Coming from a World without Objects - Reconsidering Infants' Object Individuation

Journal:	*Mind & Language*
Manuscript ID	Draft
Wiley - Manuscript type:	Original Article
Keywords:	object individuation, demonstratives, kind-bias, reference
Abstract:	Children's ability to categorize and individuate objects puts an in-principle limit on the ways in which they can learn about their environment. While research on object individuation assumes that children are able to conceptualize objects as particulars, we argue that the violation-of-expectation paradigms that are commonly used in this context only illustrate that children are capable of feature-based discrimination, while this operationalization is inconclusive with regard to children's actual ontology. In contrast, we propose that children start out navigating the world with a feature-based ontology and only later become able to individuate objects by understanding them as individuals with unique spatiotemporal coordinates or histories. We propose that object individuation is a cognitively demanding achievement resting on a uniquely human form of enculturation, namely the acquisition of a feature-independent sign-system with strictly inter-defined terms. We argue that the acquisition of deictic demonstratives (this / that) or local adverbs (here / there) is the fundament for developing spatiotemporal object individuation. Our account reinterprets a central tenet of natural pedagogy. Infants' kind-bias might stem from their inability to refer to particulars rather than from an act of conceptual generalization. We conclude outlining empirical expectations for behavioral operationalizations of key ramifications of our proposal.

Introduction

The advent of language acquisition and cultural learning lies in infants' capacity to triangulate, i.e., in the ability to share attention to an object with an adult interlocutor. Within this referential triangle of child, caregiver, and object, infants can learn *about* the environment by following their interaction partner's actions and emotions as well as by connecting the language they hear to salient and relevant aspects of a situation. Developmental research has made remarkable contributions to our understanding of the social-cognitive processes involved in the coordination of minds necessary for triangulation. The "social side" of triangulation (child-adult) — its functional principles and effects for the development of shared intentionality — has been extensively investigated (Baldwin & Moses, 2001; Gliga & Csibra, 2009; Tomasello, 2010; Tomasello & Carpenter, 2007; Tomasello, Carpenter, & Liszkowski, 2007). The "referential side" (child-object) of triangulation, however — its mechanics and impacts on skills of referring to objects — has received much less attention (but see Mattos & Hinzen, 2015). Instead, the "referential side" of triangulation is considered a precondition of shared intentionality. An implicit assumption seems to be that in situations of shared attention the child manages to conceptualize the object in the same way as her adult interaction partner.

While there has been research about the development of object individuation in the last decades (i.e. Spelke, Kestenbaum, Simons & Wein, 1995; Xu & Carey, 1996; Van de Walle, Carey & Prevor, 2000), only some studies focused on individuation within triangulation, i.e. on reference. Xu & Carey (1996) and Xu, Cote & Baker (2005) found that 9-month-old infants succeeded in identifying objects by features earlier when objects were verbally referred to. But above all, it was shown that pre-linguistic children, when presented with verbal information about objects in triangulation situations, do not understand this information as episodic, i.e. as information about the specific object in question (Yoon, Johnson & Csibra 2008; Csibra & Gergely (2009); Hernik, Sperber & Gergely 2018). Instead, children systematically "generalize" information they receive to any object with similar surface features. Authors conclude that young children interpret information about objects in a generic way. Building on empirical evidence, Csibra & Gergely (2009) argued that this *kind-bias* is part of a uniquely human capacity for cultural transmission that became famous under the term 'natural pedagogy'. As a central hypothesis of

natural pedagogy, the authors propose that "children expect to learn something generalizable in ostensive-referential contexts rather than just become informed about particular episodic facts that obtain only in the 'here-and-now'." (Csibra & Gergely 2009, p. 151). Interestingly, Csibra & Gergely interpret this as a human-specific evolutionary social adaptation all the while one could see it as the missing ability of young children to understand information about specific objects. This interpretation is more conservative insofar as no additional cognitive capacities are postulated. The inability to learn episodic information is simply a result of children's immature understanding of objects per se. Infants' tendency to generalize information could stem from an inability to conceive of objects as particulars. Seen in this light, what research should focus on is not young children's tendency to generalize information early on but rather how they are eventually able to acquire knowledge about particulars at all.

