Psychology of Language and Communication 2025, Vol. 29, No. 1

https://doi.org/10.58734/plc-2025-0019

Claire Prendergast
University of Oslo, Norway

What Young Children's Processing and Understanding of Compound Words Can Tell Us About Their Pragmatic Development

What can we learn by observing how children process and interpret compound terms? By integrating both linguistic and pragmatic factors, typically studied in isolation, the current study revealed children's growing adherence to linguistic norms, but also their increasing openness to unconventional reference. Across three experiments employing a picture selection task for referent selection, young children were presented with lexicalized and novel exocentric and endocentric compound nouns. Examining age-related differences in referent selection, Experiment 1 (baseline), found a preference for conventional and semantically transparent referents, increasing with age. Experiment 2 showed that an individual speaker influenced referent selection across both age groups, with 5-year-olds showing more accommodation of the speaker's intended meaning. Experiment 3, examining gaze behaviour, indicated that both 3- and 5-year-olds decompose lexicalized compound terms similarly to novel compounds. This research highlights the interplay between language and social development, showcasing key stages in children's pragmatic development.

Key words: social cognition, pragmatic development, compound processing

Address for correspondence: Claire Prendergast
IFIKK, University of Oslo, Norway.
E-mail: prendecn@gmail.com

As a linguistic construction, compound terms (e.g., *firetruck*) are highly productive in many languages including Germanic languages (English, German, Norwegian etc.). In languages that feature them, young children produce them abundantly in the first years of life and they act as a foundational structure or *pivot schema* in much of toddler and child speech (Tomasello, 2003). While previous work has shown children's early sensitivity to compound word structure (Clark et al., 1985), less is known about how pragmatic development shapes this process.

Research into young children's interpretations of noun-noun compounds in English has found evidence for sensitivity of the head-modifier relationship (e.g., the noun *fire* modifying head-noun *truck)* and subcategorization properties of compound terms from around 2-3 years of age (Clark et al., 1985; Mellenius, 1996). Children's understanding of the subcategorization properties of compounds has been further teased apart in more recent research in terms of the shared relational features of the constituent nouns in the compound (Snape & Krott, 2022), suggesting that 3 and 4 year olds tend not to consider the *higher-order relational* blend of the constituents (i.e., the functional or attachment relationship between the constituents; a *firetruck* is a truck used to fight fires) and more heavily rely on *non-relational* aspects such as an attributive matching of shared features in compound meaning resolution (e.g., a *firetruck* is a truck the same colour as fire). On the other hand, 5-year-olds, akin to adults, tend to place more weight on the *higher-order relational* structure between constituents in resolving whole compound meaning, suggesting a developmental shift around this age.

These studies used word learning paradigms and picture selection tasks to test young children's compositional ability or the *generative* aspect of their language acquisition with respect to *endocentric* compounds (i.e., where, in English, the right constituent noun specifies the referent that is then modified by the left constituent) by combining familiar and novel nouns into novel compounds (e.g., *moonbird*, referring to a type of bird). When interpreting these findings, one might consider age-related differences attributed to cognitive development more generally, such as improvements in *relational memory* between ages 4 and 6 (i.e., the ability to hold in mind and map the dependencies and links between different referents; Ngo et al., 2018) or the achievement of ternary status in *relational knowledge* between ages 4 and 5 (i.e., the ability to hold in mind at least three referents when computing a structural mapping of referent properties; Halford et al., 2010), which would both facilitate increased *higher-order relational* processing of compound terms. But what happens when children encounter lexicalized compounds, whose compound referent is not entailed by this head-modifier relationship, as is the case for so-called *exocentric* compounds (e.g., *bookworm* in reference to a lover of books)?

The Role of Semantic Transparency

Indeed, it is very often assumed that referents for lexicalized exocentric compounds are memorized as single lexical entries and are not *decomposed* (Li et al., 2022). However, it is important to consider that lexicalized compounds, despite becoming *entrenched* over time (i.e., ingrained meanings resistant to change or reinterpretation; Langacker, 1987), were once novel to all language learners. In light of this, if children do indeed demonstrate sensitivity to the morphosemantics of novel compound words, as previous research suggests (Clark et al., 1985), it follows that they likely process lexicalized compounds in similar ways because, for young language learners, their conventional meanings are simply less *entrenched.* While children have been shown to favour the most conventional referents, they have also been shown to prioritize semantic transparency when inferring compound meanings (i.e., how clearly the meaning of the compound can be inferred from the meanings of its constituent parts, Clark, 1995). Furthermore, while the figurative origins of lexicalized compounds can be obscured by historical drift, in many cases, much meaning can be retrieved compositionally. One example of this is "*belly button"* (Benczes, 2006, p. 96), which upon hearing the compound, adults will likely think of its synonym "navel', whereas children, who have not yet acquired the term "navel', might instead think of a "button" and then map this onto a "belly" as an attributive feature (i.e., something that resembles a button on a known part of the body), just like the 2-4-year olds did for novel compound formation in the studies by Snape and Krott (2022).

The Role of Contextual Constraint

The acquisition of lexicalized compounds is also likely constrained by the contexts in which the compounds are encountered over time in reference to specific entities and the frequency of such references, which aligns more with a usage-based theory of acquisition (Tomasello, 2009). For example, one can easily imagine a scenario in which a young child hears the infrequent compound *ladyfinger* in a café context and assumes that it is being used in reference to the finger of a nearby female. *Butterfly*, on the other hand, used in reference to the insect, is an extremely frequent and salient referent in children's picture books etc., and would be less likely to be interpreted *non-relationally* (i.e., relation-LOCATION; a fly alighted on butter). However, a *non-relational* interpretation such as this could also be a plausible referent in a café context where both constituent nouns: butter and flies, might be present. But if *butterfly* is instead uttered in a public garden, it becomes even less likely to be processed compositionally, unlike *ladyfinger* in any other context where there is at least one female with fingers. As Jackendoff (2010) explains in the context of compound meaning acquisition, "The task for the language learner, then, is to learn the lexicalized compounds and to acquire the rules for interpreting novel compounds. These two processes are evidently interdependent" (p. 420). Therefore, *contextual constraint* together with the *distributional frequency of meanings* are likely determining factors influencing how young children interpret and acquire lexicalized compound meanings across development. In essence, these factors are how *conventionality* is referred to in the current study, which assumes that conventionality itself is bound by broader needs for social coordination in communication.

The Role of Conventionality

The conventionality of exocentric compounds thus likely gradually begins to constrain compositional processing across early development. As conventionality is linked to social metacognition, in the current study, we aim to go beyond examining how young children generate compound word meanings solely through novel compounding, as studied previously (Snape & Krott, 2022), by assessing if and under what conditions, children *decompose* lexicalized compound terms, while also assessing cases where compounds are polysemes (both referents are conventional) and novel (neither referent is conventional). Moreover, focusing on the role of conventionality, we embark on the exploration of whether, how, and to what extent the process of *decompositionality* (Bien et al., 2005) is influenced by broader social development associated with age.

