Lingua. 1988; 75:203–229.

Ashby W. An acoustic profile of Right-Dislocations in French. Journal of French Language Studies. 1994; 4:127–145.

Arunachalam S, Escovar E, Hansen MA, Waxman SR. Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. Language and Cognitive Processes. 2012; 28:417–425. [PubMed: 24163490]

Arunachalam S, Waxman SR. Meaning from syntax: evidence from 2-year-olds. Cognition. 2010; 114:442–446. [PubMed: 19945696]

Bencini GML, Valian VV. Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. Journal of Memory and Language. 2008; 59:97–113.

Choi Y, Mazuka R. Young children's use of prosody in sentence parsing. Journal of psycholinguistic research. 2003; 32:197–217. [PubMed: 12690831]

Christophe A, Millotte S, Bernal S, Lidz J. Bootstrapping lexical and syntactic acquisition. Language & Speech. 2008; 51:61–75. [PubMed: 18561544]

Christophe A, Gout A, Peperkamp S, Morgan J. Discovering words in the continuous speech stream: The role of prosody. Journal of Phonetics. 2003; 31:585–598.

Dautriche, I. Master thesis. 2012. The role of prosody in toddlers' interpretation of verbs' argument structure.

De Cat, C. Interpretation, Syntax, Acquisition. Oxford University Press; 2007. French Dislocation.

Demuth K, Moloi F, Machobane M. 3-Year-olds' comprehension, production, and generalization of Sesotho passives. Cognition. 2010; 115:238–251. [PubMed: 20153462]

Demuth K, et Tremblay A. Prosodically-conditioned variability in children's production of french determiners. Journal of Child Language. 2008; 35:99–127. [PubMed: 18300431]

Dryer MS. Order of subject, object, and verb. The world atlas of language structures. 2005:330–333.

Ferreira F. The misinterpretation of non-canonical sentences. Cognitive Psychology. 2003; 47:164–203. [PubMed: 12948517]

Fisher C. Structural Limits on Verb Mapping: The Role of Analogy in Children' s Interpretations of Sentences. 1996; 81:41–81.

Gerken L, Jusczyk PW, Mandel DR. When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. Cognition. 1994; 51:237–265. [PubMed: 8194302]

Gertner Y, Fisher C. Predicted errors in children's early sentence comprehension. Cognition. 2012; 124:85–94. [PubMed: 22525312]

Gibson E, Bergen E, Piantadosi ST. The rational integration of noise and prior semantic expectation: Evidence for a noisy-channel model of sentence interpretation. Proceedings of the National Academy of Sciences. in press.

Göksun T, Küntay AC, Naigles LR. Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. Journal of child language. 2008; 35:291–323. [PubMed: 18416861]

Gordon P, Chafetz J. Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. Cognition. 1990; 36:227–254. [PubMed: 2265528]

Hollich G. Supercoder: A program for coding preferential looking (version 1.5). 2005

Ito K, Jincho N, Minai U, Yamane N, Mazuka R. Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. Journal of Memory and Language. 2012; 66:265–284.

Kern S. Lexicon development in French-speaking infants. First Language. 2007; 27:227–250.

Lee JN, Naigles LR. Mandarin learners use syntactic bootstrapping in verb acquisition. Cognition. 2008; 106:1028–1037. [PubMed: 17537422]

Levin, B.; Rappaport Hovav, M. Argument realization. NewYork: Cambridge University Press; 2005.

Levy, R. A noisy-channel model of rational human sentence comprehension under uncertain input; Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing; 2008. p. 234-243.

Lidz J, Gleitman H, Gleitman L. Understanding how input matters: verb learning and the footprint of universal grammar. Cognition. 2003; 87:151–178. [PubMed: 12684198]

Maratsos M, Fox D, Becker J, Chalkley M. Semantic restrictions on children's passives. Cognition. 1985; 19:167–191. [PubMed: 4017515]

Messenger, K.; Fisher, C. Mistakes weren't made: 3-year-olds understand passives with novel verbs; Poster presented at SRCD Biennial Meeting; Montreal. 2011.