In this article, we are going to critically discuss the assumption that infants make reference to particular objects in the same way as adults do. We present arguments to highlight that reference to objects is a complex capacity which researchers should not simply presuppose infants to have. Rather it appears unlikely that very young children partition the world into objects and their properties and relations. We mainly follow Evans (1982) and Tugendhat (1976) in this line of argument. Second, a number of studies is discussed which, starting with Spelke et al. (1995), are commonly taken to show that infants conceive the world as consisting of objects and their properties and relations. It is argued that this interpretation is not warranted and that the empirical findings are well consistent with an alternative and more parsimonious interpretation according to which infants do not have an objects-and-properties ontology and are only sensitive to perceptual features, their similarities, and patterns of features. Indeed, as Hirsch (1997) argues, such experiments, which show that test subjects behave in ways that would be expected from any subject sharing our ontology, are in principle incapable of showing that someone indeed shares our ontological commitments. Similar behaviour can be brought about by subjects with radically different ontologies. Only differences of behaviour can be considered as indicators of *different* ontologies. Thus, what researchers should be doing is devising experiments that aim to show that test subjects violate expectations derived from our ontological commitments in a systematic way. We conclude with suggestions for experimental protocols that would provide

evidence for children possessing an ontology that is *different* from ours. Finally, we highlight the relevance of our discussion for related debates.

The complexity of reference

From a common-sense perspective, it seems counterintuitive that young children lack the ability to individuate objects while displaying the ability to classify objects according to their feature similarities. Also, the view that object individuation precedes object classification according to features, i.e., being able to produce terms for basic level categorizations when seeing an object, receives a lot of support from an influential philosophical tradition. For example, Augustine described language acquisition as memorizing name-object assignments. According to his account, adults refer to objects and name them in a spatiotemporal world, and children learn a language by learning these correspondences between names and objects. For Augustine, the ability to refer to singular objects is just taken for granted. The Augustinian account, however, has been sharply criticized. In particular, Wittgenstein (1969) points out that many of the words we use simply do not refer to spatiotemporal objects or events and thereby shows that such a theory is untenable.

Moreover, questions as to how reference to objects is established by using proper names have caused frequent concerns in twentieth century analytic philosophy (cf. e.g. Frege, 1892; Russell, 1905; Searle, 1958; Strawson, 1959; Quine, 1960; Donnellan, 1966; Tugendhat, 1976; Kripke, 1980; Evans, 1982). A key challenge is to explain how proper names receive their meaning. Fregeans interpret proper names not as semantically simple but as complex expressions. In order to establish which object a proper name (e.g., Angela Merkel) stands for, we have to take recourse to descriptions of the particular object (e.g., Chancellor of Germany). Subsequently, this idea has been challenged. Kripkeans argue that proper names only have their referent as their meaning and no additional descriptive content. The referent, in turn, is fixed by an original act of naming. All future uses of the same name must stand in an appropriate causal connection to this original act. Tugendhat and Evans, then, give a broader picture including different kinds of singular terms, not just proper names, and including different mechanisms for fixing a referent, not just original acts of naming and causal chains. Evans (1982), for instance, makes room for descriptive contents again, by suggesting a hybrid theory. Evans as well as Tugendhat (1976) do not only discuss the

functions of proper names and different kinds of descriptions but also those of demonstratives, pronouns, as well as spatial and temporal adverbs. Their remarks make clear that reference to objects is a highly complex capacity which is key to understanding propositional thought and should certainly not be taken for granted.

Here is an adaptation of the central line of argument from Evans (1982) and Tugendhat (1976) that leads to the view that reference to objects is a complex capacity. Its acquisition requires learners to go through a series of steps:

1. Referring to features is cognitively less demanding than referring to objects. In order to refer to features, one needs the ability to perceptually discriminate and compare features, to register feature similarities, distinctions and changes and to react systematically to different features in specific ways. Reference to objects presupposes reference to features but not vice versa.

2. Reference to and individuation of objects requires a notion of identity which in turn presupposes a frame of reference that provides identity criteria. Common objects are minimally considered as units that move as connected wholes. They are solid, cohesive bodies moving continuously in space and time (cf. Cacchione et al. 2017, p. 580). This minimal notion of an object presupposes that places are considered as the same or different and that the same feature can be detected in different places. Furthermore, without a notion of identity over time one cannot have a notion of movement. There has to be something that remains the same over time in order for it to move. Without a notion of identity, one could only notice processes of feature changes.