Language conventions can be understood in terms of *codification* and *scope restriction* (Kalish & Sabbagh, 2007) and the current study uses this general framework to address the role conventionality may play, more specifically, in young children's interpretations of compound nouns. *Codification* pertains to the level of establishment of the conventional referent *across contexts* as some conventional compounds might be more temporary and situational (i.e., less *codified*) than others. Notably, children are not fully integrated into conventional communities, which may impact their flexibility in adhering to conventions when compared to adults (Kalish & Sabbagh, 2007). The *degree of codification* in conventional language could influence the acquisition process, in so far as children may treat ad hoc and more established conventional referents differently. *Scope restriction*, on the other hand, refers to the capacity of conventions to encompass both flexibility and stability *within a particular context*. Recognizing *the scope restriction* of a convention involves understanding the contextual factors that define its reach. The pivotal question for developmental acquisition thus pertains to how children discern the limitations of a convention's applicability and identify the principles that apply across or within specific contexts.

The Role of the Speaker

More generally, conventional language use enables the representation of collective intentions and the maintenance of community norms facilitating smooth communication between community members (Kalish & Sabbagh, 2007). This implies that as children accumulate more social experience, they will adhere more closely to, and anticipate, the most conventional language expressions from speakers. However, since individual speakers commonly deviate from using the most conventional expressions, precisely to add new information (Kallen et al., 2022), children must be simultaneously flexible in their interpretations. One such example is the use of nonliteral language which is often employed to achieve overarching social objectives (e.g., employing irony to regulate affect; Goodman & Frank, 2016) and which young children are known the struggle with, as exemplified by their literal interpretations (Winner et al., 1988). Broader social cognitive development, such as the ability to model other's beliefs (i.e., "theory-of-mind'; Wellman, 2017), which tends to show up more reliably when using standardized false-belief tests in children between 4 and 5 years old, may be the age at which children begin considering the superordinate social goals that speakers may have in communication. So, although children around age 5, may be anticipating conventional language use due to their greater social experience, compared to age 3, they may also be simultaneously more accommodating of unconventional language use when it occurs due to their greater ability to consider an individual speaker's intentions.

This proposal finds further support in the knowledge that conventionality in language use comes from the "enforcement" of normative social practices at large (Prendergast, 2024; van den Herik, 2021). In the broader social domain of coordinated action, for instance, 5-year-olds will protest norm deviations in game playing by focusing on the violator, as well as the violations. In contrast, 3-year-olds care most about the violations and are less inclined to consider the role of the violator in their assessments (Köymen et al., 2014). Therefore, age-related differences in broader social development may be strong predictors of patterns in pragmatic inferences made by young children in the case of compound interpretation. This implies that 5-year-olds may be more attuned than 3-year-olds to conventional language use (i.e., the normative meaning of the utterance), but also more attentive to an individual speaker's intended meaning if it deviates from the conventional one.

Additionally, since young children are thought to follow the principle of mutual exclusivity (ME) when inferring referential meaning (i.e., a one-to-one mapping of words and their referents; Kalashnikova et al., 2019), once the conventional referent is known, they may rely solely on that meaning (Clark, 1980), strictly rejecting any competing alternative referents. With age, children are thought to become more flexible in allowing referents to share the same labels (Sia & Mayor, 2022), which would lead to the same predictions in increased openness to unconventional referents at age 5 compared to at 3. Moreover, in related experiments, seeing the speaker as opposed to just hearing them, has been shown to lead to more ME flexibility in older children (Kalashnikova et al., 2016). Since mutual exclusivity has been used to tap *sensitivity to speaker informativity* in models of pragmatic inferencing in children (Bohn et al., 2021), this suggests that, with age, considering the role of the speaker (and their intentions) to a greater extent may explain key developmental differences between these age groups in their interpretations.

Finally, we might expect to see some of the same patterns that were found in previous studies showing differences in 3- and 5-year-olds" processing of lexicalized noun-noun compounds, as we do for novel ones. This entails that for 3-year-olds, the relational structure of lexicalized compounds may be under-analyzed (Snape & Krott, 2022), while 5-year-olds, who have shown better comprehension of the relational structure of the compound blend in previous work, may be more flexible in their interpretations of the whole compound term.

The current study thus sought to connect and integrate morphological, pragmatic, and cognitive lines of research into a unified approach by looking at how children navigate the tension between deconstructing forms (morphological analysis) and reconstructing meanings (pragmatic inference).

The Current Study

Therefore, in a series of three experiments, through the frame of *codification* and *scope restriction*, we probed the ways in which *conventionality* and *referential transparency* (Experiment 1), speaker intention (Experiment 2), and finally contextual constraint (Experiment 3) may affect children's processing and understanding of lexicalized and novel noun-noun compounds in Norwegian, while exploring the possible age-related differences that may exist between 3- and 5-year-olds in compound interpretations as a window onto greater pragmatic development. A fourth experiment was run but is omitted from the final report (see Appendix S5 for more details).

Compounding in Norwegian

Compounding is highly productive in the Norwegian language, and it bears most of the compound properties associated with other Germanic languages such as English (Eik, 2019). Noun-noun compounds are the most frequent form of compounds found in Norwegian and are usually agglutinated (Eiesland, 2016). While compounding in Norwegian, unlike English, is mostly a morphological process, as opposed to syntactic one (Fares, 2019), the semantic relationship between the head and modifier nouns explored in the current study aligns across both languages in terms of exocentric and endocentric reference and the frames of analysis of the current study which are conventionality and referential transparency.

Conventionality and Transparency

In the current study, a distinction between conventional and unconventional compound reference is based on the lexical entry for the *whole compound term*, as it usually and commonly refers to the entity represented in the visual stimuli (i.e., the most frequent and expected visual referent for that whole compound term). For instance, a compound term such as "butterfly" is classified as *conventional* when paired with an image of the colourful insect because this visual representation closely matches the *codified* meaning of the term. Conversely, if "butterfly" were to be paired with an image of "a regular black fly alighting on a block of butter,' it would be classified as *unconventional* because the visual representation diverges from the *codified* meaning of the whole compound term.

Furthermore, the *referential transparency* of the compound-image pairs is considered by first breaking down the compound term into its constituent nouns and then analyzing how well the visual representation reflects this decomposition (i.e., the constituent blend). For example, analysing the compound term *bookshelf* according to its internal *semantic transparency* in the first instance (e.g., *book+shelf* should refer to a shelf on which to place books) and then mapping this onto an available referent. If the referent image for the compound noun *bookshelf* clearly depicts a shelf with books on it this results in high overall *referential transparency* for the compound-image pair compared to an image showing stacks of books upon which other objects are placed, for example, which would have a lower rating. This *visual representational quality* of the referent images (i.e., the mapping from word to image) is important in the current study due to the use of picture selection to assess referent choice in children.