Naigles L. Children use syntax to learn verb meanings. Journal of child language. 1990; 17:357–374. [PubMed: 2380274]

Naigles, L.; Gleitman, H.; Gleitman, L. Syntactic Bootstrapping and Verb Acquisition. In: Dromi, E., editor. Language and Cognition: A Developmental Perspective. NJ: Ablex; 1993.

Noble CH, Rowland CF, Pine JM. Comprehension of argument structure and semantic roles: Evidence from infants and the forced-choice pointing paradigm. Cognitive Science. 2011; 35:963–982. [PubMed: 21545486]

Pinker S, Lebeaux DS, Frost LA. Productivity and constraints in the acquisition of the passive. Cognition. 1987; 26:195–267. [PubMed: 3677572]

Reinhart T. Pragmatics and linguistics: An analysis of sentence topics. Philosophica. 1981; 27:53–94.

Rossi, M. L'intonation. Le système du français: description et modélisation. Paris: Ophrys; 1999.

Sekerina IA, Trueswell JC. Interactive processing of contrastive expressions by Russian children. First Language. 2012; 32:63–87. [PubMed: 24465066]

Snedeker J, Yuan S. Effects of prosodic and lexical constraints on parsing in young children (and adults). Journal of Memory and Language. 2008; 58:574–608. [PubMed: 19190721]

Yuan S, Fisher C. "'Really ? She Blicked the Baby ?'" Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. Psychological Science. 2009; 20:619–626. [PubMed: 19476591]

Yuan S, Fisher C, Snedeker J. Counting the nouns: Simple structural cues to verb meaning. Child Development. 2012; 83:1382–1399. [PubMed: 22616898]

Zhou P, Crain S, Zhan L. Sometimes children are as good as adults: The pragmatic use of prosody in children's on-line sentence processing. Journal of Memory and Language. 2012; 67:149–164.

Transitive dialogue and (Dislocation dialogue)

A: Hey, il va daser (,) le bébé!	Hey, he will dase (,) the baby
B: Ah bon?il va daser (,) le bébé?	Oh really? he will dase (,) the baby?
A: Oui, et ils ont dasé (,) les canards	Yes, and they dased (,) the ducks
B: C'est vrai, ils ont dasé (,) les canards	That's right, they dased (,) the ducks

Intransitive dialogue

A: Hey il dase !	Hey he is dasing !
B: Ah bon? il dase ?	Oh really? He is dasing?
A: Oui et ils ont dasé	Yes and they dased !
B: C'est vrai, ils ont dasé	That's right they dased !

Figure 1.

Sample of dialogue pre-exposure in Experiment 1 for the 3 conditions: transitive, intransitive and dislocated. The dialogues were split in two 28s-clips containing 4 sentences each and separated by a 3 second black screen. Transitive and dislocated dialogues were similar in every respect but the prosody.

