



The Role of Saliency in Learning First Words

Eugenia Wildt*, Katharina J. Rohlfing and Ingrid Scharlau

Faculty of Arts and Humanities, Paderborn University, Paderborn, Germany

In word learning, one key accomplishment is the reference, that is, the linking of a word to its referent. According to classical theories, the term *reference* captures a mental event: A person uses a word to mentally recall a concept of an entity (an object or event) in order to bring it into the mental focus of an interaction. The developmental literature proposes different approaches regarding how children accomplish this link. Although researchers agree that multiple processes (within and across phonological, lexical, and semantic areas) are responsible for word learning, recent research has highlighted the role of saliency and perception as crucial factors in the early phases of word learning. Generally speaking, whereas some approaches to solving the reference problem attribute a greater role to the referent's properties being salient, others emphasize the social context that is needed to select the appropriate referent. In this review, we aim to systematize terminology and propose that the reason why assessments of the impact of saliency on word learning are controversial is that definitions of the term *saliency* reveal different weightings of the importance that either perceptual or social stimuli have for the learning process. We propose that defining early word learning in terms of paying attention to salient stimuli is too narrow. Instead, we emphasize that a new link between a word and its referent will succeed if a stimulus is *relevant* for the child.

Keywords: word learning, saliency, relevance, joint attention, joint action

OPEN ACCESS

Edited by:

Emily Mather,
University of Hull, United Kingdom

Reviewed by:

Carmel Houston-Price,
University of Reading,
United Kingdom
Yang Zhang,
University of Minnesota Twin Cities,
United States

*Correspondence:

Eugenia Wildt
eugenia.wildt@uni-paderborn.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 28 November 2018

Accepted: 01 May 2019

Published: 15 May 2019

Citation:

Wildt E, Rohlfing KJ and
Scharlau I (2019) The Role of Saliency
in Learning First Words.
Front. Psychol. 10:1150.
doi: 10.3389/fpsyg.2019.01150

INTRODUCTION

Studies vary in their suggestions regarding how links emerge between words and referents: Some explanations build on research in the field of the psychology of perception (see **Table 1**) confirming that saliency (or salience) can capture attention quite effectively. Bottom-up saliency (Latin *salire*: to jump) is by definition a property in objects that makes them stand (or jump) out of the surrounding context. For instance, a red jacket stands out in the context of a crowd of black jackets (Itti and Koch, 2001: 194). Besides visual features such as color, luminance, orientation, motion, or size, auditory properties such as loudness, pitch, or spectral shape can attract attention in a mainly bottom-up way (e.g., Treue, 2003). Whereas approaches taking advantage of saliency propose that during word learning, salient properties capture and sustain infants' attention for an object during the time in which the referent is being labeled, other explanations emphasize social interaction and its goals and how reference is established for this among partners. Our aim is to systematize the use of terms and sensitize the reader to the difference between *saliency* and *relevance*. We argue that relevance is achieved by embedding the child's perspective into a social environment—that is, a history of joint actions. In the following, we first review studies that focus on the role of saliency for word learning.

MECHANISTIC VIEW ON WORD LEARNING

The mechanistic view (e.g., Smith, 1995; Smith et al., 1996; Samuelson and Smith, 1998; Spencer et al., 2011; see also Plunkett, 1997, for a review) proposes that general mechanisms such as memory and attention drive the word learning process. The child's task is to link the heard phonological form with an entity (an object or event) that he or she is attending to visually. Supporting studies have boiled word learning down to essentially two problems: First, the learning environment is ambiguous [a problem already identified by Quine (1960)], because it offers many potential referents and distractors (Trueswell et al., 2016). Second, even if infants identify the referent correctly, they have difficulties in sustaining their visual focus on the referent because of their still maturing attentional skills (Yu and Smith, 2016). Thus, the mechanistic view centers on the question how infants solve the ambiguity problem and sustain their attention on a referent in order to successfully establish a new word-object link.

One approach considers saliency to be a precondition for recruiting infant's attention for an object, which in turn, if temporally synchronized with labeling, will establish a link between a word and its referent in the infant's memory (Gogate and Bahrick, 1998). We shall collectively name these studies *associationist* (Hollich et al., 2000: 12), although they operationalize saliency differently (see below). A related position assigns a greater role to child's growing experience by postulating *constraints* and *principles* of learning and is presented in more detail below.

SALIENCY—A PROPERTY IN OBJECTS

The associationist account postulates that infants solve the ambiguity problem because of their preference for salient objects or because they assume that adults will label that object which is the most interesting from the infant's point of view. This account implies that for early word learning, infants rely more on perceptual saliency than on social stimuli (Moore et al., 1999; Hollich et al., 2000). This was demonstrated by Pruden et al. (2006) who operationalized social stimuli as the gaze of the experimenter and compared its effect with that of perceptual saliency: In a "coincidental condition" (p. 269), the experimenter gazed at and simultaneously named a salient object; in a "conflict condition," the experimenter gazed at and labeled a "boring" object while a salient object was present. Infants in the coincidental condition spent more time looking toward the salient object, indicating that they mapped the new word onto the intended referent. The conflict group also looked longer at the salient distractor, indicating that the label was mismatched with the salient object. The authors viewed their results as evidence that young infants weight object saliency higher than social cues in their word learning process. Only in the course of development, do infants "move from learning words associatively to learning words based on the social cues a speaker emits" (p. 278). However, Golinkoff and Hirsh-Pasek (2006) noted that it

is still not known how child development reveals a qualitative change toward weighting social cues more than perceptual cues.