3. Insofar as spatiotemporal coordinates provide the identity criteria for common objects, the acquisition of a notion of identity is initiated by learning a spatiotemporal coordinate system. Such an abstract coordinate system is different from feature-based spatial orientation capacities and provides the means to conceptualize entities independent of their features. It opens the possibility to track objects regardless of changes in their appearance.

4. A unique opportunity for the establishment of such a coordinate system can be seen in the acquisition of a binary substitution system of strictly inter-defined signs which completely differs in its usage rules from the sign system used for features (i.e. general terms/predicates). The usage of the local adverbs "here" and "there" as well as the demonstratives "this" and "that" are not bound to features but depend on the relative position of speakers in communication situations. In order to use these deictic terms correctly, one has to refer to the same feature by two binary, interdefined, and *different* terms from different positions in space. A child knows the meaning of "this" and "here" only if it realizes that from another position what is called "this" has to be called "that" and "here" has to be called "there."

5. The understanding of such usage rules leads to the omission of attention to features and establishes a primary level of identifiability in space and time. Realizing that two different signs can stand for each other leads to the conceptual construction of one referent for both. As the usage rules do not depend on features but on spatiotemporal position, their common referent cannot be a feature but has to be something else: a place or an object. Thereby, demonstratives become the first singular terms that children acquire, which is a crucial step in the development of propositional thought altogether. This binary substitution system forms the basis of the feature independent sign system as a whole.

6. The usage of demonstrative terms depends on spatiotemporal relations between speakers and what they talk about. Based on this, children develop a spatiotemporal coordinate system and establish a feature independent sign-based system to interpret featural data in a feature-independent way, namely, in terms of spatiotemporal position.

7. This allows children to discriminate features not only by their featural similarities but also by their spatio-temporal location. Features and combinations of features which can be discriminated by their places in time and space can be reidentified and appear as objects.

Against the background of these theoretical considerations, it appears that reference to objects and, thus, the individuation of objects itself is a highly complex matter closely related to the reference to and individuation of places in space and time (Mattos & Hinzen, 2015). Reference to features, on the other hand, is cognitively simpler. Following the principle that the attributed

cognitive capacities should not be more complex than required for an explanation of the observed phenomena, we should refrain from assuming from the outset that young children refer to objects. The tendency to interpret information generically is just what one would expect of a cognitive agent that is unable to refer to objects. Nevertheless, Csibra & Gergely (2009) interpret the "kind-bias" of young children, as well as adults, as an additional cognitive achievement and not as a lack of ability (cf. Csibra & Shamsudheen, 2015).

The tendency to take individuation of objects for granted might also stem from the assumption that it is empirically well demonstrated. But do the empirical data really support this assumption?

Objects first?

Researchers can retort to a large corpus of empirical studies in order to substantiate the claim that infants have the ability to identify objects in the first year. The standard interpretation of the findings is known as the object-first hypothesis (Xu & Carey, 1996). According to this hypothesis, children organize information in their visual field much like adults do, namely, in terms of space, object, and movement (Xu, 2007; Xu, Carey, & Quint, 2004). Therein, spatiotemporal information about location and motion is not only seen as the main information adults use to identify objects (Kahneman, Treisman, & Gibbs, 1992; Pylyshyn, 2001), but also as the basis on which infants learn to individuate objects (Spelke et al., 1995; Xu & Carey 1996; Wilcox & Baillargeon 1998). In accordance with the object-first hypothesis, the natural pedagogy account claims that young children are able to individuate objects outside the context of communicative actions (Yoon, Johnson & Csibra, 2008; Hernik, Sperber & Gergely, 2018). Only when receiving ostensively communicated information from a caregiver, children somehow suppress taking the information to refer to a particular object and interpret it generically.

In order to get a better understanding of the empirical findings and their interpretation as individuation of particular objects we need to have a look at how those studies operationalize children's ability to spatiotemporally individuate objects.

