Language Sciences 46 (2014) 100–114

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

What kind of concepts need language?

Jill de Villiers

Psychology Department, Smith College, Northampton, MA 01063, USA

ARTICLE INFO

Article history:
Available online 10 July 2014

Keywords:
Events
Language and thought
Shadowing
Negation
Dual task

ABSTRACT

The question addressed in this paper is whether the language faculty provides new possibilities for concepts. Drawing on work from Hinzen and Sheehan (2013) and others, the claim is that I-language allows humans to see resemblances among events that are similar in only an abstract way mediated by their linguistic description, such as the hierarchical form <boy chase dog> that unites the class of events in which a boy chases a dog, not vice-versa. Experimental data are presented from five preliminary studies. These include a study of 61 preschool-aged children in an act-out procedure with and without verbal description, and a second with 63 adults in a picture choice procedure under conditions of verbal versus rhythmic shadowing. The methodological difficulties are discussed, particularly the use of rhythmic shadowing as an adequate control. Three further studies are reported that use adult participants in an implicit concept formation paradigm in an eyetracker, using a dual-task procedure with only verbal shadowing. These studies have the goal of discovering whether verbal shadowing selectively impairs certain abstract concepts and not others. Together the results suggest that verbal shadowing is selectively disruptive of the recognition of similarity across perceptually diverse events sharing a propositional description, but not of the formation of other abstract concepts.

1. Introduction

In this paper I explore the ramifications of the claim that I-language is a medium for thought. Specifically, I review some alternative proposals for a propositional theory of thinking, and then turn to the arguments by Hinzen (2009, in press) and Hinzen and Sheehan (2013). The outcome of that reasoning is a highly specific and testable claim about the new conceptual capacity that human language brings about. In the final half of the paper I describe some initial attempts to test it empirically with children and adults, and make a call for more experimentation with infants and animals.

Fodor (1975) proposed that thought in general required a propositional format, because any other format, such as imagery or unstructured symbols, would be insufficient to support negation, fine grained detail, conjunction and disjunction, conditional reasoning and so forth. Only propositions can have a truth value, that is, can be judged true or false. Fodor considered that to be key for reasoning. These ideas are elaborated in Fodor (2008) in which he reiterates his stance. Work in linguistics by Chomsky emphasized linguistic properties such as systematicity and productivity. Systematicity is evident when someone who can understand “the man kissed the baby” can also understand “the baby kissed the man”. Productivity is evident by the fact that a person with a finite lexicon and a grammar can make novel sentences. For Fodor, it was evident that natural language is systematic and productive because it is an expression of thoughts: thoughts also have systematicity and productivity.

E-mail address: jdevilli@smith.edu.

http://dx.doi.org/10.1016/j.langsci.2014.06.009
0388-0001/

Fodor also argued that it was impossible to imagine language acquisition itself in the absence of a conceptual system of the same degree of richness, underlying and preceding the acquisition of words and sentences. That is, the concepts had to be present to allow the mapping onto them of words and sentences from the conventional natural language that the child learned. Concepts could not just arise *with* the words and sentences, they had to be the driving force for acquiring the terms and expressions. In fact, concept are unlearnable on Fodor's view, they are merely triggered by "stereotypes" (Fodor, 2008), but the classes, the boundaries of the concept, are innate.

Fodor's solution was the Language-of-Thought (LoT) hypothesis, in which all thought is entertained in a propositional form akin to natural language, but shared by dogs, infants, and other nonverbal creatures, that is, any species that engages in reasoning. The Language of Thought was of necessity *innate*, because it was unlearnable in itself but formed the foundation for developing the particular human languages that infants learn. It was held also to be the translation mechanism for bilingual individuals who could convert propositions in LoT into either language.

The LoT hypothesis has been rejected by many contemporary philosophers of mind, for a variety of reasons. First, there is great resistance to the idea that all of our concepts are innate. Second, it has been argued that some phenomena of the mind that have been given a propositional description might be subject to a more reductive analysis, potentially solvable by non-symbolic connectionist architecture (Smolensky, 1990). Thirdly, Hutto (2008) argues that defenders of LoT provide no account of how it came into existence: neither how its symbols got their references nor the origins of its syntax.

Several philosophers (Davies, 1998; Frankish, 2004; Segal, 1998; Collins, 2001) have taken the position that the mechanisms of propositional reasoning can be achieved using the structures provided by natural language, in which case it seems unnecessary to duplicate the machinery for LoT. The properties that make the positing of a LoT desirable for human thinking are precisely those that could implicate natural language in thinking, namely a richly structured representation of meaning and form. Several philosophers of mind have taken Fodor's basic premise of a sentential basis for thought as correct, but have argued that a LoT would be redundant, given natural language. Here is Hutto's (2008) take on the matter:

> "Here my interest is only in the idea that in order to have a content-involving propositional attitude one must have facility with complex linguistic forms – sentences of some kind or another. Only sentences have the right syntactic and semantic properties for expressing propositional contents…. More than this, because they have complex structures, sentences are purpose-built for doing logical work: they can be parsed at their meaningful joints and this kind of internal organization makes them ideally suited to fund computationally tractable inferential operations" (p. 87–88).

Notice, however, that there is a price to pay for this move, in that we leave infants and dogs without a representational medium for their complex thoughts. Any apparently complex behavior or reasoning in non-linguistic beings must be argued away as reducible to something that does not entail propositional representation (Hutto, 2008).

Carruthers (2002) argues that *some* modes of human thinking require natural language representation, but he contends that the critical representations are not necessarily fully expressed sentences, but rather, Logical Form representations. By Logical Form Carruthers means the level of language before it is converted into Phonological Form, so the kind of Logical Form that linguists, rather than logicians, entertain. It can be argued (Collins, 2000) that LF is a level of *syntax, not semantics*. LF encodes those features of syntax that are relevant to semantic interpretation, for example, scope assignment to quantifier and *wh* phrases. In short, LF is the input to semantics. Hence, an "LF representation" is a well-formed labeled phrase marker containing all information relevant to the semantic interpretation of the string.

In contrast to Fodor, Carruthers does not believe that all thought or reasoning entails a propositional format, but taking the lead from Spelke (2003), he argues that only thinking that needs cross-modular connections in the mind requires language. Spelke and Carruthers contend that language serves as a kind of bridge device between modules that need to combine to solve higher order reasoning tasks. Information from different modules is integrated for non-domain-specific thinking. Spelke's illustration is the integration of information from two spatial modules: one concerned with landmarks, and one concerned with directional information. By combining them, a child can represent a space as something like:

> "To the left of the blue wall"

and hence solve a problem of navigation. Spelke's evidence comes from developmental data, and also from the curious case of human adults who are engaged in shadowing language, a task that effectively ties up their language capacity from being engaged in the navigational task (Hermer-Vasquez et al., 1999). Shadowing rhythmic tapping was the matched control, which did not interfere with location judgment. The intriguing possibility is that language is the modality-neutral means of internal communicating across diverse modules, or across different sensory inputs.[1]

Carruthers argues that the representation for reasoning of this kind is Logical Form, but if LF is spelled out into phonological form, we experience it as conscious thinking. The kind of thought in which Carruthers argues language is involved is that which is semantically evaluable, that is, having an associated judgment of truth or falsity. In addition, thought of this kind has structured components, e.g. embeddings or other structural sub-units. Most importantly, the thought is related to other thoughts by its *content*.

It is important to contrast his view with that of Fodor (1975) and other theorists. Carruthers categorically denies that thought is impossible without language, because nonverbal organisms especially in the social realm do seem to him to engage

[1] The findings are not without controversy, see Learmonth et al. (2002).

in complex reasoning. This is a point of contention with Povinelli and Vonk (2004) and Hutto (2008). Carruthers denies that language is a mere *conduit* for transmission of thought, and he proposes that language has a role to play beyond being the carrier of thought in the medium of cultural ideas (e.g. Dennett's 1991 cultural memes). Carruthers also rejects Fodor's idea that natural language plays the role of a mere mental tool, namely its role in executive function skills, for assisting memory, planning and the like.

Why did Fodor reject natural language as the medium? Collins (2000) attributes Fodor's skepticism about the role of natural language in thought to his focus on E-language as opposed to I-language, that is, to his consideration of externalized language forms rather than internal language forms. The more contemporary view of I-language versus E-language has replaced the old competence–performance distinction. However, it introduces certain new difficulties.