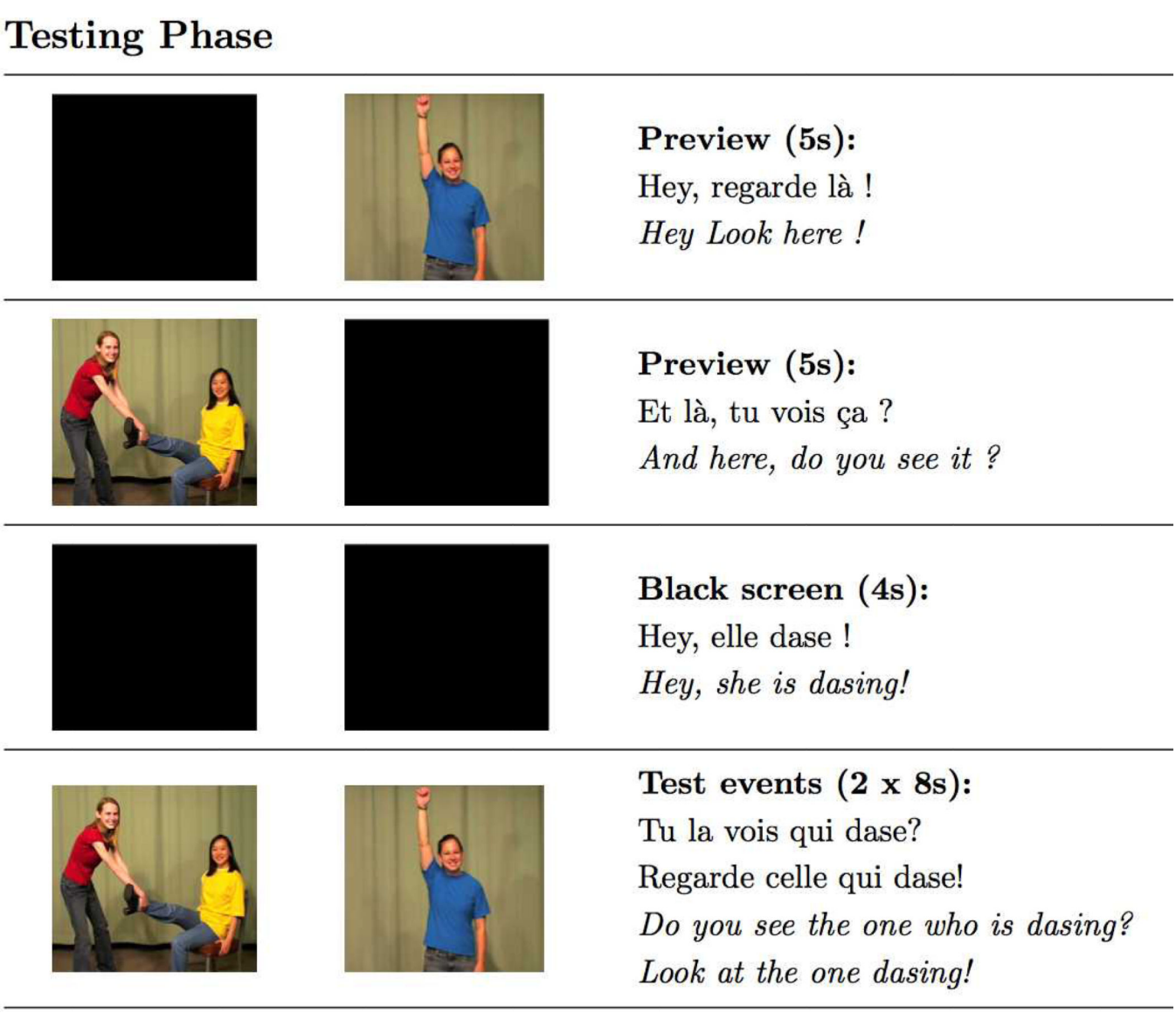

Figure 2.
Detailed procedure of the testing phase of Experiment 1. Action videos are first presented for 5 seconds each before two 8-second test events showing both videos simultaneously.