Crucially, even though Pruden et al. (2006) did not define the term *saliency* explicitly in their study, they did vary it methodologically in terms of attention-grabbing properties—that is, in terms of objects that recruit and hold infants' visual attention: The salient stimuli were brightly colored and could either make a noise or move, and they were paired with boring objects (dull color, neither motion nor noise). These properties are consistent with the psychological view on bottom-up saliency.¹ Other studies extended this notion: For example, an object's saliency increases if it is larger (Smith et al., 1996; Pereira et al., 2014) or more centered in the infant's visual field in comparison to other toys, rotates on a turntable (Moore et al., 1999), moves (Werker et al., 1998; Houston-Price et al., 2005, 2006), or is illuminated (Axelsson et al., 2012). It should be noted that even though infants attend to a moving object during training, they exclude it as referent at test if its movement is not consistent (Houston-Price et al., 2006). In this way, saliency is regarded as a "bottom-up sensory input that is clean" (Yu and Smith, 2012: 258), meaning that only one object dominates the visual view. In natural environments, parents can facilitate their infants' word learning if they establish such visually optimal moments by bringing the target object more to the fore. Furthermore, infants often create situations on their own in which referential ambiguity is low (e.g., during toy play by exploring one object at a time). Object naming in these moments is associated positively with word learning (Yu and Smith, 2012; Pereira et al., 2014).

With regard to the mechanistic view, results indicate that word learning is driven by general processes of attention and memory that are recruited via attention-grabbing features in objects. This saliency effect can be explained by cognitive learning mechanisms being facilitated by the diminishment of competitors and the unambiguous determination of the referent (Axelsson et al., 2012).

SALIENCY—A PROPERTY GENERATED BY THE PERCEIVER

Without opposing the role of perception, another perspective highlights infants' experience with word learning episodes. Being exposed to referents and their labels, the cognitive demand, namely to map the label onto some features of the referent, gives rise to necessary cognitive operations. Accordingly, children make use of *constraints* and *principles* (Markman, 1994) that narrow down referent selection. From this position, the salient property in objects derives from the knowledge (and experience with the labeling task) of the perceiver. Such constraints and principles as the *whole-object assumption* (Markman and Wachtel, 1988; Woodward, 1992) or *mutual exclusivity* (Markman and Hutchinson, 1984; Clark, 1987; Markman, 1989;

¹Note that the notion of pure bottom-up capture that depends only on stimulus features and not on current goals is a contested issue in experimental psychology; the conditions under which irrelevant salient stimuli capture attention are still not fully understood.

TABLE 1 | State of research on saliency in early infancy.

Topic of research on salient properties	Results	References
Patterned vs. Monochromatic stimuli	5-day-olds prefer black-and-white patterns over monochromatic surfaces	Fantz, 1963
Development of color preference: Chromatic vs. Achromatic	Newborns and 3-month-olds prefer chromatic over achromatic stimuli; 3-month-olds prefer red and yellow over blue and green	Adams, 1987
Dynamic vs. Static stimuli	5-month-olds prefer moving over static female faces	Wilcox and Clayton, 1968
Dynamic objects in combination with chromatic stimuli	14-week-olds' attention to moving stimuli depends on color of target object: When visual field is a mixture of red and green stimuli, infants have difficulties in attending to a green target object that starts to move; when a red object starts to move, infants readily switch their attention to it	Nagata and Dannemiller, 1996
Change of color vs. Change of rotation speed	9-month-olds notice a salient difference in color change, but not in a change of rotation speed	Kaldy and Blaser, 2013
Development of recognizing changes in salient features	6-month-olds require larger differences in color, shape, or luminance to identify an object change compared to 9-month-olds	Kaldy and Blaser, 2009

TABLE 2 | State of research on social sensitivity in infancy.