Consider the recent view expressed by Chomsky (2010) about the evolution of language. Chomsky contends that language evolved first and foremost as a tool of thought, and secondarily as a communication system. He believes that a mutation occurred that made I-language (internal language) available, and that the extra reasoning power it enabled made the mutation favorable. When it had spread sufficiently throughout a group, the stage was set for it to develop into E-language, with a phonology and shared conventions. This position is strongly criticized by Pinker and Jackendoff (2005), who argue that essential properties of language can be traced to communicative pressures. Hutto (2008) makes a parallel set of arguments about the public, normative pressures on the meanings of symbols and their arrangement in syntax, though he is arguing about the insufficiency of LoT in that regard.

What is this I-language? Hinzen (2006) describes it as the capacity for merging categories and thus creating meanings that do not exist without hierarchical syntactic structure. It is the basic structuring mechanism of human language, supposedly unique to the human mind. In this way it does not resemble LoT, since LoT is species-neutral. But in other ways there are parallels: both I-language and LoT are argued to be unlearnable, perfect, pre-structured, and necessary for E-language. Yet LoT was construed independently of language, as the representational basis of thought. For Chomsky, I-language seems to be serving the role in thought that Fodor's LoT did in providing structured representations that are systematic, recursive, and compositional.

Yet there are questions. To return to the question of LF as the medium of complex thought, e.g. Carruthers argues that LF can do the jobs for which arguments for LoT were most persuasive, e.g. handling the propositional attitudes, negation, conjunction, scope and so forth. But Collins (2000) and Segal (1998) argue that for the best representation of propositional attitudes (beliefs, knowledge). what is really needed are the Total Forms of natural language, beyond LF and complete with all kinds of phonological and orthographic detail. However, if we accept this idea about Total Form, we have begun talking about the kinds of detail that cannot be innately specified but are in fact, acquired by exposure to particular E-languages.

If I-language is completely specified in advance by virtue of us being human, then some argue that it provides humans with the medium for complex thought without any need for learning (see Chomsky, 2010). The details of particular language Total Forms would be irrelevant, unless they were used in conscious reasoning. But if by I-language is simply meant *all* the mental, internal, computational aspects of a natural language, then I-language should get more elaborated and sophisticated as language acquisition proceeds and is molded by E-language.

Consider now the specific proposal by Hinzen and Sheehan (2013), in which he discusses naturalization of the concept of Truth. He argues persuasively that truth is a property that emerges internally, from the syntax of natural language, not from reference or external considerations. In particular, he argues that anything less than a clause cannot have a truth value, that is cannot be evaluated as true or false: Noun phrases, Small Clauses, Infinitive Complements all lack the necessary structure. Truth comes about by way of Tense in a clause, but it alone is insufficient: a full CP phrase with Force is what is needed.

> "Truth is associated to only one of the phases just mentioned: the one that is most 'complete'…. The moment that completeness is reached, and the syntactic object generated is evaluated semantically, the derivation is cancelled: structure-building and recursion stop…..in particular, syntactic structures need to obtain semantic (referential, extensional) values. In this system, truth functions as a contingent ultimate limit imposed on recursive growth: it is the signal that calls a derivation to a halt". (p. 3, Hinzen, 2009; see also Hinzen, in press)

First, there is a claim that structure building proceeds in a cycle, and once a complete structure is built, it is passed off to the semantic component for evaluation.

> "An object described through a particular *nominal* predicate will be located in *space* (NP); an event described through a particular *verbal* predicate is located in *time* (*v*P); and a proposition described through a tensed proposition is located in *discourse* (CP)" (p. 24 Hinzen, 2009; see also Hinzen and Sheehan, 2013).

The fully constructed clause is thus the skeleton on which truth depends. There must be a full Thematic structure, and all of the arguments of the event-head have been appropriately filled (e.g. PUSH, Agent, Patient). If this develops into a noun (Bob's pushing of Alan), then it will never be evaluated for truth. But with a verb, once tensed ("Bob pushed Alan"), and merged with C, the structure becomes something that can be asserted in discourse and evaluated for truth value. Hinzen argues that the contact with the outside world of reference that this last step demands in no way contradicts the claim that the preconditions for this evaluation lie strictly in the grammar.

Let me sumup so far. All and only humans have the capacity for I-language, that is, syntactic structures that permit the representation of structured meanings at sentence-level complexity and above. The I-language has the properties of Universal Grammar that Chomsky describes, and allows language learning in a finite amount of time in human and only in human

species. The full acquisition of the I-system takes time, because it entails interacting with a community of speakers who convey the conventions to express ideas in a particular language. E-language thus shapes I-language in the developing child. As this is accomplished, I-language connects to the world. The nature of this relation is uniform across languages and conditions of development, if considered at an abstract level. In adults, I-language might operate even in the absence of E-language or its impairment or loss of systematic connections, as in aphasia (cf. Varley and Siegal, 2000). In shadowing tasks involving human language, I-language is entailed and so the normal course of reasoning with it is suspended (Hermer-Vasquez et al., 1999).

However, Hinzen's argument is more global. He argues that only a human brain comes equipped with the machinery necessary to create syntactic structures, and hence assertions. Even the simplest tensed sentence should be beyond the reach of animal cognition. In his work, he describes how the syntactic operations of external and internal Merge create a conceptual object that is a new entry on the evolutionary stage, e.g. (BOY (CHASE DOG)). How should this be interpreted, empirically?

Consider the following progression:

a) Many animals and all humans might be capable of forming the nominal concepts BOY, DOG, that is, recognizing that there are similarity classes in the world that correspond (roughly) to those words. Note however that even here, there are proposals that humans may use lexical codes to help perceptual discrimination at the level of DOG, CAT (Gilbert et al., 2008). Furthermore, the work by Xu (2007) suggests that early labels play a role in infants' ability to recognize objects as different sortals.
b) Some animals and all humans might be capable of forming the concept CHASE, though already here we enter a domain of considerable uncertainty. What is a verb, even a simple action verb, unfettered by particularities of e.g. certain agents? This is discussed in more depth below, but consider for certain "movie-type" – and intransitive verbs (Maratsos and Deak, 1995), it might be a movement pattern recognizable via point-light displays on the joints (Golinkoff et al., 2002). However, this is surely a case in the limit.
c) Only humans can recognize the event "BOY CHASE DOG", because it is a structured syntactic object whose arguments are in a certain specified relation to one another: it is not enough that the elements BOY, CHASE and DOG be present in the scene.

> "The algebraic relations in question hold between lexicalized and appropriately grammaticalized (in particular, syntactically categorized and phrasally configured) 'concepts' retrieved from long-term memory" (p. 10, Hinzen, 2009).

What does "algebraic' mean, here? As in Marcus (2001), it means that language constructs expressions that are constituted of variables, not particulars. Just as BOY does not refer to a specific boy but instead to the set of all boys, the composed form BOY CHASE DOG can potentially refer to an infinity of possible events in which a boy is chasing a dog. This linguistic structure thus *brings similarity* to a set of events in the world that had no unity before. This is a strong claim, and a good one in the sense that it is empirically testable.

In what follows I will try to flesh out the claim by considering particular empirical studies, and some preliminary evidence from children and adults in my laboratory.

2. Background to studies

As said, a fundamental claim about the nature of human language is that it is algebraic in nature, that is, the rules in question operate on variables. Take for instance, rules of syntax. The rules that govern sentences operate not on words, but on abstract entities like subject, verb, and object. Of course these are matters of theoretical dispute, and some researchers in language development have argued that the child might begin with rules operating over narrower scope, even to the level of words (Tomasello, 2000).

However, words themselves are variable in their reference: common nouns and verbs do not pick out individuals but classes or sets of entities or events. The word "dog" is not used just for one dog, but for any member of the class. It is the same with a verb like "push": not just one event of pushing, but any event with the right quality. There is good evidence that children learning nouns make this leap of abstraction strikingly early, and may even extend proper nouns to conform, as in calling all men "Daddy", or all dogs "Nunu" (de Villiers and de Villiers, 1979). It is widely acknowledged that at least the perceptual classes for the concepts underlying these noun generalizations are also established in the first year of life, for example, infants have been shown to dishabituate when a series of different exemplars of a basic object level entity such as "cow" switches into say, "dog" (Eimas and Quinn, 1994).