Considering *referential transparency* also helps to better understand the *scope restriction* of conventional compound referents. For instance, a mapping of the decomposed compound term "butter+fly" onto an image referent showing "a regular fly alighting on a block of butter', making it an unconventional but endocentric reference, is relatively higher in *referential transparency* than a mapping onto the conventional referent of the colourful winged insect. The *referential transparency* of this conventional exocentric referent is relatively lower because the head noun refers to a different type of *fly*-insect not depicted in the referent image. However, even exocentric compound referents may share attributive features with the constituent head and modifier noun referents (e.g., *bellybutton*). In this case, the head-noun *fly* also has-wings and is a similar colour to the modifier, *butter*, which may allow it to be partially retrieved. The current study thus investigated the referent preferences that L1 Norwegian-speaking 3- and 5-year-olds may have for compound terms that vary in *conventionality* and *referential transparency.*

Materials and Methods

All participants were recruited for the current study through existing participant databases and Meta Advertisements (Facebook and Instagram). Information on participant ethnicity was not recorded for the purposes of this study. All participants had Norwegian as a first language due to one or both parents being native-Norwegian speakers and/or full-time *barnehage* (kindergarten) in Norwegian. Some participants had one or more additional languages spoken at home (Experiment 1; N=8, Experiment 2; N=12, Experiment 3; N=11). These included English (9), French (4), Spanish (3), German (2), Danish (2), Russian (2), Italian (2), Urdu (2), Dutch (1), Swedish (1), Icelandic (1), Polish (1), Lithuanian (1), Punjabi (1), Vietnamese (1), Nepali (1) and Indonesian (1). The sample was thus representative of a typical European capital city. Ethical approval for the experiments was granted by the Norwegian Data Protection Services (NSD).

Across all three experiments with children, we used a picture selection task with a binomial choice outcome for 18 noun-noun compound terms in Norwegian. We used data from 17 of the items for the analysis (see Appendix S1.1 for information on one excluded item whose constituent head noun was a homonym). We selected the compound words for the current study to ensure an early age of acquisition (AoA) for the compound meanings in addition to their constituent noun meanings, as well as the ability to clearly visually represent both exocentric and endocentric interpretations of the compounds for the picture selection task. These included compounds with highly common lexical items such as "horse', "bear', "apple', "star" etc. Children's familiarity with the references was further confirmed by their parents.

Compound Classification and Ratings

Compound-image pairs were first rated as *conventional* and *unconventional* by members of the research team through lexical searches on Norsk ordbok (*Ordbøkene.no*, 2022) where *conventional* referents had to be listed and *unconventional* referents, absent. Ratings for the *referential transparency* for each compound-image pair were further obtained by surveying native Norwegian-speaking adults through the Prolific platform (N=23). Adult participants first rated the compound words linguistically for their internal semantic transparency on a Likert scale slider bar ranging from -10 (not at all transparent) to 0 (somewhat transparent) to 10 (completely transparent), and secondly in terms of their visual transparency in reference to the images used in the experiments on the same scale levels, culminating in a mean rating on both dimensions, the mean of which was then calculated to provide an overall rating of *referential transparency* for each compound term (for more detailed information see Appendix S2). The mean ratings of *referential transparency* for each compound-image pair are shown in Figures 1 and 2. To further contextualize the study, compound terms were classified according to Jackendoff (2010) (e.g., exocentric/endocentric), allowing some balance in the presentation and comparison of items within each block for analysis purposes.

All experiments were run using Experiment Builder on an Eyelink 1000 Plus eye-tracker in remote mode. We began the experimental session with Set 1 of the Norwegian version of the BPVS II (Lyster et al., 2010). This was a practise round for the picture selection task to follow (i.e., pointing to the referent on screen that best represents the word heard). Next, a 5-point calibration was performed for each child participant.

Experiment 1: Baseline Measure for Choice Preference

Procedure

In Experiment 1 (baseline), children were informed that they would be seeing some images and hearing some words that the experimenter (posing as a kindergarden teacher) had been learning with her kindergarden class, and that the experimenter would like to find out what kind of things the child participant knows by inviting them to take part in a picture selection task. For each pair of images shown onscreen, participants heard a pre-recorded audio file prompting them to point to the image that best shows the compound word uttered (e.g. Can you point to the image that best shows a [butterfly]?) Please see Appendix S3 for more detailed information.

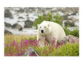
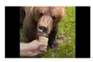
The experimenter then selected the image that the participant pointed to with the mouse, moving the experiment onto the next item. If the child's answer was considered invalid, the experimenter clicked in the middle of the screen to move the experiment on. Invalid answers included pointing before the prompt had completed playing or not pointing to one image after approximately one minute, or up to three additional prompts by the experimenter. When necessary, the experimenter encouraged the participants to point to an image by emphasising that there were no right or wrong answers and that their answers should be based on what image *they* thought *best* shows the word spoken.

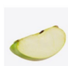
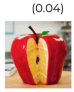
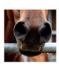
Compound-image pairs were presented in two blocks. *Block 1.1* showed *items varying in conventionality* (items shown in Figure 1) to tap referential choice preference as governed by *conventionality.* The items were presented in randomized order within each subgroup: *1.1.1* which showed lexicalized exocentric and endocentric compounds (6 items) pitted against unconventional endocentric and exocentric referents higher in *referential transparency*; and *1.1.2* which presented lexicalized endocentric compounds (4 items) pitted against unconventional exocentric and endocentric referents lower in *referential transparency*. For the analysis, we combined the 10 items to track choice preferences for conventional or unconventional referents.

We expected that both 3- and 5-year-olds would have acquired all the constituent noun meanings, as well as the lexicalized exocentric compound meanings. To support this assumption, parents were asked to indicate whether they thought their child knew the lexicalized exocentric referent for compounds on a scale of -1 (sure the child does not know referent) to 0 (unsure if they do or don't know referent) to 1 (sure they know referent). Parents indicated that they expected their children to know these referents to a satisfactory degree with the lowest mean ratings interpreted around the label anchor "unsure'. The mean ratings are shown in brackets for each age group beside the lexicalized exocentric compounds in Figures 1 and 2 (See Appendix S1.2 for variability in ratings). These ratings combined with children's choice behaviour supports the assumed AoA of the items.

Figure 1

Compound-Image Pairs from Block 1.1. Item-Pairs Varying in Conventionality Across Two Subgroups 1.1.1 and 1.1.2

1.1.1 Lexicalized exocentric and endocentric compound pairs (Age: MEAN parent's likelihood rating of knowledge)	**Conventional Term** (MEAN Referential Transparency)	**Unconventional Term** (MEAN Referential Transparency) *Experiment 2 Choice	**Experiment 3 Non-related Distractor**
Sjøhest // "Sea+Horse" // Seahorse (3: 0.40, 5: 0.74)	Attributive exocentric (-0.18) 	Subordinate endocentric (0.43) 	
Sjøstjerner // "Sea+Stars" // Starfish (3: 0.84, 5: 0.98)	Attributive exocentric (3.48) 	Subordinate endocentric (0.73) 	
Sommerfugl // "Summer+Bird" // Butterfly (3: 1, 5: 1)	Attributive exocentric (-3.23) 	Subordinate endocentric (3.04) 	
Sukkertøy // "Sugar+Clothing" // Candy (3: 0.06, 5: 0.43)	Attributive exocentric (-3.09) 	Attributive endocentric (N/R) 	
Piggsvin // "Spike+Pig" // Hedgehog (3: 1, 5: 1)	Attributive exocentric (0.78) 	Attributive exocentric (NR) 	
Isbjørn // "Ice+Bear" // Polar Bear (3: 0.98, 5: 1)	Subordinate endocentric (2.41) 	Coordinate endocentric (2.64) 	