Topic of research	Results	References
Visual preferences	Newborns prefer to look at face-like compared to non-face-like stimuli; preference for faces declines during second month of life Infants discriminate their mother's faces from foreign ones	Fantz, 1963; Johnson et al., 1991 Field et al., 1984; Bushnell et al., 1989; Pascalis et al., 1995
Listening preferences	Infants prefer listening to human speech compared to rhesus calls, as well as speech compared to non-speech sounds; this listening bias for human speech is innate From birth, infants prefer listening to infant-directed compared to adult-directed speech	Vouloumanos and Werker, 2004; Vouloumanos and Werker, 2007; Vouloumanos et al., 2010 Fernald, 1985; Cooper and Aslin, 1990; Pegg et al., 1992
Following others' gaze direction	Infants prefer human faces and are sensitive to direct eye contact from birth 6- to 18 month-olds follow others' gaze direction if the referent is in their visual field Three mechanisms serve joint attention in the first 18 months Infants do not reliably follow others' gaze direction for joint visual attention before the age of 18 months Differences in the capacity of gaze following at 6 months of age relate to vocabulary development	Farroni et al., 2002 Butterworth and Cochran, 1980 Butterworth and Jarrett, 1991 Moore and Corkum, 1998 Morales et al., 1998
Responses to others' gestures	12-month-olds but not 9-month-olds follow gestures to targets behind them 4.5-month-olds follow dynamic, but not static pointing gestures	Deák et al., 2000; Flom et al., 2004 Rohlfing et al., 2012
Responses to aspects in social interactions	2-month-olds show organized facial expressions that are responsive to maternal communication Infants are more responsive to mothers in live compared to replayed videotape sequences Right from early on, infants are socialized as participants in interactions Early vocal exchange between infant and mother has a turn-taking format Infants are sensitive to other's contingent actions	Trevarthen, 1985 Murray and Trevarthen, 1986 de Leoin, 2000 Masataka, 2003; Gratier et al., 2015 Kaye, 1977; Jaffe et al., 2001; Striano et al., 2005
Understanding social events	14-month-olds show understanding of social intentions; they imitate intentional actions more than accidental actions	Carpenter et al., 1998

Merriman et al., 1989; Waxman and Kosowski, 1990; Golinkoff et al., 1992, 1994) are of different natures. While the whole-object assumption guides infants to the object's features that they need to select for labeling, *mutual exclusivity* describes a bias in infants that prevents them from linking new words with already named objects because of the underlying assumption that objects can bear only one label (Markman, 1989, 1990). Indications of mutual exclusivity have been observed in infants as young as 10 months (Mather and Plunkett, 2010), but the cognitive basis of such a bias remains disputed: Is mutual exclusivity based on the knowledge of familiar versus novel labels for objects (e.g., Mervis and Bertrand, 1994) or rather the knowledge of an object's novelty (e.g., Merriman et al., 1995; Mather and Plunkett, 2010; Horst et al., 2011)? At this point, we would like to note that we limit our considerations here to studies investigating the novelty bias.

To determine this bias, Mather and Plunkett (2012) presented 22-month-old infants with two stimuli: Both were name-unknown, but only one of them was truly novel. The children had become familiar with the other object through pre-exposure to it. Results showed that after having heard a new word, infants' attention, in the form of looking time, increased more toward the novel object compared to the familiar one. The authors concluded that mutual exclusivity is a novelty-based mechanism, because it seems to be cognitively easier for infants to search for a perceptually novel object in the environment rather than retrieving all familiar object names. Hence, when presented with a novel word, the most novel object will appear to be the most salient one to an infant, thereby facilitating the mapping process. The novelty bias is a good example for a learning constraint, although we do not suggest that all constraints and principles are attributable to saliency in the same way as novelty.

One aspect is crucial to this position: Rather than being a salient object emerging through bottom-up attention, saliency is generated by the perceiver through top-down processes (Connor et al., 2004). When facing the ambiguity problem, infants rely on their prior knowledge to identify whether not only words but also objects are actually novel. The ability to hear a new label and to map it onto a novel, name-unknown object, characterizes infants as active learners in their environment. This perspective holds that word learning is a cognitive process that also affords a top-down mechanism (including past memory) in the perceiver.

SOCIAL-PRAGMATIC AND INTERACTIONIST VIEW ON WORD LEARNING

Whereas in the mechanistic view, word learning is dependent on perceptual and attentional constraints, social-pragmatic theories (e.g., Bruner, 1983; Tomasello, 2000; Csibra and Gergely, 2006) claim that from early on, infants are sensitive to social cues (see Table 2).

Even though both social-pragmatic and interactionist perspectives agree that word learning is an inherently social process, they differ in how referents become salient during social interaction. We shall elaborate on this difference in the following.

SALIENCY—A PROPERTY EMERGING FROM SOCIAL PERCEPTION

One line of social-pragmatic studies (Tomasello and Akhtar, 1995; Baldwin et al., 1996; Carpenter et al., 1998) proposes that