The evidence also seems compatible with the idea that children by the second year of life take a verb to be generalizable to new events despite the change in agents and affected objects, accumulating verbs such as "push", "hit", "eat", "like" and so forth (Tomasello and Merriman, 1995). The conceptual underpinnings are harder to evaluate, as infants tend to dishabituate to changes in the agent even when the action is kept constant in an event. For example, children aged 3 years can respond to the invariance afforded by point-light displays on actors where the agents change, provided the actions are a relatively stable shape – the kind Maratsos and Deak (1995) labeled "movie-verbs" (e.g. waving, skipping, *not* helping, fixing). Finding evidence that very young children can observe the similarity across novel simple motion events despite changes in the actor is more disputed (Waxman et al., 2009). There is a growing body of evidence that suggests that learning of actions and their

associated word forms may be quite demanding and take place relatively late in the word learning process (Hirsh-Pasek and Golinkoff, 2006). The work usually assumes that in order to learn words for actions infants must already have a firm conceptual grasp of those actions and the events they are a part of. That is, the difficulty in verb learning may stem from the mapping problem, not from any inherent challenge in conceptualizing actions and their events (Gentner and Borodtisky, 2009). According to Gentner, 'It is not perceiving relations but packaging and lexicalizing them that is difficult' (1982, p. 326).

Packaging means generalization, and children as old as three years may not find it easy to generalize a new transitive verb to a new object. For example, Imai et al. (2008) showed adults and children (3- and 5-year-olds) an actor doing a novel action to a novel object. They heard either a novel noun ("Look! There is an X!") or a novel verb ("Look! There is Xing!"). In generalization trials, two different events were shown, either with the same action done to a new object, or with a new action on the familiar object. The instruction was either to "Find the X" (noun condition) or "Find X'ing" (verb condition). It was no problem for adults and five year olds to generalize novel nouns and verbs beyond the particular events in which they had been introduced. This was not true for 3-year-olds with the verbs, which they did not generalize readily. The same result has been found in English, Mandarin, and Japanese (Imai et al., 2008). Waxman et al. (2009) describe this as being "captured" by the objects. Yet it is also likely that the meaning of the verbs is in flux, as the physical manifestations of say "jumping" or chasing" will undoubtedly change shape with the agents involved (see Hirsh-Pasek and Golinkoff, 2006).

There is some evidence that labeling the verb makes a considerable difference in whether children can see the class of events as the same (Pruden et al., 2005, 2008). Unlike common nouns, in which the category boundaries in the world may be already be clear prelinguistically, perhaps what counts as a common action may need an assist from language. Lexical variation across languages seems to confirm this conclusion, as languages as close as Spanish and English vary in what features they put in the verb as opposed to its satellites. English tends to encode manner of action in the verb, and direction in its prepositions or adverbs, and Spanish does the reverse (Talmy, 1975, 1985; Slobin, 1987; Hirsh-Pasek and Golinkoff, 2006). It would make sense for infants to keep an open mind about what to class together until the verbs in the input help them to define what the verbs should encode. One step in this event kind abstraction process appears to be in place by the child's second year.

But nouns and verbs are only half the story. A simple reversible sentence such as "the boy chased the dog" is a descriptor of a potentially infinite set of events in which myriad boys and dogs are entailed in acts of chasing. The crucial thing that it also encodes (in English, a language heavily dependent on word order) is that the boy in each case be the agent of the action, and the dog, the patient, or thing affected. The evidence is very strong that children as young as the second year of life attend to this directional difference when a sentence is spoken (Naigles, 1990). But what would the conceptual underpinnings consist of? We pose here a fundamental question: is there any sense in which we can recognize the resemblance across multiple "boy-chase-dog" events, prior to language encoding them that way? After all, we could ask the question about whether children can prelinguistically recognize the class of boys and dogs, and the answer seems to be yes. We can ask if children can prelinguistically recognize different acts of chasing, and the answer is probably, but not definitively, yes. Now we ask: is there any conceptual unity to the set of events of boys chasing dogs, that carves them out of from the set of *all* acts of chasing involving boys and dogs? Or does syntax give us something new, conceptually? We cannot easily think our way back into this problem, as adults, because we encode events automatically in language.

3. Study 1: Imitation of transitive reversible events: the role of language

The notion of a three term reversible, transitive event is a fundamental concept, yet we know little about how children acquire such concepts and their linguistic forms. Can infants learn to form such concepts of events like "*boy chasing dog*", prior to language? For instance, does the child see two such events as similar when the physical properties of the participants and therefore actions change? This capacity is necessary for language, in which sentences refer to classes of like events, but does this sort of abstraction of event kinds occur without language as a guide?

The kind of events that Golinkoff and Hirsh-Pasek assume to be present prior to language acquisition and the kinds of actions abstracted by 2-year-olds in Waxman et al.'s studies fall short of this, in that they do not involve a transitive action with two participants that are reversible in agency. Even if some kinds of events, for instance those involving figure–ground relations or a single actor performing a non-transitive action are available pre-linguistically, there might be others, namely reversible transitive events, which are only abstracted in the presence of language.

Nevertheless there is some evidence that even very young children under the age of two can represent and hold in mind complex events. Bauer and Mandler's (1992) work on children's recall of events suggests that children can hold in mind both the perceptual features of a particular event instance while also making a generalization about the participant kinds. Critically, 25-month-old children performed 6-step event sequences with analogous tokens of participant kinds (*bear*, *train*) at high levels immediately after the event was modeled with different participant objects. This result demonstrates that even 2-year-olds can generalize particular tokens of an object kind as participants in events, and can do so while correctly inserting the objects into a precisely-ordered sequence. Furthermore, even 11.5-month-olds infants are able to reproduce 2- and 3-step event sequences involving enabling relations in the correct temporal order, suggesting that they may indeed possess the memory capacities to reason about more complicated 3-term transitive events (Bauer and Mandler, 1992). What remains to be determined is the effect of the critical property of *reversibility*, that is, the role played by syntax in a three-term relation. In all the above studies, the actions were irreversible.

In Study 1, (reported in Hobbs et al., 2011) we asked whether young children could generalize an event to new participants. The methodology was inspired by a study by Bonawitz et al. (2010) on how young children encode causation. In that study,

Fig. 1. A still showing the demonstration from the procedure in Study 1.

the ability of 2 and 4 year olds to understand a causal connection was tested under one of two conditions, one in which participants observed co-occurrence of two events, and one in which participants also heard it described causally.

In our study, sixty-one 2-, 3-, and 4-year-old children participated in an event abstraction task. Their task was to watch and then replicate an event by acting it out with toys, but *the toys were not the same tokens*. Children watched an experimenter enact reversible, transitive events on a small stage covering a laptop keyboard (see Fig. 1) with two characters: one toy animal and one toy human. For example, one such event was a boy pushing over a tiger. After an event was demonstrated twice, the original toys were removed from sight and the child was presented with a tray containing four toys – different tokens of the same character *kinds* – say, a new boy and a new tiger – and two distractors, one human (e.g. a girl baby) and one an animal (e.g. a horse). Children were invited to "do the same thing", that is, to take the corresponding exemplars of a boy and a tiger and repeat the act in which a boy pushed over a tiger.

The first two events served as warm-up trials, in which children got up to three chances to generalize the event, with additional enactments if they failed to do it. Following these two warm-up trials children received no feedback during the 8 test trials. Children viewed the series of events in one of two orders.

As in Bonawitz et al., children were randomly assigned to one of two conditions – Language (L) and No-Language (NL). Children in the L condition heard an accompanying linguistic description of each event while it was enacted by the experimenter (e.g. "Look, a boy is pushing a tiger"). Children in the NL condition heard simply "Watch my show" before each event was enacted. As a "reinforcement", a laptop screen (under the experimenter's control) showed a series of stars and music if the child's reenactment matched the original event.

Fig. 2 shows some basic results on how many times the children of different ages enacted the eight events correctly under the two instruction conditions (L and NoL) The results revealed a highly significant effect of age, unsurprisingly, with older children performing much better than younger children ($F(2,55) = 27.0$, $p < 000$). However, there was also a significant effect of Language condition ($F(1,55) = 5.1$, $p < .03$). Children were much better at re-enacting the event when they heard the event described – the Language condition – than in the "No-language" condition where they received no linguistic scaffolding. There was a significant interactions of condition with age ($F(2,55) = 3.1$, $p = .05$), As the graph shows, 2-year-olds did not improve in the language condition; 3-year-olds benefited the most from the language condition; 4-year-olds were advantaged by language but to a lesser degree.