1.1.2 Lexicalized exocentric and endocentric compound pairs	**Conventional Term** (Referential Transparency)	**Unconventional Term** (Referential Transparency) *Experiment 2 SPEAKER ChOICE	**Experiment 3 Non-related Distractor**
Epleskive // "Apple+Slice"	Subordinate endocentric (7.24) 	Subordinate exocentric (-1.98) 	
Eplekake // "Apple+Cake"	Subordinate endocentric (7.15) 	Coordinate endocentric (0.04) 	
Hest(e)nese // "Horse+Nose"	Subordinate endocentric (6.53) 	Attributive exocentric (-6.03) 	
Hest(e)ører // "Horse+Ears"	Subordinate endocentric (6.85) 	Attributive exocentric (-6.60) 	

If the children did not know the referent for the lexicalized compounds, they would not choose them above chance in Experiment 1 (baseline). Additionally, choice behaviour at ceiling across both conditions of Experiment 3 when both referents are presented onscreen with a non-related distractor further support this assumption even though, in this case, knowledge of the non-related distractor shown onscreen could also have led to the correct identification through mutual exclusivity. However, taken together, all three sources of evidence can be considered consistent in support of the assumption.

We next investigated the preference for *referential transparency* in *Block 1.2 items varying in referential transparency* (items shown in Figure 2) with two further sub-groupings holding *conventionality* constant including: *1.2.1 polysemic compounds* (3 items; one item excluded as head noun had more than one meaning) featuring pairs of conventional exocentric and endocentric referents, as well as *1.2.2 novel compounds* (4 items) with pairs of unconventional exocentric and endocentric visual referents analysing the 7 items together to determine the role of *referential transparency* on choice behaviour.

An early version of the preregistration for Experiment 1 was registered on OSF (see Appendix S6 for more information). Data sets from all experiments are available on OSF: https://osf.io/4thjc/.

Hypotheses

In Experiment 1 (baseline), (1.1) it was predicted that *conventionality* would determine choice of referent for lexicalized noun-noun compounds to a greater extent than *referential transparency* across age groups (Clark, 1995). For polysemic and novel compounds, (1.2) we made a prediction that *referential transparency* would lead to differences in choice preferences between 3-year-olds and 5-year-olds based on the possible tendency to default to *referential transparency* when both referents are conventional (polysemes) or when no conventional referent exists (novel). We predicted that 5-year-olds would choose referents highest in *referential transparency* to a greater extent than 3-year-olds, since those were the options showing *higher-order relational* structure mappings between constituents (Snape & Krott, 2022).

Power Analysis

The sample size for Experiment 1 (baseline) was estimated by the age-related differences expected in choice preferences, according to hypothesis (1.2). As such, a target sample size was initially estimated for a two-sample proportions test on choice of referent with an estimated moderate effect size of $\delta = 0.4$ and indicated that N=46 (n=23 3-year-olds) would result in power of alpha = 0.8. The sample size for all other experiments in the study was determined by this power analysis for Experiment 1 (baseline). Where possible excluded participants were replaced by new participants to reach the required sample size. However, reduced time and resources, and participant recruitment challenges meant that in

Figure 2

Compound-Image Pairs from Block 1.2 Item-Pairs Varying in Referential Transparency Across Two Subgroups 1.2.1 and 1.2.2. Where Conventionality was Held Constant (i.e., Polysemes and Novel Case

1.2.1 Polysemic compounds (Age: MEAN parent's likelihood rating of knowledge EXOCENTRIC REFERENCE)	**Higher transparency Term** (MEAN Referential Transparency)	**Lower transparency term** (MEAN Referential Transparency) **Experiment 2 SPEAKER ChOICE*	**Experiment 3 Non-related Distractor**
Hest(e)hale // "Horse+Tail" // Ponytail (3: 0.65, 5: 0.71)	Subordinate endocentric (9.19) 	Attributive exocentric (-2.65) 	
Hestehov // "Horse+Hoof" // Dandelion (3: 0.27, 5: 0.59)	Subordinate endocentric (6.64) 	Attributive exocentric (-7.84) 	
Løvetann // "Lion+Tooth" // Dandelion (3: 0.63, 5: 0.86)	Subordinate endocentric (7.74) 	Attributive exocentric (-7.21) 	

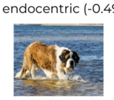
1.2.2 Novel compounds	**Higher transparency Term** (Referential Transparency)	**Lower transparency Term** (Referential Transparency) **Experiment 2 SPEAKER ChOICE*	**Experiment 3 Non-related Distractor**
Sitronpinne // "Lemon+Stick"	Subordinate endocentric (3.24) 	Attributive exocentric (-6.96) 	
Sjøhund // "Sea+Dog"	Subordinate endocentric (-0.49) 	Attributive exocentric (-1.85) 	
Kattebrød // "Cat+Bread"	Coordinate endocentric (1.79) 	Surbordinate endocentric (-0.32) 	
Kuball // "Cow+Ball	Coordinate endocentric (2.93) 	Subordinate endocentric (-1.60) 	

Table 1

Proportion of Choice of Conventional and Unconventional Referents from Block 1.1 Across All Experiments

	3-year-olds			5-year-olds		
	Prop	SE	[95% CI]	Prop	SE	[95% CI]
Experiment 1						
Conventional	.7957	.0266	[.7386, .8429]	.8318	.0252	[.7764, .8757]
Unconventional	.2043	.0266	[.1571, .2614]	.1682	.0252	[.1243, .2236]
Experiment 2						
Conventional	.7143	.0312	[.6494, .7714]	.6000	.0330	[.5338, .6627]
Unconventional	.2857	.0312	[.2286, .3506]	.4000	.0330	[.3373, .4662]
Experiment 3						
Target	.9755	.0108	[.9424, .9898]	.9842	.0090	[.9522, .9949]
Distractor	.0245	.010	[.0102, .0576]	.0158	.0090	[.0051, .0478]

Table 2

Proportion of Choice of Higher Referential Transparency and Lower Referential Transparency Referents from Block 1.2 Across All Experiments

	3-year-olds			5-year-olds		
	Prop	SE	[95% CI]	Prop	SE	[95% CI]
Experiment 1						
Higher Transparency	.6584	.0374	[.5817, .7276]	.6169	.0392	[.5377, .6903]
Lower Transparency	.3416	.0374	[.2724, .4183]	.3831	.0392	[.3097, .4623]
Experiment 2						
Higher Transparency	.6803	.0385	[.6006, .7507]	.5714	.0399	[.4920, .6473]
Lower Transparency	.3197	.0385	[.2493, .3994]	.4286	.0399	[.3527, .5080]
Experiment 3						
Target	.9592	.0163	[.9120, .9816]	.9925	.0075	[.9485, .9989]
Distractor	.0408	.0163	[.0184, .0880]	.0075	.0075	[.0011, .0515]

Experiments 2 and 3, the full sample size of N=46 was not achievable.