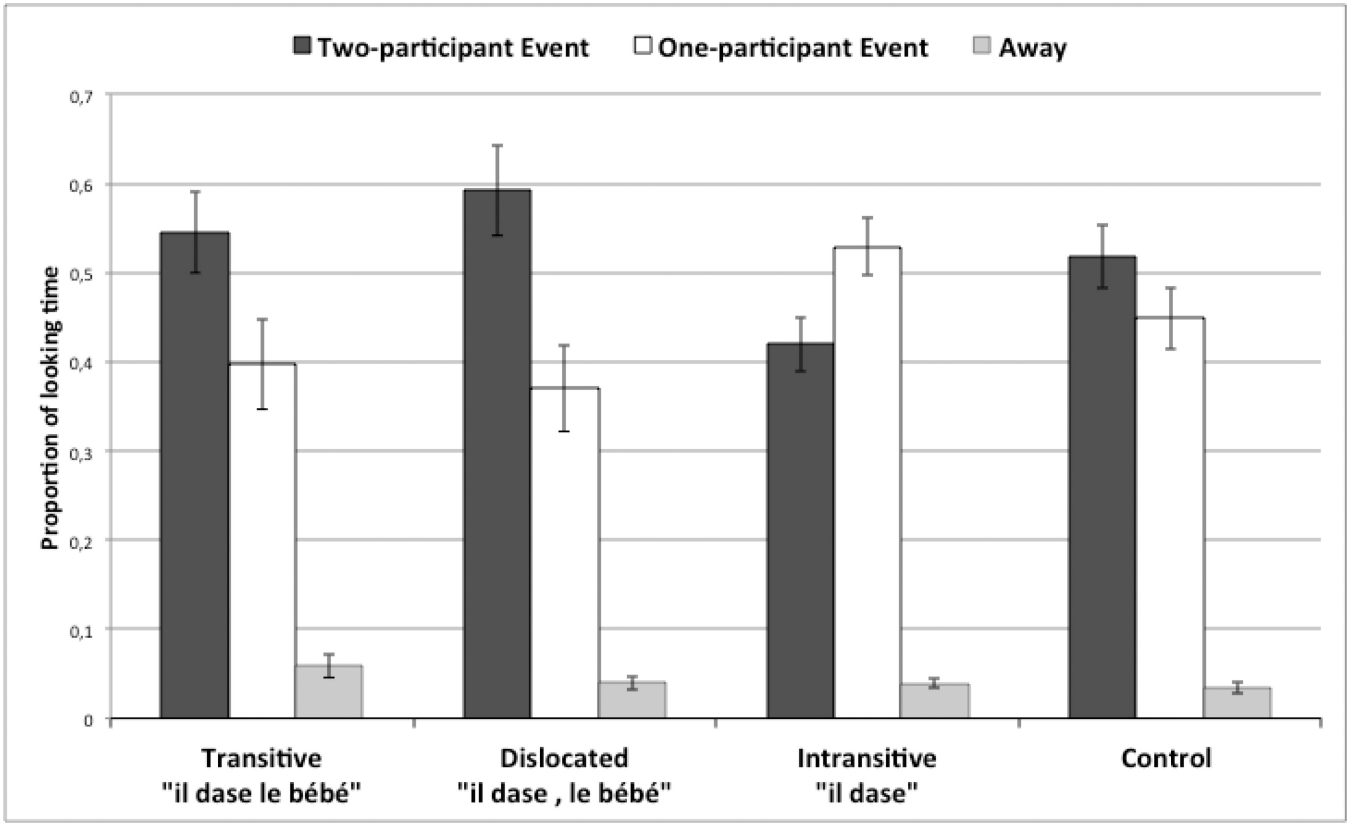

Figure 3.

Proportion of looking and look-away times, averaged across the two test events phases, in the test trial in Experiment 1. Error bars indicate standard error of the mean per condition. Toddlers in the dislocated condition behaved as did toddlers in the transitive condition, and markedly differently from toddlers in the intransitive condition: they looked more at the two-participant video than at the one-participant video.

Figure 4.

For each verb the same pair of video was played for toddlers in the dislocated or in the transitive conditions. The DP-patient video corresponded to the transitive interpretation of the sentence, in which the referent of the post-verbal DP in the test sentence was the patient in the action. The DP-agent video corresponded to the dislocated interpretation, in which the referent of the post-verbal DP was the agent of the action.