Furthermore, when linguistic *component* (nouns, verb, subject-object) was considered in the enactment, there was a highly significant main effect ($F(3,55) = 48.3$, $p < 000$), and interaction of component with age ($F(3,55) = 17.6$, $p < .000$). The 3 and 4 year old children were much better at generalizing to new actors, namely choosing another tiger and boy, than at choosing the correct verb (pushing) and direction of action (boy as agent, tiger as patient). For example a 3 year old might correctly select a new boy and a new tiger but then just position them next to each other with no ensuing act.

These results are certainly compatible with the idea that language provides an assist in holding onto a reversible transitive event to reproduce it with new exemplars. The 2 year olds could not yet benefit from the linguistic description, though the three years olds benefitted greatly. The suspicion is that the four year olds were encoding the events themselves, and needed less the assist that the linguistic encoding gave them.

Much more empirical work is needed on this paradigm. For example, what is the role of syntax versus the verb alone as an anchor: would providing the verb label alone in the language condition be sufficient? We did not test the children's language level, but merely assumed the four year olds would be more competent than the twos at describing the whole transitive events. A new study should differentiate the successors and failures by seeing who could describe the event themselves. We argue that this is a potentially fruitful line of questions that have not yet been addressed in development, namely how children see the resemblance across events and how language – their own or a caregiver's – assists in this generalization.

4. Study 2: Adults recognizing event similarity while shadowing

In a second study we turned attention to whether adult subjects could retain a memory of a pictured reversible transitive event while shadowing language, in other words, with their language faculty tied up. This methodology has been widely used in cognitive psychology studies to test whether language is needed for another task, as in the spatial navigation studies by Hermer-Vasquez et al. (1999), and by Newton and de Villiers (2007) to explore false belief reasoning in adults. The usual design is to have a control interfering task that has been matched to verbal shadowing for its general attentional demands on some second task, usually some kind of pattern recognition (Newton and de Villiers, 2007) or visual attention (Hermer-Vasquez et al.) task. In both of these studies, the control interference task was rhythmic pattern shadowing, using tapping as the performance. First, the two interference tasks are calibrated for general attentional difficulty on an independent task, that is, one that does not have any language requirements. Then, participants in the main study are assigned to either the verbal shadowing or this alternative non-verbal shadowing to assess whether having the language faculty "tied up" causes difficulty with the main task, over and above general attentional load. This was the procedure followed in Study 2.

A separate group of 30 adult participants aged 18–22 provided the baseline data for comparing rhythmic and verbal shadowing. For the primary task participants searched arrays of 16 faces to determine whether the faces were all oriented alike or if one face in the array was rotated in a direction 180° opposite to the rest. This task, though attentionally taxing, does not involve any reasoning about events, or mental states, and does not require verbal encoding.

Participants completed the visual search task while simultaneously engaged in another task. As this experiment was part of a larger study, participants underwent more conditions of interference than just the two that are of interest here. We used a repeated measures design, counterbalancing order of conditions. Interference conditions included rhythmic tapping, shadowing the same rhythms by saying *bababa* (i.e. repeating a nonsense phoneme to the rhythm), shadowing English, shadowing Swahili, or undergoing no interference. Subjects' reaction time and judgment were recorded for each of 32 arrays of faces.

We compared error rates and response times for the visual search task under different conditions of interference; here we present the data from only the English verbal shadowing and rhythmic tapping conditions. Overall, error rates were very low and did not discriminate among conditions: rhythmic shadowing and verbal shadowing had the same number of errors ($t(29) = 1.2$, $p = .23$). A further independent-samples t-test revealed that reaction times on the visual search task were not significantly different under conditions of verbal shadowing versus rhythmic tapping interference ($t(29) = -.382$, $p = .705$).

The results from reaction times show that the rhythmic tapping and English verbal shadowing tasks were equivalent, and reaction times increased by ~350 ms relative to no interference for each condition. Thus these particular English verbal shadowing and rhythmic tapping tasks are very well matched for difficulty and attentional demands on a control task.

In Study 2, the procedure was different for the adults than the child subjects in Study 1. Instead of asking adults to act out a "like" reversible transitive event with toys, participants were asked to select the picture that best fit a sample picture out of a set of options. For example, after seeing a still picture of a man pushing a woman in a wheelchair for three seconds, subjects had to hold this original picture in mind for six seconds while viewing a blank screen then choose one picture out of a set of foils. These foils were for instance:

a) Another man pushing another woman in a wheelchair (target: extension to new instances)
b) A woman pushing a man in a wheelchair (foil: reverse agent and patient)
c) A woman pushing a baby in a stroller (foil: right action, wrong agent and patient)

For half the participants in each group, we added a fourth distractor:

d) A man pushing a baby (foil: the wrong patient but correct verb)

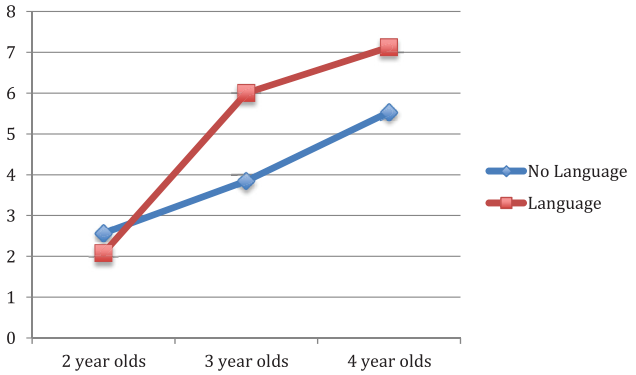

Fig. 2. Data from Study 1 showing the effects of age and condition on correct responses to generalizing a transitive action.

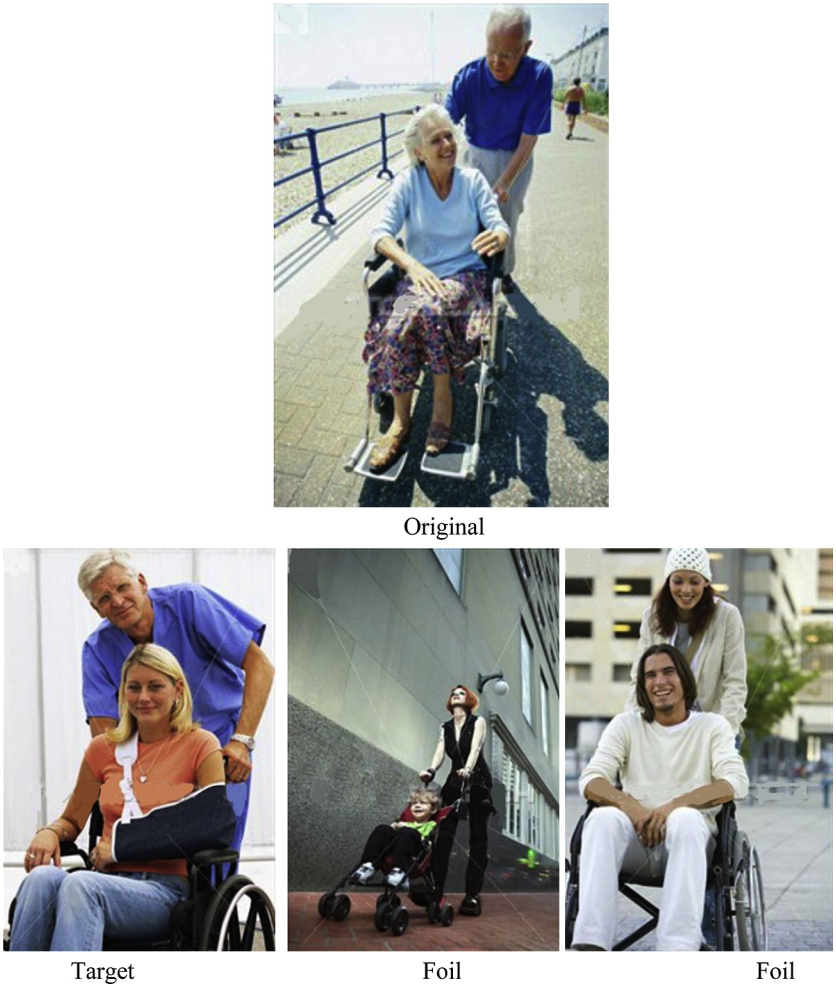

Fig. 3. A sample set of target and foils from Study 2.

All of these were natural photographs taken from the internet, and as such were as full of irrelevant details, colors and angles as any set of real life events. Each participant saw 20 stimuli each followed by a set of foils after six seconds (see Fig. 3). Before the main task began, and with no shadowing, each subject was shown a single instance of an event stimulus then the choices. If the subject chose wrongly, the right choice was pointed out. For example, if the subject chose the woman pushing the man, the experimenter pointed out that the original had a man pushing a woman. Through this training, participants recognized that the significant dimensions were e.g. gender and age and agency. Following that explanation, the procedure commenced, with the assigned interference task (or none) added.