Participants

A total of 27 3-year-olds (N = 9 females, M_{Age} = 3 years, 5 months, 27 days) and 23 5-year-olds (N = 8 females, M_{Age} = 5 years, 5 months, 21 days) participated. The final sample (N = 45) was then reduced by exclusions which amounted to four 3-year-olds (n = 2 parent influenced answers, n = 1 would not point, n = 1 wrong age) and one 5-year-old (n = 1, would not point).

Data Preparation and Analyses

Invalid trials amounted to 0.06% of the total amount of trials. Proportion of choice across all compound items was calculated. Choice data is provided in Tables 1 and 2.

Since picture selection tasks are discrete choice models, we modified our original plan to use two-sample tests of proportion of choices, opting for a mixed logit model in Stata (see Appendix S4 for more details). to enable us to model the effects of both the alternative-specific covariates (*referential transparency* for each compound item) and case-specific covariates (age) on choice behaviour (conventional vs. unconventional/higher transparent vs. low transparent) across participant and compound item, which we modelled as random intercepts through data structuring (i.e., command *cmset* and *cmxtmixlogit*; StataCorp LLC, 2023]).

Model 1.1 – *Conventionality* predicts choice of lexicalized referent to a greater extent than *referential transparency* across both age groups.

$\text{ChoiceConventional}_{ijg} \sim \text{referential transparency} + \text{age}_k + (1 \mid \text{item}_i) + (1 \mid \text{participant}_j)$

Model 1.2 – Differences in choice behaviour for polysemes and novel compounds at 3 and 5 years accounted for by *referential transparency.*

$\text{ChoiceTransparent}_{ijg} \sim \text{referential transparency} + \text{age}_k + (1 \mid \text{item}_i) + (1 \mid \text{participant}_j)$

Results

Model 1.1. fit our prediction that *conventionality* governed choice of referents to a greater extent than *referential transparency* which didn't play a significant role in choice behaviour (β = -.0088797, SE = .023976, 95% CI [-.0558718 - .0381124], p = 0.711) across age groups for referent selection in Block 1.1. The model also indicated that 5-year-olds were more likely to choose conventional referents than 3-year-olds with a moderate range in confidence intervals (β = .6394801, SE = .322406, 95% CI [.007576 -1.271384], p = 0.047). However, both age groups had high, above chance, marginal probabilities in their choice of the conventional referents at 90% and 83%, respectively (see Table 3).

Table 3

Marginal Probabilities for Model 1.1 and 1.2

	3-year-olds			5-year-olds		
	Prop	SE	[95% CI]	Prop	SE	[95% CI]
Model 1.1						
Conventional	.8315	.0271	[.7774 .8856]	.9034	.0220	[.8598 .9470]
Model 1.2						
Higher Transparency	.6583	.0358	[.5881 .7286]	.6168	.0374	[.5434 .6904]

Model 1.2 showed that our prediction for an age-related difference in choice behaviour in relation to *referential transparency* did not hold (β = .1958615, *SE* = .2450716, 95% *CI* [-.28447 - .6761929], $p = 0.424$). However, *referential transparency* did significantly impact choice of referent overall (β = .1154971, *SE* = .0232713, 95% *CI* [.0698862 - .161108], $p < 0.001$). The marginal probabilities for choice of referent highest in *referential transparency* were at 66% for 3-year-olds and 62% for 5-year-olds (see Table 3), which showed that, in fact, 3-year-olds were slightly more inclined to choose image-referents higher in *referential transparency* than 5-year-olds, a trend opposing the direction predicted, although not to significant degree.

Discussion

In Experiment 1 (baseline), the data revealed that in early childhood, in line with Clark (1995), preferences for lexicalized compound referents are largely driven by their *conventionality*, or in other words, due to their *codification*. This factor appears to guide children's interpretations when asked to select the most suitable referents, even when these referents have lower *referential transparency* than the alternative options shown onscreen. Furthermore, in this experiment, 5-year-olds (90%) displayed tendency to select conventional exocentric referents more often than 3-year-olds (83%), even though all the lexicalized compound referents used in the study were expected to have been acquired by age 3. This suggests that 5-year-olds more strictly adhere to conventionality in their interpretations than 3-year-olds, showing an increase in knowledge of *codification*. Conversely, while the results might suggest that 3-year-olds were relying more on *referential transparency* than 5-year-olds, it remains that *referential transparency* did not significantly covary with choice behaviour in either age group for Block 1.1., and therefore doesn't explain 3-year-old's slightly weaker tendency to select the conventional referents in this experiment, also suggesting that it was not lack of knowledge of the conventional term that governed choice at age

3 (i.e., they defaulted to *referential transparency* to infer the meaning). As such, the age-related difference in choice is likely driven by increasing adherence to *conventionality* at 5 compared to 3, as per hypothesis 1.1.

However, as anticipated, *referential transparency* did play a significant role in shaping choice behaviour for the polysemic and novel compounds where conventionality between referent options onscreen is held constant. Yet, there were no age-related differences in choice behaviour across these 7 items. While 3-year-olds (66%) were somewhat more inclined than 5-year-olds (63%) to choose the referent with the highest *referential transparency*, this difference did not reach statistical significance. For the polysemes, when the compound label conventionally referred to both referents onscreen, children did seem to rely on *referential transparency* by selecting the endocentric compound. However, the referents highest in *referential transparency* for each pair represented a *higher-order relational* mapping, i.e., they showed a functional or attachment relationship between constituent nouns, as in Snape and Krott (2022), while the alternatives were based on *non-relational* attributive featural mappings (e.g., *kuball (cow+ball);* the image rated highest in *referential transparency* showed a cow playing with a ball (functional), while the alternative rated lower in *referential transparency* showed a ball that resembled a cow (attributive)). In this sense, both 3- and 5-year-olds showed a tendency toward compositional processing based on functional relations, and against the prediction, 3-year-olds did so to a slightly greater extent (1.2). This may be because participants had just been presented with 10 lexicalized exocentric and endocentric compounds in which *functional relational* mappings were present, although it is unclear what role this may have played in influencing choice behaviour for subsequent novel terms in the current study, it has been shown that children's compound interpretations are influenced by their familiarity with existing compound terms more generally (Nicoladis & Krott, 2007).

Experiment 2: Speaker Endorsement of Unconventional and Semantically Opaque Referents

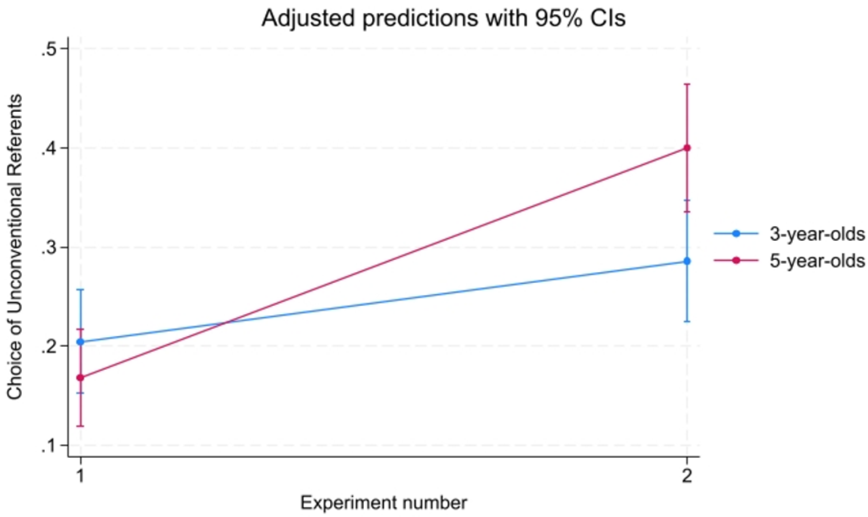
Procedure

In Experiment 2, additional characters were added to the paradigm to test whether a knowledgeable speaker's consistent endorsement of the unconventional image referents across all compounds would decrease the choice preference for conventional referents from Experiment 1 (baseline). In addition, we explored if having the speaker select the referents lowest in *referential transparency* across polysemic and novel items would affect preferences for referents highest in *referential transparency* when compared to Experiment 1 (baseline).