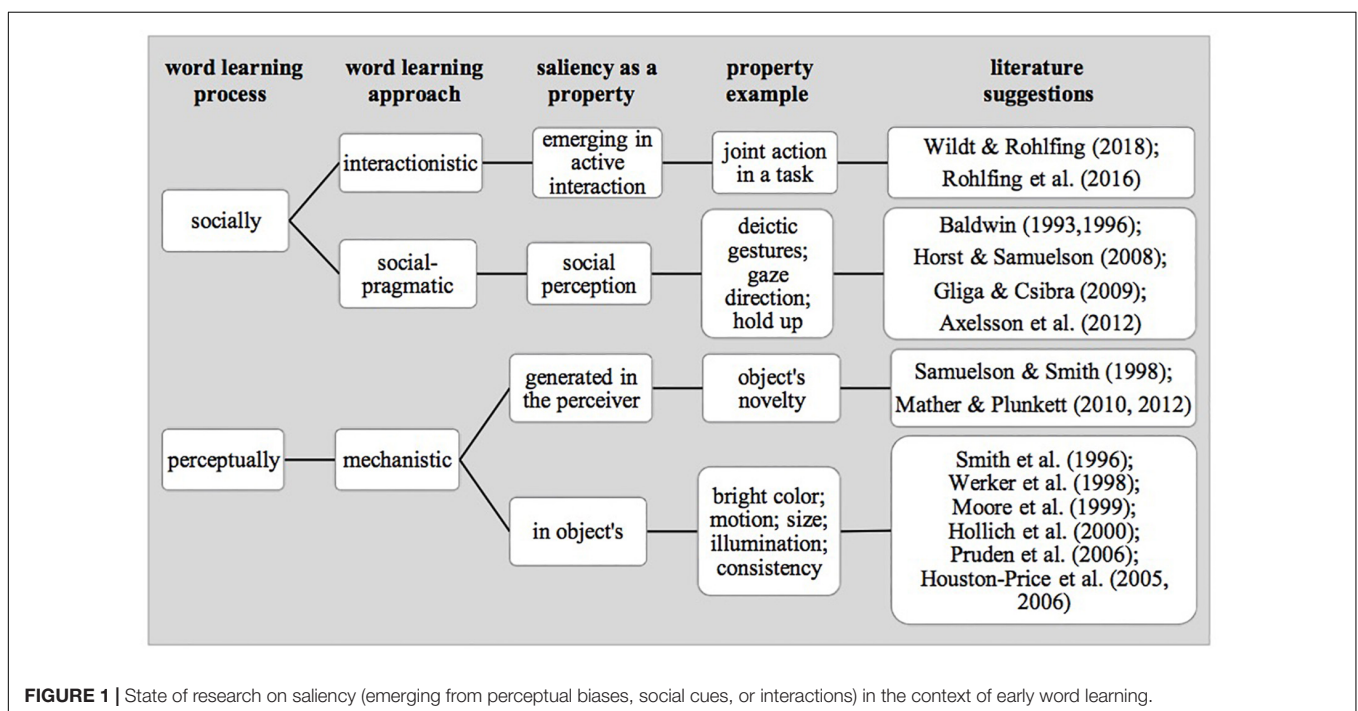


FIGURE 1 | State of research on saliency (emerging from perceptual biases, social cues, or interactions) in the context of early word learning.

because infants are especially sensitive to objects that adults single out via *ostensive means* (eye gaze, gestures, or emotions), these objects become salient and can then be linked with new words (Akhtar et al., 1996; Csibra, 2010; Axelsson et al., 2012). In a study with 24-month-olds, Horst and Samuelson (2008) used ostensive naming (i.e., addressing the child directly, holding up the target object, and pointing at it) as a form of specific social behavior that singled out the target object and reduced competition from distractors. Compared to a condition in which a non-ostensive naming was provided, retention of new words and thus long-term learning was observed only in the ostensive naming condition. It can therefore be concluded that the use of social cues not only facilitates infants' encoding but, more crucially, induces their long-term processing.

Recent research on infants' reactions to attention-directing social cues supports the assumption that caregivers' actions do not just facilitate shared attention, but might also aid learning. For example, Deák et al. (2018) found that five properties are salient from a perceiver's perspective: gaze shift, pointing gestures, speech, object sounds, and object manipulation. They argued, for example, that pointing gestures are salient due to the sweeping motion of the arm and hand (see also Rohlfing et al., 2012). Their results indicated that when used within a dyadic toy play interaction, 3- to 11-month-olds were sensitive to all five different caregiver social cues, with object manipulation being the most effective (see also Yu and Smith, 2012).

Effects of saliency as a property of social perception can be explained by natural pedagogy (Csibra and Gergely, 2009). This stipulates that from early on (and possibly even from birth, Csibra, 2010), infants are sensitive to ostensive signals. Because they understand the referential nature of communicative signals (Gliga and Csibra, 2009), they follow them to indicate the word reference (Baldwin, 1993; Baldwin et al., 1996).

In sum, this line of social-pragmatic research reveals that social cues evoke a high level of attention in the perceivers and guide them toward objects relevant for the interaction. Furthermore, research on the attention of adults makes a similar distinction between top-down attention, which is guided by goals, knowledge, or expectations; and bottom-up attention, driven by stimulus contrast or salience (e.g., Folk et al., 1992). This might point to a more general mechanism that drives word learning as well as visual attention.