Adult participants ($N = 63$) were allotted to one of three conditions at random: no shadowing ($N = 16$), verbal shadowing of a book passage ($N = 28$), and shadowing of rhythmic tapping ($N = 19$).

The two experimental groups were trained to shadow for the first two minutes of the study, to ensure that participants understood the importance of continually shadowing. In our experience the majority of adult college-level participants can grasp shadowing and begin to get fluent after a small amount of such practice.

Only participants who maintained shadowing throughout were used in the analysis, with the criterion that pauses were no longer than two seconds. The number of targets correctly chosen was then calculated. Adults with full access to their language faculty achieved approximately 85–90% correct on this task, making about 2 errors in 20 trials (Fig. 4). However, when shadowing verbal material, adult participants made 6/20 errors on average, i.e. their performance was reduced to 70% correct (a highly significant difference, $t(42) = 3.53$, $p < .0001$). From a post-hoc analysis of the stimuli, it can be argued that 70% is close to the performance one could achieve if a participant merely registered that e.g. a man was the agent of a particular event, without encoding the whole event. That might be possible *sans* syntactic encoding. There was no statistical difference in performance between participants who saw three or four choices (see Fig. 4).

The performance with verbal shadowing was significantly worse ($t(45) = 2.01$, $p < .05$) than the performance of participants engaged in rhythmic shadowing (77.5% targets correct), but rhythmic shadowing also reduces performance

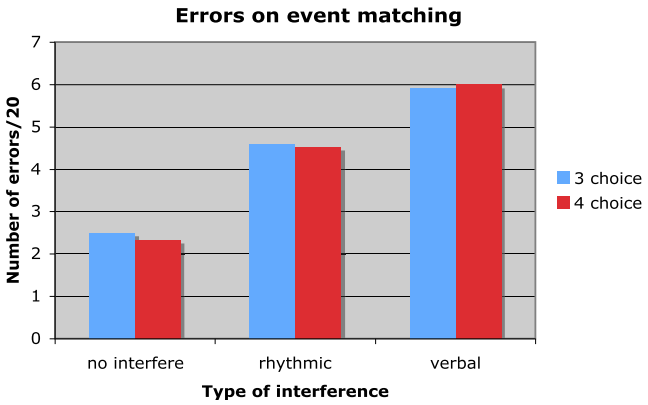

Fig. 4. Error data from adults in Study 2 matching pictures to target.

significantly compared to the no-interference condition ($t(33) = 2.7$, $p < .01$). Clearly these results are suggestive of an important role for language in recognizing event similarity, but there are executive demands of memory and retention that also come into play and cloud the picture.

In the traditional analysis, rhythmic shadowing is a control for the "attentional demands" of an interfering task. For that reason, both are compared first on some non-linguistic task. In our case, we used data from a separate group of adults whose task was to identify whether an array of 16 faces were all oriented the same way or not. We determined that the rhythmic and verbal shadowing were matched in general attentional difficulty based on RT and errors on that task. However new questions have been raised about the adequacy of this matching process, with some researchers now arguing that it is necessary to titrate the difficulty of each interfering task for each individual subject. Dungan and Saxe (2012) have also argued that one can manipulate the difficulty of rhythmic tapping so that it interferes as much as language shadowing, in their case with belief reasoning. Furthermore, shadowing rhythmic tapping has been shown to tie up the phonological loop (Larsen and Baddeley, 2003). If the phonological loop is needed for rehearsal, any task involving memory might be subject to interference from shadowing both rhythmic tapping and verbal material. Response selection in a choice task is also argued to require executive demands. A new technique is needed that gets beyond the "executive function" demands of a choice task, just as argued by Baillargeon et al. (2010) for the use of eyegaze with infants. The next studies use a procedure that removes potential confounds of executive demands such as short term memory and decision making/response selection.

5. Studies 3, 4 and 5: Implicit concept formation

In Study 3, an entirely new technique was invented to explore implicit concept formation with adults. Like work on implicit concepts in infancy, this work is motivated to make behavioral demands inconsequential, to see if *eye gaze* will be preferentially drawn to stimuli that embody some rule. In our case, we ask whether adults can form an implicit representation of a set of like events: "like" in the sense discussed above in which they are all potential manifestations of the same linguistic description.

6. Study 3: An implicit concept of like events: child kiss adult

The idea behind the studies was as follows. Suppose a person sees two still pictures side by side for a few seconds, and then one of the stills animates to become a small video. Then a second pair is presented, and one of them animates after a short while.[2] Will adults be motivated to predict the animation (i.e. by increased anticipatory looking at that side) across trials, if there is some commonality to the pictures that animate? We have shown that they will, across quite a range of different content. In Study 3 we asked whether a participant could attend to the difference between closely matched pairs of pictures, one in which a parent kisses a child, and the second with the same participants, in which a child kisses a parent. In short, the content to be attended to is a reversible transitive event, with the tokens constantly varying.

In our first effort, we made simple videos of parents and young children at a local day care center. To minimize distraction, we filmed each pair who volunteered against a plain backdrop and asked the parent to kiss the child on the cheek (so it was not "reciprocal" kissing) and then to solicit a kiss from the child in the same manner. We then cropped the videos to make matching pairs in which one depicted "child kiss parent" and the other "parent kiss child", one above the other (see Fig. 5).

[2] This lovely idea I owe to Katherine Hobbs. I had suggested flashing some kind of reward on the screen, but it was her idea to use animation as the natural reward.

Fig. 5. A pair of stills from implicit concept formation Study 3, "Adult kiss child".

Stills from these videos constituted the initial picture pair that the participants saw in a Tobii 1750 Eyetracker. The program presented the stills for three seconds and then one of the stills – the one showing the child kissing the parent – animated for three seconds. The first four such stimuli were identical, contained the same pair of participants in each still, and only varied across trials in the position of the animating still, and the left-right orientation of the parent and child. The next 21 pairs then counted as "generalization" trials, to see if the participants had formed a general rule such as "child kiss parent". In these 21 test trials the participants in each still were different individuals (as in Fig. 5). In 4 trials the foil was a picture with one element "odd", such as a parent kissing a teddy bear. These were included to demonstrate to the participants that the characters mattered. Seventeen of the test pairs were the critical stimuli to analyze, in which the difference was *child kiss adult* versus *adult kiss child*.

Participants were seated at the eyetracker and instructed simply to keep their eyes on the screen and to watch what happened. 27 female college students served as participants. 17 were in a no-shadowing condition. A second group was first trained to shadow and then did simultaneous verbal shadowing of a passage read from a book. Data were only analyzed for those ($N = 10$) participants who maintained shadowing with no more than a two second pause throughout the experiment.

We compared looking time to the target versus the non-target in the three seconds immediately prior to animation of the still, that is the *anticipatory* gaze at the right event. Comparing looking time to target versus nontarget for the nonshadowing group of subjects revealed significant differences in looking time (Paired $t(16) = 4.1$, $p = .0001$). In the absence of shadowing, adults easily learn, and are motivated, to look in the direction of the still that depicted the "same" event over trials. However, when verbally shadowing, adults did not look consistently at the right still, in fact they were reduced to chance levels (Paired $t(9) = .125$, $p = .8$). In the verbal shadowing condition, with language tied up, these participants could not track the commonality across the events.

The traditional measure of differential looking score (DLS) was also used to compare looking time to chance. DLS is the measure which takes into account total looking time (target − nontarget)/(target + nontarget), and should be zero if looking time is not different. DLS was different from zero for the non-shadowers (DLS = .19; one sample $t(1,16) = 4.468$, $p < .001$), but not remotely so for the shadowers (DLS = −.001,n.s), and a univariate ANOVA on DLS confirms the two groups are different on this measure ($F(1,26) = 5.23$, $p = .03$).