Experiment 2 therefore used the same compound-image pairs as in Experiment 1 (baseline). However, children were first introduced to two characters onscreen using a pre-recorded audio file. The design was developed to seem as

Figure 3

Marginal Probabilities for Choice of Unconventional Referents from Model 2.2 Across



naturalistic and as relevant as possible to the children to engage them in the task by situating it in a familiar kindergarden context. As such, if the participant was female, they were introduced to “Ida” as a knowledgeable native-Norwegian speaker who is teaching her classmate Max, a non-native learner, some words in Norwegian. If the participant was male, they were introduced to “Max” as a knowledgeable native-Norwegian speaker who is teaching his classmate Ida, a non-native learner, some words in Norwegian. The reason for gendered assignment was due to a possible “cooties” effect in the sample age range (Telzer et al., 2015). This is where children in the age range of 4-6 years old have been shown to hold a negative bias toward opposite-sex members. This came to light during a test session for the study when a 5-year-old participant verbalized a message consistent with this effect by saying opposite sex members are always wrong! Participants were informed that they would first be seeing what the knowledgeable native-Norwegian speaker selected as the best referent image for each compound term. Directly after being shown the knowledgeable speaker’s choice, the child-participant was asked to select the image that *they* thought best represents the compound word (see Appendix S3 for scripts and example trials).”

In Experiment 2, (2.1) we expected that choice preferences for lexicalized compounds would be significantly reduced due to the influence of the knowledgeable speaker's preference for unconventional options when compared to Experiment 1 (baseline) and that (2.2) there would be age-related differences between groups in the changed rate of choice of conventional referents across items. We further evaluated possible changes in rate of choice of referents highest in *referential transparency* (2.3).

Sample Size

A new sample of 26 3-year-olds (N=16 females, M_{Age} = 3 years, 4 months, 18 days) and 24 5-year-olds (N=16 females, M_{Age} = 5 years, 4 months, 1 days) were tested. The final sample (N=43) resulted from the exclusion of five 3-year-olds (n = 3 technical error, n = 1 pointed before shown peer choice, n = 1 verbalized choice not related to task question) and two 5-year-olds (n = 2 technical error).

Data Preparation and Analyses

Invalid trials amounted to 0.04% of the overall data. Please refer to Tables 1 and 2 for proportion of choice that was calculated across all compound items. The models for Experiment 2 were as follows.

Model 2.1 *Within age group effect of knowledgeable speaker's preference on choice behaviour for conventional compounds across experiments*

$ChoiceConventionalAge_{ijg}$ ~ referential_transparency + $Experiment\_no_l$ + (1 | $Item_i$) + (1 | $Participant_j$)

Model 2.2 *Between group effect of knowledgeable speaker's preference on choice behaviour for conventional compounds across experiments.*

$ChoiceConventional_{ijg}$ ~ referential_transparency + Age_k + $Experiment\_no_l$ + $Experiment\_no_l$:Age_k + (1 | $Item_i$) + (1 | $Participant_j$)

Model 2.3 *Between age group effect of knowledgeable speaker's preference on choice behaviour for polysemic and novel compounds across experiments.*

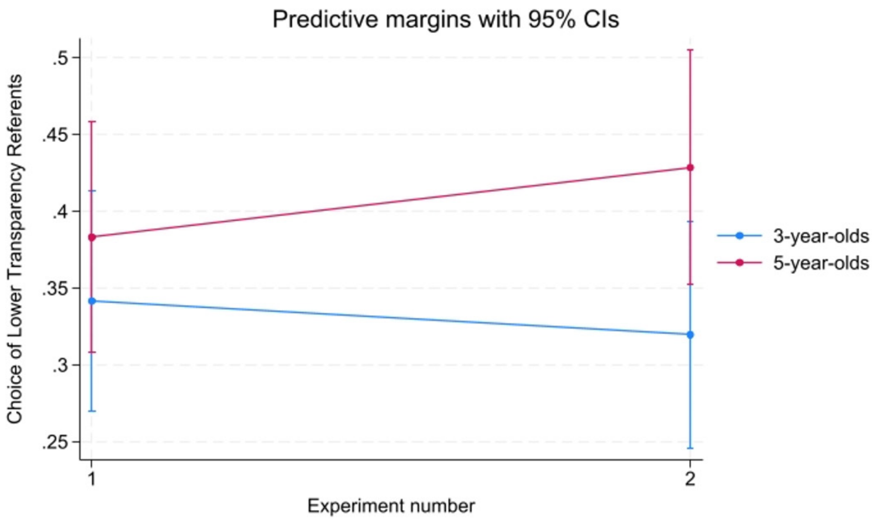
$ChoiceTransparent_{ijg}$ ~ referential_transparency + $Experiment\_no_l$ + Age_k + $Experiment\_no_l$:Age_k + (1 | $Item_i$) + (1 | $Participant_j$)

Results

With regards to Model 2.1, the findings revealed that both 3-year-olds (β = .6218155, *SE* = .2623273, 95% *CI* [.1076634 - 1.135967], $p = 0.018$) and 5-year-olds (β = 1.576449, *SE* =.3006667, 95% *CI* [.9871534 - 2.165745], $p < 0.001$) displayed a shift in their choice behaviours, aligning with the choice preferences of the knowledgeable speaker for unconventional referents when it came to interpreting conventional exocentric and endocentric compound pairs (Block 1.1.).

Figure 4

Marginal Probabilities for Choice of Lower Transparency Referents from Model 2.3



In Model 2.2, the data indicated that 5-year-olds exhibited a stronger tendency than 3-year-olds to shift their choices to *unconventional* referents in accordance with the knowledgeable speaker's choices in Experiment 2 compared with Experiment 1 (baseline). Specifically, there was a significant negative main effect with age (β = -.639389, *SE* = .3223821, 95% *CI* [-1.271246 - -.0075316], *p* = 0.047) and a main effect of experiment (β = .6212894, *SE* =.2622104, 95% *CI* [.1073666 - 1.135212], *p* = 0.018), together with an interaction between age and experiment (β = .9556623, *SE* = .398972, 95% *CI* [.1736916 - 1.737633], *p* = 0.017) in the data (See Figure 3). Post-hoc chi-square tests run on the contrasts between predictive margins on all interaction levels of age and experiment were statistically significant in the model. At age 3, the predicted probability of choosing unconventional referents increased by 0.105 (95% *CI* [0.019, 0.192]) when moving from Experiment 1 (baseline) to 2. At age 5, this resulted in an increase of 0.244 (95% *CI* [0.162, 0.327]). Moving from age 3 to 5 in Experiment 1 (baseline), the predicted probability decreased by 0.072 (95% *CI* [-0.141, -0.002]) in choice of unconventional referents. Finally, moving from age 3 to 5 and Experiment 1 (baseline) to 2, the predicted probability increased by 0.172 (95% *CI* [0.084, 0.261]).