AWAY FROM SALIENCY—RELEVANCE EMERGING IN INTERACTIONS

In comparison to the line of research mentioned above, interactionist theory turns away from the view that infants build references by observing salient entities. Instead, a successful object-referent association is attributed to infants' *engagement* in an interaction toward a joint goal. Proponents of this approach emphasize the importance of a pragmatic frame that is established by repetitive participation in an ongoing social event leading to interactional experience and particular communicative acts (in the form of joint attention) that serve word learning (Bruner, 1983; Rohlfing et al., 2016).

Along these lines, Wildt and Rohlfing (2018) investigated word learning in 10-month-olds by directly comparing the associationist, the social-pragmatic, and the interactionist approaches. A stimulus set consisting of one perceptually highly salient and one less salient object was presented on a screen within an intermodal preferential looking paradigm (IPLP). Infants' visual attention was estimated by an eye tracker. Infants' interest for the stimulus set was assessed as the baseline. This confirmed that all infants preferred the salient stimulus. After that, all infants participated in an interaction: While both stimuli lay side by side on a table, the experimenter gazed at the "boring" stimulus, repeatedly demonstrated its function, and labeled it ostensively. After ostensive naming, infants were divided into two groups and encouraged to play with the less salient object. One group explored the object's function on their own without any further input from the experimenter; the other group explored it in an interaction. The specific manipulation in the second group was for the experimenter to provide the infants with support in manipulating the boring object to achieve a joint goal. Here, we were contrasting the impact of joint attention with the impact of joint action on word learning. Our prediction was that infants would match a new word to the boring object because it was demonstrated as relevant for the interaction. Even though the salient object was always in the infants' field of view, subsequent visual attention provided evidence that infants' engagement was crucial to establish the reference. These results contradict the associationist approach and demonstrate that young infants' early attention is not driven merely by perceptually salient objects. In addition, the results on the contrast between the joint action and the joint attention condition reveal that establishing joint attention is not sufficient to develop a word-referent link for a boring object, because infants in this condition still preferred the salient object as at baseline. In contrast, an interaction in which infants participated had a stronger effect on infants' perception, because they no longer showed a preference for the salient object. In this vein, while the contribution of pointing behavior to later language development is recognized, recent studies reveal that this relationship might rely on pointing to interaction-relevant entities (Białek et al., 2018). Again, we want to point out a possible parallel to theories of attention in adults that are as yet unexplored: Selection-for-action approaches assume that attentional selection mainly serves action control (e.g., Neumann, 1987). Even though there is no particular reference to interactions in these approaches, interactions and actions may share the crucial property that action capacities are limited. Further research needs to clarify whether saliency emerging from interactions accords with theoretical approaches on selection for action in adults.

Taken together, the interactionist approach differs in its perspective on word learning by claiming that infants need not only perceptual or social saliency but also an active experience in an interaction such as achieving a relevant function or a joint goal. Hence, a referent becomes relevant by "charging" it with a rich meaning from actions that have taken place within a social interaction (Nomikou et al., 2016) and that draw on what the child knows (Bloom et al., 1993) rather than merely being salient in a joint attention or visual attention scenario.

DISCUSSION

We surveyed how the term *saliency* is used in three theoretical approaches to the study of word learning. Systematizing its uses, we registered that they differ in terms of the prioritization of the non-social versus the social information that is recruited as a cue in the process of word learning (see **Figure 1**).

The difference in mechanistic approaches lies in the internalization of saliency: Whereas from the associationist view, bottom-up mechanisms drive infants' attention toward attention-grabbing properties, the approach using constraints and principles suggests that a perceiver's top-down driven attention is based on past knowledge. The recruitment of the experience in perceiving an object is what makes it salient; or to put it in more appropriate terms: relevant.

This review addressed two other approaches that consider saliency to be socially driven. This means that infants are sensitive to social cues from early on, and that word learning is inherently social. Whereas in one line of social-pragmatic studies, saliency is attributed to infants' social perception and their responsivity to cues such as eye gaze and pointing or ostensive labeling, a second interactionist line of studies claims that the pragmatic frame of a joint action is needed for the infant to recognize the relevance

of an object for the joint goal of this interaction. In this latter view, a word-object link becomes charged with interwoven words and actions. Therefore, infants must be additionally embedded in joint actions and achieve a joint goal to see the purpose of a word and to involve their memory processes.

Taken together, it becomes clear that a unified theory is lacking. As a first step, we propose that future research should focus more on the term *relevance* rather than *saliency*, because it better encompasses the (social) context in which infants attend to objects.

AUTHOR CONTRIBUTIONS

EW and KR developed the framework and wrote the mini-review. This paper benefits from the expertise of IS in visual saliency for information processing.

FUNDING

This work was made possible within Beethoven project EASE funded by the NCN-DFG collaboration (RO 2443/5-1).