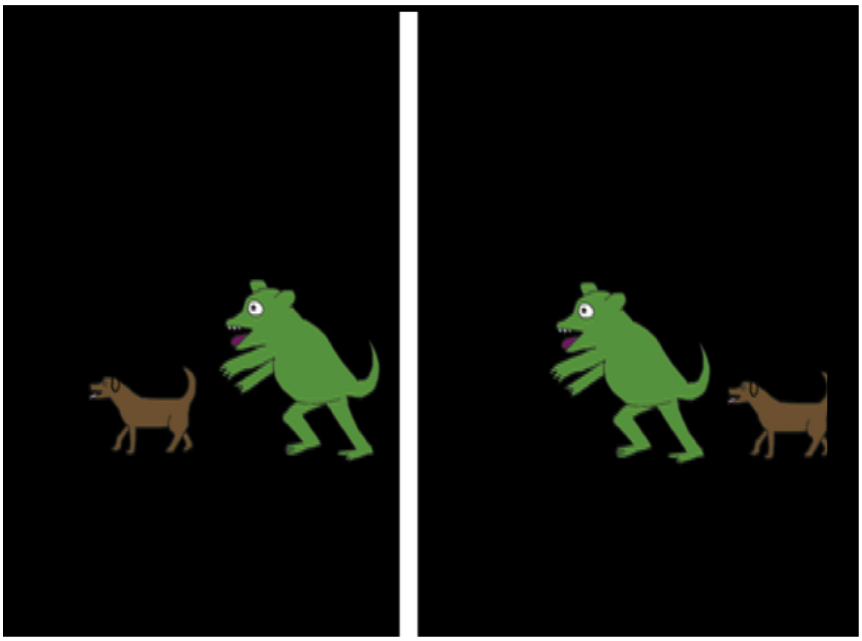

Fig. 6. A pair of stills from implicit concept formation Study 4, "Monster chase dog".

Participants were asked after the study what they thought it was about, and even though the non-shadowers had tracked the event type successfully with their anticipatory gaze, they seemed unaware of the underlying similarity across events.

On the one hand it is compelling that adults formed this implicit concept of the same event with natural stimuli of considerable variation, but on the other hand, there is a lack of control with natural photographs: could participants' attention be drawn to something less than the event itself, such as "adult with lips pursed"? For that reason Study 4 carefully controlled the form of the videos by using animation.

7. Study 4: An implicit concept of a like event: monster chase dog

For Study 4, the variance in natural events was sacrificed for the control afforded by animation. With the same design as Study 3, simple animations of monsters and dogs were created, in which both creatures moved across the scene once the animations started. Each scene consisted of a pair of stills in which a monster was chasing a dog or a dog was chasing a monster. After 2 s the target animated. The creatures were simple line drawings varying in size, color, and with small variations in shape consistent with remaining the same species (see Fig. 6 for an example). The pair either faced in a left to right direction or a right to left direction. Each subject watched 40 videos: 10 training videos in which the same target changed orientation and position, 20 novel monster-chase-dog vs. dog-chase-monster videos, and 10 with the non-target as e.g., monster-chase-cat or elephant chase-dog as a control to ensure that the subjects were recognizing both the agent and the patient rather than only one element. 32 female college students served as participants, 16 with no shadowing and 16 who were successful with verbal shadowing, that is, those participants who maintained shadowing with no greater than a two second pause. We analyzed the data from the periods right before animation – assuming that the first second of the still would be needed to look at both sides, and then the next second would be sensitive to anticipation of the animation.

The results replicated Study 3 with more tightly controlled stimuli: adults without interference were easily able to discover and track the similarity across events, looking more in anticipation at the target picture than the non-target (Paired $t(15) = 2.99$, $p = .009$), but not at all while verbally shadowing (Paired $t(15) = 1.65$, $p = .12$). The DLS was highly significant compared to chance (0) for the non-shadowers ($t(15) = 3.08$, $p < .008$) but not for the shadowers ($t(15) = 1.42$, $p = .18$). A univariate ANOVA on DLS confirms the two groups are different on this measure ($F(1,30) = 4.37$, $p < .05$).

Studies 3 and 4 used a technique that allowed us to look at implicit event concepts that are based on fundamental syntactic representations of the sort that might be provided by I-language, and the results reveal that adult subjects are easily drawn to track the equivalence across events when the tokens change, though virtually none of then could express afterwards what they were doing. Many subjects express a false generalization, such as "It was always the one on the left that moved". There was no requirement to consciously direct attention in this task, even though we did instruct subjects to watch carefully. Nothing about the implicit concept task requires explicit rehearsal, nor response selection, so executive demands should be minimal. But it is still possible that shadowing might have the result it does not by tying up the language faculty, but instead by re-deploying attention away from the main task. Since rhythmic shadowing has several strikes against it as a control, Study

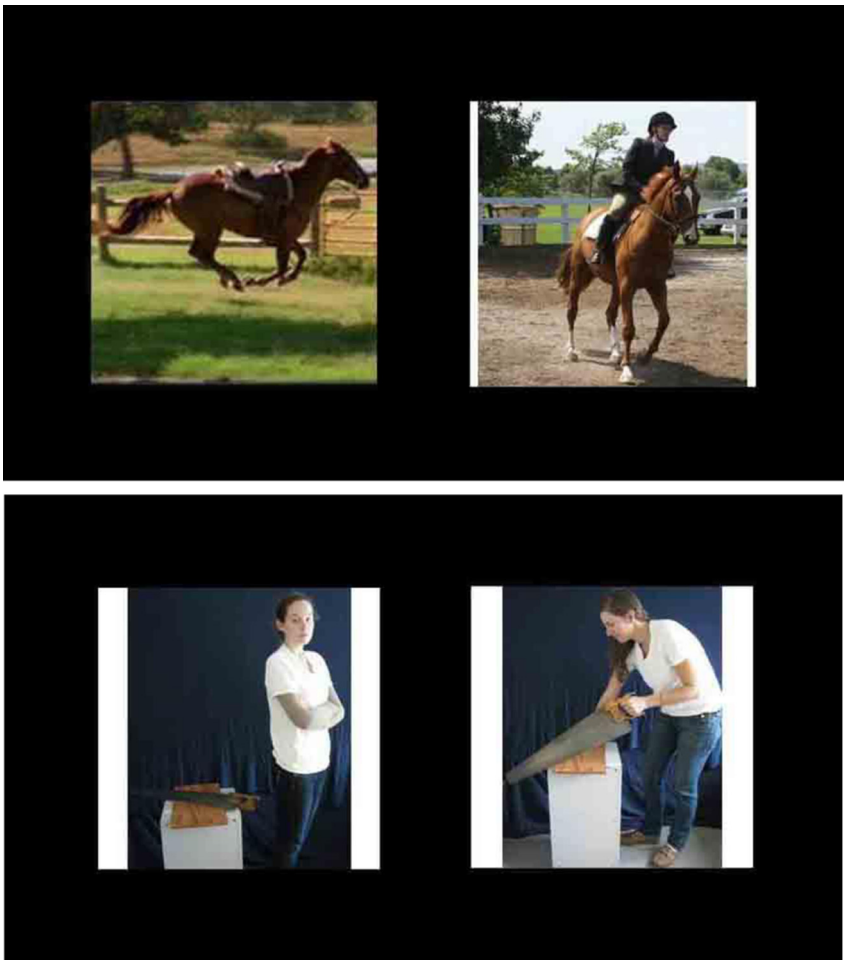

Fig. 7. Two sample pairs of negation in Study 5, showing nonexistence and refusal.

5 approached the question in an entirely new way, by asking what kind of abstract concepts does linguistic shadowing NOT disrupt? That is, if shadowing just causes participants to relax attention to the stimuli, then they should not succeed at stimuli that are abstract or difficult, even if they are not united by any kind of propositional description. However, if shadowing specifically ties up the ability to unite stimuli via propositions, then it might not interfere with other abstract implicit concept formation.

8. Study 5: The implicit concepts of negation, and natural kinds

Ann Nordmeyer[3] (2011) used the same design as in Studies 3 and 4 to do a critically different experiment using verbal shadowing. Instead of looking for a difference in the disruptive effects of verbal versus rhythmic shadowing, she compared two abstract concepts and predicted on theoretical grounds that one would be disrupted by verbal shadowing and one would not.

The starting point was an interest in negation: a linguistic concept *par excellence.* Negation was one of the key arguments for why thought must have a propositional rather than merely an imagistic basis, because one cannot imagine a negative picture. Fodor used the example "Picture a man who is not standing next to a giraffe". It is easy to draw such a picture, but the only thing that makes it a picture of THIS is with the associated verbal description. Without that, it could be just a man, or a man standing alone, or a man standing alone with a mournful giraffe at some distance, or even a man sitting next to a giraffe…..Pictures don't suffice. At the same time, negation demands a "context of plausible denial" (Wason, 1965), that is, the negation is a contrast with an affirmative: a cancellation of a presupposition, a failure, something missing that was expected. That is, there are varieties of negative meaning painstakingly elucidated for child language by Pea (1980) and for adult

[3] Ann Nordmeyer was a student at Smith College who did an honors thesis with the author in 2011, and is now a graduate student at Stanford.