For Model 2.3, the data showed that neither 3-year-olds nor 5-year-olds

(β = -.1875037, *SE* = .2397275, 95% *CI* [-.6573609, .2823536], *p* = 0.434) were inclined to distinctly shift their preferences from Experiment 1 (baseline) in line with the speaker's selection of lower transparency referents for polysemic and novel compounds in Experiment 2 (β = .1029998, *SE* = .2476345, 95% *CI* [-.3823549, .5883546], *p* = 0.677) (See Figure 4).

Discussion

While *conventionality* played a greater role in determining choice behaviour than *referential transparency* for lexicalized compound terms, as shown in Experiment 1 (baseline), it seems that when a knowledgeable speaker is added to the referential context, young children's referent choices, particularly those of 5-year-olds, are open to change to unconventional options in line with the speaker's choice (2.1). The significant difference between age groups here likely stems from how much weight children assign to an individual speaker's intended meaning in context, according to their sensitivity of the *scope restriction* of a conventional label. In a context where the unconventional referent is higher in *referential transparency*, as in lexicalized exocentric compounds, this effect of openness to an alternative unconventional meaning showed up in the data (2.2).

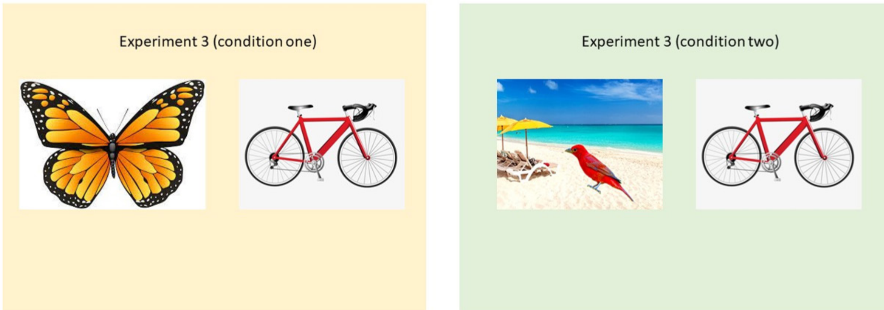
However, no such effect of influence was observed in the polysemic and novel items. This seems to support the idea that where conventional meanings exist, they may be up for renegotiation in certain contexts. In other words, where no such wider conventional referent exists or is a polyseme, children may be more likely to follow their own intuitions, relying most on the *referential transparency* of the compound-image pair. Another possible explanation is that since the block with polysemic and novel compounds came after the block with lexicalized exocentric compounds, the children may have started to question the speaker's reliability after witnessing a sequence of unconventional choices (Ronfard & Lane, 2018). However, they were reminded of the speaker's "knowledgeable" status in between blocks, so interpreting this data pattern by considering the unique and joint presence and absence of *conventionality* across blocks remains valid (2.3).

Experiment 3: Gaze Behaviour Differences due to Contextual Constraint

In Experiment 3, we focused on gaze behaviour differences to target vs. an alternative referent onscreen at test. As such, we introduced two experimental conditions. Condition one had the conventional/transparent options from each image-pair (items from the leftmost column in Figures 1 and 2) as the target referent presented alongside an unrelated distractor as an alternative (items from the rightmost column in Figures 1 and 2). Condition two had the unconventional/less transparent options from each image-pair (items from the middle column in Figures 1 and 2) as the target referents presented alongside the same unrelated distractors from the rightmost column as an alternative. Participants were ran-

Figure 5

Example Presentation of Referents for Compound "sommer+fugl" (summer +bird; butterfly) Across Conditions one and two in Experiment 3



domly assigned to either condition. See Figure 5 for an example.

Therefore, by comparing gaze behaviour data from both conditions one and two in Experiment 3 with the gaze behaviour from Experiment 1 (baseline), we aimed to gain insights into the effects of *contextual constraint* on children's processing of compounds. *Contextual constraint* refers to the degree of ambiguity caused by the presence of an unconventional but plausible alternative referent (i,e., higher referential transparency), that competes with the conventional target referent for attention during processing, and to a greater extent than a non-related distractor referent. The *context* of the lexical decision task is therefore shaped by the availability of options simultaneously shown onscreen. In other words, if the alternative shown onscreen is higher in *referential transparency*, even if unconventional (as in the café example of *ladyfinger* given in the introduction), then this will result in different distributions of gaze between target options and unconventional/lower transparency alternatives when compared to alternatives that are completely non-related (Experiment 3). Additionally, this provides evidence for *decompositional* processing of compound terms. Finally, evidence for the assumed early AoA of the lexicalized exocentric compounds is captured in choice behaviour (i.e., children should be at ceiling in their selection of the target conventional referents when they are presented alongside unrelated distractors). It also works in support of the acceptability of the unconventional alternatives presented in Experiment 1 if these referents are also chosen above chance when presented alongside a non-related distractor.

In Experiment 3, we predicted that gaze behaviour would be significantly different from Experiment 1 (baseline) due to the role of *contextual constraint* on referent selection. In other words, the presence of an unconventional referent higher in *referential transparency* (*sommerfugl;* a bird in summer) being shown alongside a conventional referent comparatively lower in *referential transparen-*

cy (*sommerfugl;* a butterfly) in Experiment 1 (baseline), would result in differential gaze distribution when compared to Experiment 3 condition one (*sommerfugl;* a butterfly vs. a bicyle) and condition two (*sommerfugl*; a bird in summer vs. a bicycle).

Sample Size

A new sample of 25 3-year-olds (N= 13 females, M_{Age} = 3 years, 5 months, 2 days) and 20 5-year-olds (N= 10 females, M_{Age} = 5 years, 4 months, 12 days) were tested. The final sample (N=41) was reached after exclusions which amounted to three 3-year-olds (n = 3 technical error, n = 1 pointed before shown peer choice, n = 1 verbalized choice was not related to task question) and one 5-year-old (n = 1 technical error).

Data Preparation and Analyses

Invalid trials amounted to 0.02% of the data from Experiment 3. We first looked at percentage choice of the conventional referent (condition one) and the unconventional referent (condition two) over the non-related distractor images (See Table 1).

EDF files containing gaze data produced by Experiment builder were converted to ASC files generating large raw sample data for each participant. These were next processed to contain data from a period of interest for each trial from the onset of the target word to the end of trial which was brought about by eventual choice of referent. The raw data were next parsed into fixation data using Gazepath in R since I-VT filters used in eye-tracker software programs are designed for parsing high quality data from adults and may not be suitable for lower quality data collected from children due to head movement. Fixation is a type of eye movement indicated by gaze velocity and dispersion thresholds, which implies moments in time during which information is encoded into working memory. The Gazepath algorithm is specifically designed to cater to both individual differences in gaze behaviour and low-quality data often present in infant and child samples for fixation detection (van Renswoude & Visser, 2017).