Spelke, Kestenbaum, Simons, Wein (1995) provide the classical looking-time violation-of-expectation paradigm that is seen as the basis of the object-first hypothesis. In their studies, 4-

month-old children watched an object appear from a central position in front of a stage. Two occluders were on the stage with some distance to each other. An object came out from behind the left occluder and moved to the left side of the stage and disappears behind the left occluder again. Then, an identical-looking object came out from behind the right occluder, moved to the right side of the stage, and returned behind the right occluder. Finally, the occluders were removed, showing the infants either an object behind each occluder (expected event) or an object only behind one occluder (unexpected event). 4-month-old children looked longer when there was only one object behind one of the occluders. These findings were replicated using similar looking-time procedures (e.g., Xu & Carey, 1996; Aguiar & Baillargeon, 1999; Wynn, 1992; among others).

Van de Walle, Carey, and Prevor (2000) established a violation-of-expectation manual-search procedure to investigate children's abilities to identify objects. In their studies, children were presented with a box and two objects analogous to the procedure of the looking-time studies by Xu & Carey (1996). In both studies, there was a first condition in which both featurally different objects were simultaneously visible after being taken out of a box by the experimenter before they were put back into the box. In addition, in the Van de Walle et al. (2000) study, there was a second condition in which the featurally different objects were not simultaneously visible. One object was taken out and put back by the experimenter before the second object was taken out from the box and put back again. In both conditions, infants were allowed to search the box in the test phase and either found both objects (expected event) or only found one object (unexpected event). The dependent measure was the duration of children's search efforts. 12-month-olds searched longer after the unexpected than after the expected event in both conditions, whereas 10-month-olds did so only in the first condition in which both objects were visible at the same time.

Yoon et al. (2008) and Hernik et al. (2018) used a classical looking-time violation-of-expectation procedure to investigate effects of verbal communication on object identification operationalised as a change-detection task. 9-month-old children saw an object which an adult either reached for (condition 1) or pointed to while communicating ostensively (condition 2). After a short occlusion, the object either changed its location or its features. Children looked

longer in the reaching-for condition when the object changed location after occlusion than when it changed features. In the pointed-to/ostensive communication-condition children looked longer when the object changed its features.

These findings can be interpreted as follows. As children appear surprised when presented with an unexpected number of objects (cf. Spelke et al., 1995; Xu et al., 1996; condition one in van de Walle et al., 2000), researchers conclude that they represent objects independently of their features. And due to the difference between the reaching-for and pointing-to conditions in Yoon et al. (2008) and Hernik et al. (2018), the tendency to interpret communicated information generically is taken to be due to ostensive communicative cues. Children behave exactly as would be expected from someone who can individuate objects.

However, there is a general concern with the standard interpretation of these findings. From the observation that children behave in line with our expectations and ontology, it cannot be concluded that they actually have the same ontology. The sensory environment can be perceived and interpreted in various meaningful ways which could all result in similar expectations (Hirsch 1997). This means that similar expectations about what will happen next or about what will be the result of an event can be constructed on the basis of entirely different ontologies.

The fact that young children are surprised or not surprised by the "same" events that we would be surprised by simply shows that, for example, this kind of discontinuity, interrelation, or featural outcome is what they expect or do not expect to be confronted with within their own ontologies. From showing that children have expectations comparable to ours, we learn about their behaviour but nothing about their ontology. This is what Hirsch (1997) points out as a fundamental difficulty of all experiments trying to show that children have a similar ontology by measuring the similarity of their expectations.

There is indeed an indefinite number of possible interpretations of what is going on in a cognitive system which do not assume that reference is made to objects in space and time. Hirsch (1997) suggests that the observed subjects might, for instance, be Quineans or Humeans or Strawsonians. Quineans, for instance, would perceive certain discontinuous space-time portions of reality (Quine 1960). Humeans would structure the world in momentary events and their

interrelations (Hume, 1739). While within a Strawsonian early child ontology the perceived environment would consist of placed features and feature changes (Strawson 1959). However, while similarities in expectations do not have to be the result of a shared ontology, differences in expectations can be considered as indications of different ontologies.

What are the consequences of this basic argument for interpreting the results of the experiments? Assuming that young children conceive the world in a similar way to Strawsonian feature placers, i.e., discriminate features and detect feature changes, in case of the experiments designed by Spelke et al. (1995) and in the respective first condition of van de Walle et al. (2000) and Xu et al. (1996) children would perceive certain feature changes in their visual field. Dissimilar visual features could, for example, interact in a predictable way and form a regular feature pattern. What we call "occlusion events" could simply be conceived as well-ordered changes of feature patterns by feature placers - corresponding to what can be taken to be likely according to their learning history. Other events, such as those presented in the test conditions, are statistically unlikely and thus surprising - in their ontology as well as in ours.