REFERENCES

- Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behav. Dev.* 10, 143–150. doi: 10.1016/0163-6383(87)90029-4
- Akhtar, N., Carpenter, M., and Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Dev.* 67, 635–645. doi: 10.1111/j.1467-8624.1996.tb01756.x
- Axelsson, E. L., Churchley, K., and Horst, J. S. (2012). The right time at the right time: why ostensive naming facilitates word learning. *Front. Psychol.* 3:88. doi: 10.3389/fpsyg.2012.00088
- Baldwin, D. A. (1993). Early referential understanding: infants' ability to recognize referential acts for what they are. *Dev. Psychol.* 29, 832–843. doi: 10.1037//0012-1649.29.5.832
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., and Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Dev.* 67, 3135–3153. doi: 10.1111/j.1467-8624.1996.tb01906.x
- Białek, A., Białecka-Pikul, M., Filip, A., and Broda, M. (2018). Relevance matters: eighteen month-olds' use of relevant informative pointing as a predictor of two-year-olds' language abilities/La relevancia es importante. El uso de gestos deícticos relevantes e informativos por parte de niños de año y medio como factor predictor de las capacidades lingüísticas a los dos años. *Infanc. Aprendizaje* 41, 1–28.
- Bloom, L., Tinker, E., and Margulis, C. (1993). The words children learn: evidence against a noun bias in early vocabularies. *Cogn. Dev.* 8, 431–450. doi: 10.1016/s0885-2014(05)80003-6
- Bruner, J. S. (1983). *Child's Talk: Learning to use Language*. New York, NY: W. W. Norton & Company.
- Bushnell, I. W. R., Sai, F., and Mullin, J. T. (1989). Neonatal recognition of the mother's face. *Br. J. Dev. Psychol.* 7, 3–15. doi: 10.1111/j.2044-835x.1989.tb00784.x
- Butterworth, G., and Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *Int. J. Behav. Dev.* 3, 253–272. doi: 10.1177/016502548000300303
- Butterworth, G., and Jarrett, N. (1991). What minds have in common is space: spatial mechanisms serving joint visual attention in infancy. *Br. J. Dev. Psychol.* 9, 55–72. doi: 10.1111/j.2044-835x.1991.tb00862.x
- Carpenter, M., Akhtar, N., and Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behav. Dev.* 21, 315–330. doi: 10.1016/s0163-6383(98)90009-1
- Clark, E. V. (1987). "The principle of contrast: a constraint on language acquisition," in *Proceedings of the 20th Annual Carnegie Symposium on Cognition*, ed. B. MacWhinney (Hillsdale, NJ: Erlbaum).
- Connor, C. E., Egeth, H. E., and Yantis, S. (2004). Visual attention: bottom-up versus top-down. *Curr. Biol.* 14, R850–R852.
- Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595. doi: 10.1111/j.1467-8624.1990.tb02885.x
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind Lang.* 25, 141–168. doi: 10.1111/j.1468-0017.2009.01384.x
- Csibra, G., and Gergely, G. (2006). Social learning and social cognition: the case for pedagogy. Processes of change in brain and cognitive development. *Attent. Perform.* XXI 21, 249–274.
- Csibra, G., and Gergely, G. (2009). Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153. doi: 10.1016/j.tics.2009.01.005
- de León, L. (2000). The emergent participant: interactive patterns in the socialization of Tzotzil (Mayan) infants. *J. Linguist. Anthropol.* 8, 131–161. doi: 10.1525/jlin.1998.8.2.131
- Deák, G. O., Flom, R. A., and Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Dev. Psychol.* 36, 511–523. doi: 10.1037/0012-1649.36.4.511
- Deák, G. O., Krasno, A. M., Jasso, H., and Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses during object play. *Infancy* 23, 4–28. doi: 10.1111/inf.12204
- Fantz, R. L. (1963). Pattern vision in newborn infants. *Science* 140, 296–297. doi: 10.1126/science.140.3564.296
- Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9602–9605. doi: 10.1073/pnas.152159999
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–195. doi: 10.1016/s0163-6383(85)80005-9
- Field, T. M., Cohen, D., Garcia, R., and Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behav. Dev.* 7, 19–25. doi: 10.1016/s0163-6383(84)80019-3