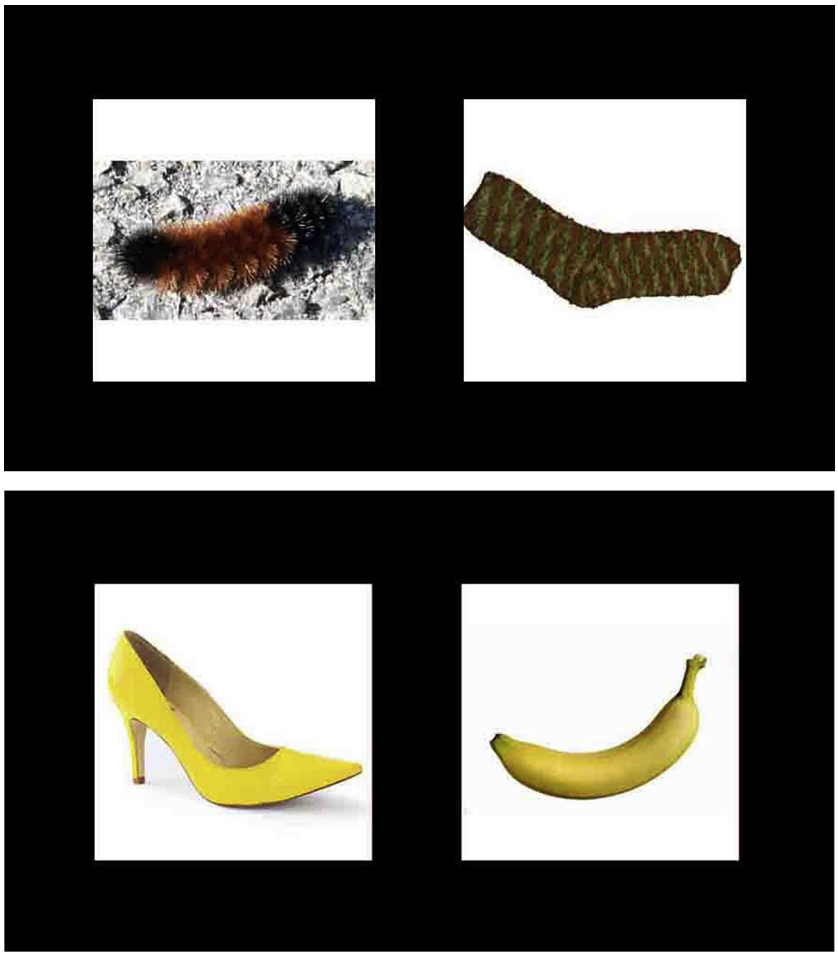

Fig. 8. Two sample pairs of natural kinds versus artifacts from Study 5.

semantics by Horn (1989) and others. Nordmeyer reasoned that seeing equivalence across contexts of plausible denial was an abstract task, and one most probably mediated by I-language.

To test this in adults, Nordmeyer set up pairs of pictures, one of which set up a contrast context for the other (see Fig. 7). In 7a, the horse does not have a rider but looks like it should, as in 7b. In 7c, a woman is not sawing, unlike the woman in 7d. She had examples representing refusal (e.g. the sawing), incompleteness (horse that has lost its rider), not-functioning-as-expected (fires contrasted with a match going out instead of lighting the paper), She represented the various forms of negation in this way, and the rule was that the one corresponding to the *negative* description would transform into an animation.

In this case, given the complexity of the stimuli we could not engineer natural videos, so we created a Monty Python facsimile: a large bare foot emerged after 3 s and squashed the (negative target) scene then retreated. This proved highly motivating to watch. There were 8 training scenes with the same four pairs varying in position and orientation, followed by 22 generalization pairs on which the differential looking time was assessed during the three seconds before the animation appeared, to measure anticipatory looking.

As the control stimuli, Nordmeyer used pairs of stimuli that were similar in color, shape and sometimes texture, but one was a natural kind and one was an artifact. These included considerable variety, not just animate and inanimate (a helicopter and a bird), but also both inanimate (a tennis ball and a green apple), but e.g. landscape features such as a fountain (artifact) and a waterfall (natural). The concept to be determined was "natural kind". This was chosen because of the likelihood that the difference is salient to us even in the absence of language, given studies with infants, and animals. However, no-one has previously asked whether the global *concept* of natural kind was available to adults, just that particular classes of natural kinds were easier to acquire than similarly diverse classes of artifacts (Herrnstein, 1985). Despite its abstractness, the concept of natural kind is not a propositional one, and so we reckoned that language should not be needed for forming the equivalence class. Hence, if adults could do this at all, verbal shadowing would not disrupt it (see Fig. 8).

Nordmeyer (2011) and our subsequent follow-up tested 35 participants on the negative task, the shadowers (17) of which passed the usual criteria of shadowing without pause. On the identically configured natural kinds task, 47 participants took part, 27 non-shadowers and 20 of whom were competent shadowers.

The results were stunning confirmation that verbal shadowing is selective in its effects. Adults took about the same (small) number of trials to "catch on" to the idea of natural kinds versus negation when they were not shadowing, and then looked selectively at the still that was likely to animate rather than the non-target (Negation: Paired $t(17)$, $p = .0009$; Natural Kinds: Paired $t(26)$, $p = .001$). The shadowers in the natural kinds task still looked selectively at the natural kind stills (Paired $t(19)$, $p = .038$). The shadowers in the negation task failed completely to register the property that the negative stills had in common, looking equally often at target and nontarget (Paired $t(16)$, $p = .94$). We argue that this is because *the negative pictures have nothing in common without language mediation.*

Importantly, Nordmeyer's control can just as easily be considered a control condition for Studies 3 and 4, in that the same basic design was used, and the same methodology and participant group. This lends support to the finding that *verbal shadowing selectively interferes with the recognition of equivalence classes that depend on language*, and does not have its effects at some more general level like deploying attention away from the task, or getting in the way of response selection.

There are two broad generalizations emerging from this initial empirical work. First, some concepts have a similarity that emerges at a linguistic level, dependent on propositions (either reversible transitive events, or negation). These concepts are easily recognized by adults who have full control of their language faculty. Second, these same concepts are selectively disrupted by verbal shadowing compared to concepts that are abstract but do not depend on propositional representation. This is a new idea, and one that needs to be carefully explored empirically.

9. Conclusion

In this paper I have laid out some arguments and evidence that I-language makes possible some concepts that might not exist in a non-linguistic creature. These include the concept of the class of events describable by a three term reversible sentence, namely one in which the hierarchical arrangement of arguments defines the agency, namely, the direction of action. It also includes the concept of negation. It is not clear what else it includes, for example the concept of AND seems much too primitive to need a linguistic mediation. In addition, the process by which features bind naturally into a percept seems outside of this special set.

More broadly, new evidence suggests that relational properties suffer under verbal shadowing, whereas featural properties do not (Jung and Hummel, 2013). They argue that stimuli whose resemblance is relational (e.g. X behind Y) are disrupted, but featural resemblance (black, striped) survives shadowing. There is much new work to be done to differentiate these alternatives, but on the surface, it does not seem obvious that artifacts and natural kinds are distinguishable by featural properties.

Imagine the concept of a *curly-red-haired-young boy*. One can easily imagine a set of such pictures in which the similarity would not need propositional analysis – they would just "look alike" – after all, is that not how an animal recognizes its specific prey? However, scenes that contain relational properties such as "red on the left, green on the right" might be precisely the kind that demand linguistic representation, and linguistic encoding has been shown to help four year olds on this task (Dessalegn and Landau, 2008).

The studies I have reported here represent the beginnings of a research program. Perhaps these initial results will inspire more research, and in the process they may be proved mistaken. Clearly the research must expand beyond preschool children and college students, and there are obvious studies to be tried with primates as well as with infants. In particular, does the language faculty alone give infants the edge over primates, even when language is not yet actualized? Or does some actual natural language have to be present, as in Study 1, to assist in the recognition of similarity? There is much exciting empirical work to be done.