In the second condition of van de Walle et al. (2000) and Xu et al. (1996), in which the featurally different objects were not simultaneously visible, 10-month old children show behaviour that is different from what we would predict from our ontological perspective. This could serve as a hint for children's perceiving the world in terms of a different ontology. Apparently, given the initial visual featural situation, i.e. the perceived pattern of features, their expectations regarding the regularity of feature changes are different from what is observed. A partial change of colour and shape within a larger changing feature pattern does not seem to affect their expectation of the feature change outcome in a way we would expect from our object-perspective — given the complete feature pattern was not visible at the beginning of the experiment. A sequential experience of two feature patterns that differ in one feature (the feature we would attribute to the two objects, respectively) does not lead to the expectation of a complex feature pattern that would correspond to what we take to be two objects being taken out of a box.

Instead of sticking to an object-and-places interpretation of the children's ontology, then, it seems more reasonable to conclude that children operate with a different ontology. That 12-

month-olds again act as would be expected surely marks a developmental shift between 10- and 12-months. But again, this cannot serve as evidence for children already structuring the world according to our object-based ontology. The behaviour of 12-month-old children in the second condition of the experiments of van de Walle et al. (2000) and Xu et al. (1996) can be analysed just as their behaviour in the first condition by assuming something like the ability of feature-placing (see also Mendes, Rakoczy & Call, 2008). The shift between 10- and 12-month-olds might well consist in 12-month-olds having learned to detect smaller feature changes and to focus on smaller sections of features within a pattern in order to predict an outcome of a feature change.

Results of the experiment by Yoon et al. (2008) and Hernik et al. (2018) seem surprising compared to what we would expect, because in the pointed-at/ostensive communication condition (even if the position of the pointed-at object should have strong relevance (Hernik et al., 2018)) children seem to be invariably locked in a featural interpretation, while children in a reach-for situation first of all take into account what we would normally call "position of the object." That children do not react reference-based but use the information as pertaining to the object's features may therefore be also an indication of an ontological difference. The authors nonetheless interpret the results according to the "object-first-hypothesis": Children understand the reaching gestures as referring to the position of the object in the reaching-for condition, while they inhibit this ability in the pointed-at/ostensive-communication condition and interpret the pointing in combination with verbal ostensive cues as referring to a feature.

However, we argue that another interpretation of the results is more convincing. Let us, again, assume for illustration that young children conceive the world similar to Strawsonian feature-placers. For feature-placers, the world consists of features, not of objects in a spatiotemporal world. Pointing in combination with ostensive verbal cues by adults could be understood by feature-placing children as combined visual and auditory signs highlighting a feature. One could assume that the pointing gesture is understood to restrict the relevant featural area strongly, while a reaching gesture - due to the intention underlying the movement - serves in a well-ordered way as a hint to a feature-section as a part of a broader featural pattern. In this interpretation, both gestures would be interpreted by children in terms of features alone. What would differ would be the frame of featural reference — an interpretation that would well

converge with the just presented feature-placing interpretation of the results from van de Walle et al. (2000) and Xu et al. (1996). What is perceived as differences in location of an object from the perspective of our ontology can be detected by feature-placers as differences in feature patterns. A similar line of argument applies to what we would call differences in quantity.

By offering these alternative interpretations of the presented experiments we do not claim that children really perceive the world similar to Strawsonian feature-placers. But it shows that the object-first interpretation of these findings is not the only viable interpretation. And in light of the presented theoretical considerations it is a rather unlikely interpretation, because reference to objects is cognitively more demanding than the discrimination of features and their combinations. Following Hirsch (1997), what is needed to provide evidence for infants' ontologies are some principle-based expectations to the effect "that certain ontological schemes are more likely than others to support certain kinds of expectations."

Principle-based empirical expectations for an ontology without individuals

While we are criticizing standard interpretations of the studies discussed above for being inconclusive with regard to infants' actual ontological commitments, we do not propose that children's ontologies are fundamentally inaccessible. To the contrary, children who do not have the ability to refer to objects should differ from children who have already acquired the ability to refer to objects in the following ways.