- Flom, R., Deák, G. O., Phill, C. G., and Pick, A. D. (2004). Nine-month-olds' shared visual attention as a function of gesture and object location. *Infant Behav. Dev.* 27, 181–194. doi: 10.1016/s0163-6383(04)00017-7
- Folk, C. L., Remington, R. W., and Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 1030–1044. doi: 10.1037//0096-1523.18.4.1030
- Gliga, T., and Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychol. Sci.* 20, 347–353. doi: 10.1111/j.1467-9280.2009.02295.x
- Gogate, L. J., and Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *J. Exp. Child Psychol.* 69, 133–149. doi: 10.1006/jecp.1998.2438
- Golinkoff, R. M., and Hirsh-Pasek, K. (2006). Baby wordsmith: from associationist to social sophisticate. *Curr. Dir. Psychol. Sci.* 15, 30–33. doi: 10.1111/j.0963-7214.2006.00401.x
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., and Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Dev. Psychol.* 28, 99–108. doi: 10.1037//0012-1649.28.1.99
- Golinkoff, R. M., Mervis, C. B., and Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework. *J. Child Lang.* 21, 125–155.
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., and Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Front. Psychol.* 6:1167. doi: 10.3389/fpsyg.2015.01167
- Hollich, G., Hirsh-Pasek, K., and Golinkoff, R. M. (2000). I. What does it take to learn a word? *Monogr. Soc. Res. Child Dev.* 65, 1–16. doi: 10.1002/wcs.1421
- Horst, J. S., and Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy* 13, 128–157. doi: 10.1080/15250000701795598
- Horst, J. S., Samuelson, L. K., Kucker, S. C., and McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition* 118, 234–244. doi: 10.1016/j.cognition.2010.10.015
- Houston-Price, C., Plunkett, K., and Duffy, H. (2006). The use of social and saliency cues in early word learning. *J. Exp. Child Psychol.* 95, 27–55. doi: 10.1016/j.jecp.2006.03.006
- Houston-Price, C., Plunkett, K., and Harris, P. (2005). 'Word-learning wizardry' at 1;6. *J. Child Lang.* 32, 175–189. doi: 10.1017/s0305000904006610
- Itti, L., and Koch, C. (2001). Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., et al. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monogr. Soc. Res. Child Dev.* 66, i–viii, 1–149.
- Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40, 1–19. doi: 10.1016/0010-0277(91)90045-6
- Kaldy, Z., and Blaser, E. (2009). How to compare apples and oranges: infants' object identification tested with equally salient shape, luminance, and color changes. *Infancy* 14, 222–243. doi: 10.1080/15250000802707088
- Kaldy, Z., and Blaser, E. (2013). Red to green or fast to slow? Infants' visual working memory for "just salient differences." *Child Dev.* 84, 1855–1862. doi: 10.1111/cdev.12086
- Kaye, K. (1977). "Toward the origin of dialogue," in *Studies in Mother-Infant Interaction*, ed. H. R. Schaffer (London: Academic Press), 89–119.
- Markman, E. M. (1989). *Categorization and Naming in Children*. Cambridge, MA: MIT Press.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cogn. Sci.* 14, 57–77. doi: 10.1016/0364-0213(90)90026-s
- Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua* 92, 199–227. doi: 10.1016/0024-3841(94)90342-5
- Markman, E. M., and Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cogn. Psychol.* 16, 1–27. doi: 10.1016/0010-0285(84)90002-1
- Markman, E. M., and Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cogn. Psychol.* 20, 121–157. doi: 10.1016/0010-0285(88)90017-5
- Masataka, N. (2003). *The Onset of Language*. Cambridge: Cambridge University Press.
- Mather, E., and Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *J. Exp. Child Psychol.* 105, 232–242. doi: 10.1016/j.jecp.2009.11.004
- Mather, E., and Plunkett, K. (2012). The role of novelty in early word learning. *Cogn. Sci.* 36, 1157–1177. doi: 10.1111/j.1551-6709.2012.01239.x
- Merriman, W. E., Bowman, L. L., and MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monogr. Soc. Res. Child Dev.* 54, 1–132.
- Merriman, W. E., Marazita, J., and Jarvis, L. (1995). "Children's disposition to map new words onto new referents," in *Beyond Names for Things: Young Children's Acquisition of Verbs*, eds M. Tomasello and W. E. Merriman (Hillsdale, NJ: Erlbaum), 147–183.
- Mervis, C. B., and Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Dev.* 65, 1646–1662. doi: 10.1111/j.1467-8624.1994.tb00840.x
- Moore, C., Angelopoulos, M., and Bennett, P. (1999). Word learning in the context of referential and saliency cues. *Dev. Psychol.* 35, 60–68. doi: 10.1037//0012-1649.35.1.60
- Moore, C., and Corkum, V. (1998). Infant gaze following based on eye direction. *Br. J. Dev. Psychol.* 16, 495–503. doi: 10.1111/j.2044-835x.1998.tb00767.x
- Morales, M., Mundy, P., and Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behav. Dev.* 21, 373–377. doi: 10.1016/s0163-6383(98)90014-5
- Murray, L., and Trevarthen, C. (1986). The infant's role in mother-infant communications. *J. Child Lang.* 13, 15–29.
- Nagata, Y., and Dannemiller, J. L. (1996). The selectivity of motion-driven visual attention in infants. *Child Dev.* 67, 2608–2620. doi: 10.1111/j.1467-8624.1996.tb01878.x
- Neumann, O. (1987). "Beyond capacity: a functional view of attention," in *Perspectives on Perception and Action*, eds H. Heuer and A. F. Sanders (Hillsdale, NJ: Lawrence Erlbaum Associates), 361–394.
- Nomikou, I., Schilling, M., Heller, V., and Rohlfing, K. J. (2016). Language-at all times. *Interact. Stud.* 17, 120–145.
- Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., and Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant Behav. Dev.* 18, 79–85. doi: 10.1016/0163-6383(95)90009-8
- Pegg, J. E., Werker, J. F., and McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: evidence from 7-week-old infants. *Infant Behav. Dev.* 15, 325–345. doi: 10.1016/0163-6383(92)80003-d
- Pereira, A. F., Smith, L. B., and Yu, C. (2014). A bottom-up view of toddler word learning. *Psychon. Bull. Rev.* 21, 178–185. doi: 10.3758/s13423-013-0466-4
- Plunkett, K. (1997). Theories of early language acquisition. *Trends Cogn. Sci.* 1, 146–153. doi: 10.1016/S1364-6613(97)01039-5
- Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., and Hennon, E. A. (2006). The birth of words: ten-month-olds learn words through perceptual saliency. *Child Dev.* 77, 266–280. doi: 10.1111/j.1467-8624.2006.00869.x
- Quine, W. V. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rohlfing, K. J., Longo, M. R., and Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Dev. Sci.* 15, 426–435. doi: 10.1111/j.1467-7687.2012.01139.x
- Rohlfing, K. J., Wrede, B., Vollmer, A. L., and Oudeyer, P. Y. (2016). An alternative to mapping a word onto a concept in language acquisition: pragmatic frames. *Front. Psychol.* 7:470. doi: 10.3389/fpsyg.2016.00470
- Samuelson, L. K., and Smith, L. B. (1998). Memory and attention make smart word learning: an alternative account of Akhtar, Carpenter, and Tomasello. *Child Dev.* 69, 94–104. doi: 10.1111/j.1467-8624.1998.tb06136.x
- Smith, L. B. (1995). "Self-organizing processes in learning to learn words: development is not induction," in *Basic and Applied Perspectives on Learning, Cognition, and Development. The Minnesota Symposia on Child Psychology*, Vol. 28, ed. C. A. Nelson (Mahwah, NJ: Lawrence Erlbaum Associates), 1–32.
- Smith, L. B., Jones, S. S., and Landau, B. (1996). Naming in young children: a dumb attentional mechanism? *Cognition* 60, 143–171. doi: 10.1016/0010-0277(96)00709-3
- Spencer, J. P., Perone, S., Smith, L. B., and Samuelson, L. K. (2011). Learning words in space and time: probing the mechanisms behind the suspicious-coincidence effect. *Psychol. Sci.* 22, 1049–1057. doi: 10.1177/0956797611413934