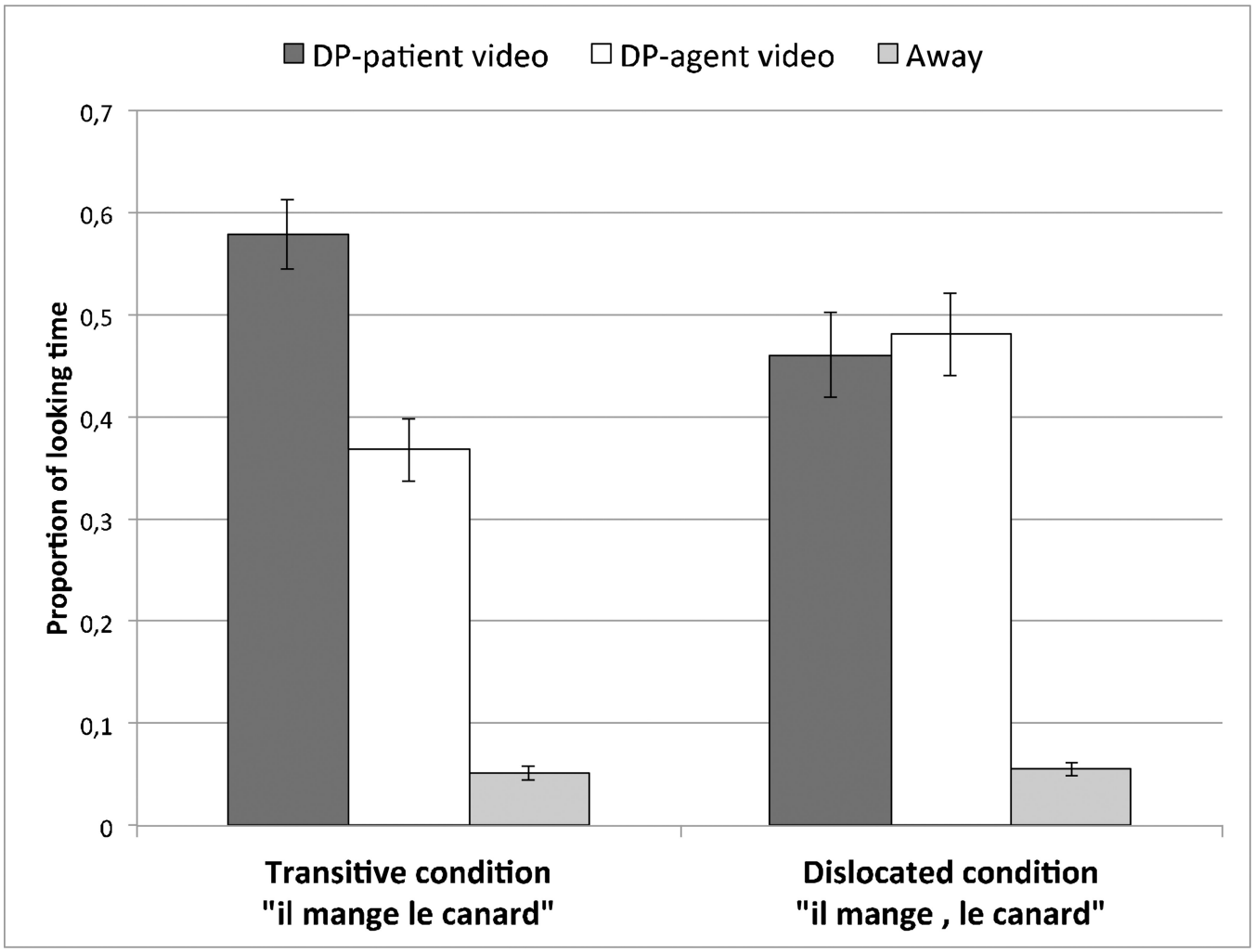

Figure 5.
Proportion of looking time toward the DP-Agent video and the DP-Patient in Experiment 2. Error bars indicate standard errors of the mean per condition.