Using history, function and possession as markers of object reference

Neither visual nor other featural information alone can be used to reidentify an object from an array of perceptually indistinguishable objects. Feature discrimination solely detects differences and similarities between features in different layouts. In the case of objects that are indistinguishable with regard to surface features, children who are not able to individuate objects at all should be unable to behave discriminatingly towards several such objects.

Also, children who do not have the ability to individuate objects are not able to reidentify objects — neither by perceptible features nor by their spatiotemporal position, function or other

relational properties (such as possession or valuation). For an external observer, participants' behaviour might look just like an actual reidentification of objects on the basis of featural information when, for example, featural differences and object boundaries overlap. However, a feature-independent notion of an object is required for being able to bind features to conceptualize an object as an individual. Proper featural reidentification presupposes non-featural reidentification. Otherwise only a co-occurrence of features would be detected. That a child is able to use featural information to reidentify what we take to be an object can at best be taken as inconclusive evidence for an ability to properly reidentify objects - i.e. without using featural information.

Because children have the ability to discriminate features by differences and similarities, effort should be geared to avoid confusion between, on the one hand, *identifying objects* by features or locations and, on the other, *detecting featural similarities*. In order to avoid this confusion, children's dealing with relational properties, function, and feature integration should be investigated — situations in which combinations of features and object boundaries do not overlap and in which properties of the broader featural layout cannot be used to pick out the target object. Arguably, relational properties, functions, and non-continuous combinations of features that need to be integrated over time can only be used to identify objects insofar as they are taken to somehow adhere to a "bearer" — which is an object. To illustrate: functions do not manifest in a scene's features when they are not currently performed. Likewise, relational properties need not surface in a situation's features. For instance, featural information would not be sufficient to pick out relational properties when the relational property — such as possession — holds to only some of a number of featurally indistinguishable objects. And in the case of drastic featural change — such as a toy rabbit that can be turned inside out to become a carrot — children would need a reference frame within which the dissimilar features could be integrated. If no additional featural information is available — such as the featural layout of the scene — an object would be required for such an integration of features.[1]

[1] Please note that object recognition proper could be more demanding than Rakoczy suggests and would actually involve an understanding aspectuality in our view.

During an experiment, such properties and relations could, for instance, be introduced by highlighting one of a number of featurally indistinguishable objects via some sort of special treatment. For instance, objects could be singled out by (1) their spatiotemporal origin, (2) the role they had in different play situations, (3) what happened to them during a shared interaction, (4) the child's own interactions with them — including active placing, (5) change of perceptual features without occlusion. Possession should be marked behaviourally and verbally. Thereby a common history is created with one of the objects, qualifying it non-featurally relative to a set of featurally indistinguishable objects. Within an array of such objects, the special object could only be singled out by its spatiotemporal position.

As a result, children without the ability to individuate objects should not be able to reidentify an object from an array of featurally very similar objects based on information obtained during such a common history with the object — even when the object was not occluded at any time. This should indeed be the case regardless of whether there was a change in the featural appearance of an object during the common history or not and regardless of whether it can be distinguished from other objects in the array on the basis of surface features. When no differences in infants' behaviour towards featurally indistinguishable but functionally, relationally, or otherwise historically different objects can be found, the simplest interpretation would be that those infants do not distinguish objects but only features.

Outlook

The main aim of this article was to highlight the possibility that the current consensus that infants possess a concept of individual objects very early on might be an overestimation of their cognitive abilities. We argued that empirical findings illustrating that infants can track objects in violation-of-expectation paradigms can be explained on the basis of feature-discrimination alone. Importantly, from an analytical perspective, individuation of objects is a complex and challenging skill that is closely related to the ability to refer to objects and implies an understanding of identity. Whether children are able to individuate objects as particulars with unique spatiotemporal coordinates can be probed, for example, by testing whether infants are able to reidentify an object in an array of featurally indiscriminable objects.