- Striano, T., Henning, A., and Stahl, D. (2005). Sensitivity to social contingencies between 1 and 3 months of age. *Dev. Sci.* 8, 509–518. doi: 10.1111/j.1467-7687.2005.00442.x
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics* 10, 401–413. doi: 10.1075/prag.10.4.01tom
- Tomasello, M., and Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cogn. Dev.* 10, 201–224. doi: 10.1016/0885-2014(95)90009-8
- Treue, S. (2003). Visual attention: the where, what, how and why of saliency. *Curr. Opin. Neurobiol.* 13, 428–432. doi: 10.1016/s0959-4388(03)00105-3
- Trevarthen, C. (1985). Facial expressions of emotion in mother-infant interaction. *Hum. Neurobiol.* 4, 21–32.
- Trueswell, J. C., Lin, Y., Armstrong, B. III, Cartmill, E. A., Goldin-Meadow, S., and Gleitman, L. R. (2016). Perceiving referential intent: dynamics of reference in natural parent–child interactions. *Cognition* 148, 117–135. doi: 10.1016/j.cognition.2015.11.002
- Vouloumanos, A., Hauser, M. D., Werker, J. F., and Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Dev.* 81, 517–527. doi: 10.1111/j.1467-8624.2009.01412.x
- Vouloumanos, A., and Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Dev. Sci.* 7, 270–276. doi: 10.1111/j.1467-7687.2004.00345.x
- Vouloumanos, A., and Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10, 159–164. doi: 10.1111/j.1467-7687.2007.00549.x
- Waxman, S. R., and Kosowski, T. D. (1990). Nouns mark category relations: toddlers' and preschoolers' word-learning biases. *Child Dev.* 61, 1461–1473. doi: 10.1111/j.1467-8624.1990.tb02875.x
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., and Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Dev. Psychol.* 34, 1289–1309. doi: 10.1037//0012-1649.34.6.1289
- Wilcox, B. M., and Clayton, F. L. (1968). Infant visual fixation on motion pictures of the human face. *J. Exp. Child Psychol.* 6, 22–32. doi: 10.1016/0022-0965(68)90068-4
- Wildt, E., and Rohlfing, K. J. (2018). “What type of interactional presentation helps 10-month-olds to overcome the saliency-effect during referent selection?” in *Poster Session Presented at the International Congress of Infant Studies (ICIS)*, Philadelphia, PA.
- Woodward, A. L. (1992). *The Role of the Whole Object Assumption in Early Word Learning*. Doctoral dissertation, ProQuest Information & Learning, Ann Arbor, MI.
- Yu, C., and Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition* 125, 244–262. doi: 10.1016/j.cognition.2012.06.016
- Yu, C., and Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Curr. Biol.* 26, 1235–1240. doi: 10.1016/j.cub.2016.03.026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Wildt, Rohlfing and Scharlau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.