References

Baillargeon, R., Scott, R.M., He, Z., 2010. False-belief understanding in infants. Trends Cogn. Sci. 14 (3), 110–118.
Bauer, P.J., Mandler, J.M., 1992. Putting the horse before the cart: the use of temporal order in recall of events by one-year-old children. Dev. Psychol. 28, 441–452.
Bonawitz, E.B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A., Woodward, J., Schulz, L., 2010. Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. Cognition 115, 104–117.
Carruthers, P., 2002. The cognitive functions of language. Behav. Brain Sci. 25, 657–719.
Chomsky, N., 2010. Some simple evo-devo theses: how true might they be for language? In: Larson, R.K., Deprez, V., Yamakido, H. (Eds.), The Evolution of Human Language. Cambridge University Press, Cambridge.
Collins, J., 2000. Theory of mind, logical form and eliminativism. Philos. Psychol. 13, 465–490.
Davies, M., 1998. Language, thought and the language of thought (aunty's own argument revisited). In: Carruthers, P., Boucher, J. (Eds.), Language and Thought: Interdisciplinary Themes. Cambridge University Press, Cambridge.
de Villiers, P.A., de Villiers, J.G., 1979. Early Language, The Developing Child Series. Harvard University Press.
Dennett, D., 1991. Consciousness Explained. Little, Brown & Co., Boston.
Dessalegn, B., Landau, B., 2008. More than meets the eye: the role of language in binding and maintaining feature conjunctions. Psychol. Sci. 19 (2), 189–195.
Dungan, J., Saxe, R., 2012. Matched false-belief performance during verbal and nonverbal interference. Cogn. Sci. 36, 1148–1156.
Eimas, P.D., Quinn, P.C., 1994. Studies on the formation of perceptually based basic-level categories in young infants. Child Dev. 65, 903–917.
Fodor, J., 1975. The Language of Thought. Harvester Press, New York.

Fodor, J., 2008. LoT2: Language of Thought Revisited. Oxford University Press, Oxford.
Maratsos, M., Deak, G., 1995. Hedgehogs, foxes and the acquisition of verb meanings. In: Tomasello, M., Merriman, W. (Eds.), Beyond Names for Things: Young Children's Acquisition of Verbs. Psychology Press.
Frankish, K., 2004. Mind and Supermind. Cambridge University Press, Cambridge.
Gentner, D., 1982. Why are nouns learned before verbs: linguistic relativity versus natural partitioning. In: Kuczaj, S.A. (Ed.), Language Development: Language, Thought and Culture, vol. 2. Lawrence Erlbaum Associates, pp. 301–334.
Gentner, D., Borodtisky, L., 2009. Early acquisition of nouns and verbs: evidence from Navajo. In: Gathercole, V. (Ed.), Routes to Language: Studies in Honor of Melissa Bowerman. Taylor and Francis, New York, NY, pp. 5–36.
Gilbert, A.L., Regier, T., Kay, P., Ivry, R.B., 2008. Support for lateralization of the Whorf effect beyond the realm of color discrimination. Brain Lang. 105, 91–98.
Golinkoff, R.M., Chung, H.L., Hirsh-Pasek, K., Liu, J., Bertenthal, B.I., Brand, R., Maguire, M., Hennon, E., 2002. Dev. Psychol. 38 (4), 604–614.
Hermer-Vasquez, L., Spelke, E.S., Katsnelson, A.S., 1999. Sources of flexibility in human cognition: dual-task studies of space and language. Cogn. Psychol. 39, 3–36.
Herrnstein, R.J., 1985. Riddles of natural categorization. In: Weiskrantz, L. (Ed.), Animal Intelligence, vol. 7. Clarendon Press, Oxford, U.K, pp. 129–144.
Hinzen, W., 2006. Mind Design and Minimal Syntax. Oxford University Press, Oxford.
Hinzen, W., 2009. The naturalization of truth. Unpublished manuscript.
Hinzen, W., 2014. Recursion and truth. In: Roeper, T., Speas, M. (Eds.), Recursion: Complexity in Cognition. Springer, Dordrecht.
Hinzen, W., Sheehan, M., 2013. The Philosophy of Universal Grammar. Oxford University Press, Oxford.
Hirsh-Pasek, K., Golinkoff, R.M. (Eds.), 2006. Action Meets Word: How Children Learn Verbs. Oxford University Press.
Hobbs, K., Nordmeyer, A., 2011. Repeating a Reversible Transitive Event: Language Helps 3-year-olds Remember Actions. Paper given in symposium *The Role of Language in Event and Action Concepts, SRCD*, Montreal.
Horn, L.R., 1989. A Natural History of Negation. Chicago University Press, Chicago.
Hutto, D., 2008. Folk Psychological Narratives. MIT Press (Bradford Books), Cambridge, Ma.
Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R., 2008. Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. Child Dev. 79 (4), 979–1000.
Jung, W., Hummel, J., 2013. The effects of dual verbal and visual tasks on featural vs. relational category learning. In: Proceedings of the 35th Annual Conference of the Cognitive Science Society.
Larsen, Baddeley, A., 2003. Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: do they have a common source? Q. J. Exp. Psychol. 56A (8), 1249–1268.
Learmonth, A.E., Nadel, L., Newcombe, N.S., 2002. Children's use of landmarks: implications for modularity theory. Psychol. Sci. 13 (4), 337–341.
Marcus, G., 2001. The Algebraic Mind. MIT Press, Cambridge, MA.
Naigles, L., 1990. Children use syntax to learn verb meanings. J. Child Lang. 17, 357–374.
Newton, A., de Villiers, 2007. Thinking while talking: adults fail non-verbal false belief reasoning. Psychol. Sci. 18, 574–579.
Nordmeyer, A., 2011. To Be or Not to Be: Examining the Role of Language in a Concept of Negation (Honors thesis). Smith College.
Pea, R.D., 1980. The development of negation in early child language. In: Olson, D.R. (Ed.), The Social Foundations of Language and Thought. W.W. Norton & Co, New York, NY, pp. 156–186.
Pinker, S., Jackendoff, R., 2005. The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). Cognition 97 (2), 211–225.
Povinelli, D., Vonk, J., 2004. We don't need a microscope to explore the chimpanzee's mind. Mind Lang. 19, 1–28.
Pruden, S.M., Hansell, N., Albertson, G., Hirsh-Pasek, K., 2005, April. Finding the Invariant Path and Manner in Motion Events: Do Labels Help?. Poster presented at biennial meeting of the Society for Research in Child Development, Atlanta, GA.
Pruden, S.M., Hirsh-Pasek, K., Golinkoff, R.M., 2008. Current events: how infants parse the world and events for language. In: Shipley, T.F., Zacks, J.M. (Eds.), Understanding Events: How Humans See, Represent, and Act on Events. Oxford University Press, New York, NY, pp. 160–192.
Segal, G., 1998. Representing representations. In: Carruthers, P., Boucher, J. (Eds.), Language and Thought. Cambridge University Press, Cambridge.
Slobin, D., 1987. Thinking for speaking. In: Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, pp. 435–445.
Smolensky, P., 1990. Connectionism, constituency, and the language of thought. In: Loewer, B., Rey, G. (Eds.), Meaning in Mind: Fodor and His Critics. Basil Blackwell, Oxford, UK, 1991.
Spelke, E.S., 2003. What makes humans smart? In: Gentner, D., Goldin-Meadow, S. (Eds.), Advances in the Investigation of Language and Thought. MIT Press Cambridge, MA.
Talmy, L., 1975. Semantics and syntax of motion. In: Kimball, J. (Ed.), Syntax and Semantics, vol. 4. Academic Press, New York, pp. 181–238.
Talmy, L., 1985. Lexicalization patterns: semantic structure in lexical forms. In: Shopen, T. (Ed.), Language, Typology and Syntactic Description: Vol. III. Grammatical Categories and the Lexicon. Cambridge University Press, New York, pp. 57–149.
Tomasello, M., 2000. The item based nature of children's early syntactic development. Trends Cogn. Sci. 4, 156–163.
Tomasello, M., Merriman, W., 1995. Beyond Names for Things: Young Children's Acquisition of Verbs. Psychology Press.
Varley, R., Siegal, M., 2000. Evidence for cognition without grammar from causal reasoning and "theory of mind" in an agrammatical aphasic patient. Curr. Biol. 10, 723–726.
Wason, P.C., 1965. The contexts of plausible denial. J. Verbal Learn. Verbal Behav. 4 (1), 7–11.
Waxman, S.R., Lidz, J., Braun, I.E., Lavin, T., 2009. Twenty-four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. Cogn. Psychol. 59 (1), 67–95.
Xu, F., 2007. Sortal concepts, object individuation, and language. Trends Cogn. Sci. 11 (9), 400–406.