Furthermore, we highlighted that the ability to individuate objects should be explored in contexts of joint attention because of the close connection between reference and object individuation. The "referential side" of triangulation (child-object) — its mechanics and impacts on individuation skills — could be investigated by focusing on the ability to refer to objects within triangulation situations. Here, we expect systematic differences in the ways in which children interpret pointing gestures before and after they have mastered object individuation. When pointing gestures emerge around the first birthday, they serve to localize entities and events in intersubjective space and allow for a coordination of attention. We would predict that infants would use and interpret such acts of reference as merely feature-highlighting. Only with the mastery of a feature-independent system of signification, children would become able to individuate and refer to objects. This, in turn, puts new opportunities for the interpretation and production of pointing gestures in their hands. In addition to opening new modes of reference, a fully developed understanding of deictic demonstratives and local adverbs also function as the first singular terms children acquire which is crucial for the development of propositional thought (Tugendhat, 1976).

We expect a strong connection between the development of the ability to individuate and refer to objects and working memory. From an information-processing perspective, one can state that the capability to individuate objects leads to a "superior effectiveness" in information processing (Evans 1982, p. 277). The central aspect of the specific human system of gathering information is "…that we group information together, as being from the same object — that we collect information lines into bundles" (p. 126). An increase in working memory should therefore develop in parallel to the development of the ability to refer to objects independently of their features.

References

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive psychology*, 39 (2), 116-157.

Augustinus, A. (2009). *Confessiones. Bekenntnisse. Lateinisch und Deutsch.* Ditzingen: Reclam-Verlag.

Baldwin, D.A., Moses, L.J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, 10, 309–329.

Cacchione, T. & Rakoczy, H. (2017): Comparative metaphysics: Thinking about objects in space and time. In: *Handbook of Comparative Psychology*, ed. by J. Call (Ed.), pp. 579-599, American Psychological Association.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Science*, 13, 148–153.

Csibra, G. & Shamsudheen (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual review of psychology*, 66, 689-710.

Donnellan, K. S. (1966). Reference and definite descriptions. *The philosophical review*, 75(3), 281-304.

Evans, G. (1982). *The varieties of reference*. Oxford: Clarendon Press.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100(1), 25-50.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, 20(3), 347-353.

Hernik, M., Sperber, D., & Gergely, G. (2018). Infants' encoding of location and appearance of pointed-at and reached-for objects when both location and appearance are relevant. Poster presented at the "Budapest CEU Conference on Cognitive Development 2018", Budapest, Hungary.

Hespos, S. J., & Rochat, P. (1997). Dynamic mental representation in infancy 1. *Cognition*, 64 (2), 153-188.

Hirsch, E. (1997). Basic objects: A reply to Xu. *Mind & Language*, 12(3-4), 406-412.

Hume (1739). 1978. *A Treatise on Human Nature*. Oxford: Oxford University Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive psychology*, 24(2), 175-219.

Kripke, S. A. (1980). *Naming and Necessity*. Harvard: Harvard University Press.

Mattos, O., Hinzen, W. (2015). The linguistic roots of natural pedagogy. *Frontiers in Psychology*. doi: 10.3389/fp-syg.2015.01424

Mendes, N., Rakoczy, H., & Call, J. (2008). Ape metaphysics: Object individuation without language. *Cognition*, 106(2), 730-749.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1-2), 127-158.

Quine, W.V.O. (1960) *Word and object*. Cambridge, MA: MIT Press.

Russell, B. (1905). On denoting. Mind, 14(56), 479-493.

Searle, J. R. (1958). *Proper names*. Mind, 67(266), 166-173.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, 13(2), 113-142.

Strawson, P. F. (1959). *Individuals*. London: Methuen.

Tugendhat, E. (1976). *Einführung in die sprachanalytische Philosophie*. Frankfurt am Main: Suhrkamp Verlag.

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental science*, 10(1), 121-125.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child development*, 78(3), 705-722.

Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development*, 1(3), 249-280.

Wilcox, T., & Baillargeon, R. (1998). Object individuation in young infants: Further evidence with an event-monitoring paradigm. *Developmental Science*, 1(1), 127-142.

Wittgenstein, L., von Wright, G. H., Paul, D., & Anscombe, G. E. M. (1969). *On certainty* (Vol. 174). Oxford: Blackwell.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358 (6389), 749.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive psychology*, 30(2), 111-153.

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, 49(2), 155-190.

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, 16(5), 372-377.

Yoon, J. M., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, 105(36), 13690-13695.