Upon initial visual inspection, gaze data quality in the current study was at times low due to participant movement. As such, the algorithm parsed out valid fixation data from 15 3-year-olds and 21 5-year-olds from Experiment 1 (baseline). From Experiment 3, gaze data from 19 3-year-olds and 17 5-year-olds was processed. The total fixation duration on each referent was calculated for the period of interest for each compound term. Since period of interest extended from onset of the compound word until a choice was made, which took some children longer than others, outliers were identified using the *extremes* command in Stata, resulting in values above the 1.5 interquartile range (IQR) being removed from the data set. This amounted to 39 trials (4%) in total being removed across both experiments for the total fixation durations on targets with excluded values ranging from 4680-28194ms. Data from 15 trials (2%) in total were removed for total fixation durations on alternatives with excluded upper values ranging from 4364-11706 ms. From the remaining data, trials with valid fixation data jointly present across both variables, target and alternative, resulted in a total of 994 trials on which the analysis was conducted.

We ran a regression model to examine the difference between fixation duration on target relative to the alternative/distractor across experiments and conditions using ANOVA. We initially ran models with random effects (item and participant), but found that the Likelihood ratio test was non-significant, suggesting the model was not improved by the addition of these random effects. Therefore, we opted for simpler models running post-hoc pairwise comparison tests using Tukey for multiple comparisons. We did this separately for both compound blocks; Block 1.1. conventional compounds (Model 3.1) and Block 1.2 polysemic/novel compounds (Model 3.2). We looked at potential differences across age groups and since 5-year-olds were on average quicker to select referents than 3-year-olds (see Table 4 for mean trial length values), we added average time taken to choose a referent to the model (i.e., mean trial length) to further interpret these results. Finally, we looked at differences across compound blocks to assess whether there were processing differences (measured through differential gaze distribution to targets relative to the alternatives) for conventional and novel compounds (Model 3.3).

Model 3.1 *Total fixation duration on target compared to alternative/distractor between Experiment 1 (baseline) and Experiment 3 (with two conditions) for conventional compounds by age and trial length.*

Total fixation duration on target $\text{Conventional}_{ijk} \sim$ total fixation duration on alternative_{ijk} + experimental condition_{i} + age_{j} + trial length_{k} + age_{j}*trial length_{k}

Model 3.2 *Total fixation duration on target compared to alternative/distractor between Experiment 1(baseline) and Experiment 3 (with two conditions) for polysemic and novel compounds by age and trial length.*

Total fixation duration on target $\text{Transparent}_{ijk} \sim$ total fixation duration on alternative_{ijk} + experimental condition_{i} + age_{j} + trial length_{k} + age_{j}*trial length_{k}

Model 3.3. *Total fixation duration on targets compared to alternative/distractor across experiments between compound Blocks 1.1 and 1.2*

Total fixation duration on $\text{target}_{ijkl} \sim$ total fixation duration on $\text{alternative}_{ijkl}$ + experimental condition_{i} + age_{j} + trial length_{k} + compound block_{l} + age_{j}*trial length_{k}

Table 4

Proportion of Choice of Conventional and Unconventional Referents from Block 1.1 Across All Experiments

	3-year-olds			5-year-olds		
	M	SE	[95% CI]	*M*	SE	[95% CI]
Experiment 1						
Block 1.1	3138.57	174.28	[2796.56, 3480.57]	2218.77	67.89	[2085.55, 2351.99]
Block 2.1	3437.98	199.59	[3046.31, 3829.64]	2576.70	114.53	[2351.94, 2801.45]
Experiment 3 (condition one)						
Block 1.1	3688.63	241.26	[3215.20, 4162.06]	2526.96	128.05	[2275.69, 2778.23]
Block 1.2	4207.95	334.51	[3551.53, 4864.37]	2789.41	178.97	[2438.21, 3140.62]
Experiment 3 (condition two)						
Block 1.1	4172.56	208.15	[3764.09, 4581.03]	2914.29	175.12	[2570.64, 3257.95]
Block 1.2	3278.15	169.85	[2944.85, 3611.46]	2338.10	190.36	[1964.54, 2711.65]

Results

Firstly, choice percentages showed that both 3- and 5-year-olds chose conventional referents and referents highest in referential transparency for each type of compound at ceiling (see Tables 1 and 2). Average length of time in referent choice from onset of the compound word by age across both blocks and all experiments is shown in Table 4. Across all experiments, 5-year-olds were on average faster to choose referents than 3-year-olds (β = 1490.432, *SE* = 162.43, 95% *CI* [1171.686, 1809.178], $p < 0.001$).

For the gaze behaviour analyses, Model 3.1 explained approximately 19.41% of the variance in fixation duration on the target (conventional referent) across conditions and was significant (F(6, 586) = 23.52, $p < 0.001$). The findings revealed that fixation duration on the alternative/distractor referent had a significant negative effect on fixation duration on the target (β= -.3543407, *SE* = .056724, 95% *CI* [-.4657478, -.2429336], $p < 0.001$). Significant differences were observed between experimental conditions: condition one (β = 258.2692 88.98006, Tukey 95% *CI* [49.19446, 467.344]), $p = 0.01$) and condition two (β = 357.3792, *SE* = 82.4687, Tukey 95% *CI* [163.6041, 551.1544], $p < 0.001$), which generated increased fixations on conventional referents (targets) relative to alternatives (non-related distractors) when compared to Experiment 1 (baseline). However, no significant difference was found between condition one and two for gaze distribution between targets and alternatives (*β*= 99.11003, *SE* = 93.54921, Tukey 95% *CI* [-120.7008, 318.9208], *p* = 0.54). Age-related differences were also evident, with 5-year-olds having, on average, shorter fixation durations on target referents relative to alternatives compared to 3-year-olds *(β*= -771.6484. 122.321, Tukey 95% *CI* [-1011.889, -531.4074], $p < 0.001$). Additionally, there was an interaction between length of time-to-choice and age *(β*= .2163958, *SE* = .0370724, Tukey 95% *CI* [.1435849, .2892066], $p < 0.001$) which was found to have significant positive effect on fixation duration on the target relative to the alternative meaning that the longer it took 3-year-olds to choose the more they fixated on the target relative to the alternative compared to 5-year-olds (See Figure 6).

Model 3.2 explained approximately 27.89% of the variance in total fixation duration to the target (high transparency referent), and it was significant (F(6, 394) = 25.4, $p < 0.001$). Fixation duration on the alternative/distractor referents again predicted fixation duration on target referents (*β* = -.1433309, *SE* = .0584914, 95% *CI* [-.2583251, -.0283367], *p* = 0.015), and there were significant differences between Experiment 1 (baseline) and Experiment 3 condition one (*β* = 436.6491, *SE* = 95.17736, Tukey 95% *CI* [212.7345, 660.5638], $p < 0.001$) and condition two (*β* = 338.0646, 89.71318, Tukey 95% *CI* [127.005, 549.1243], *p* = 0.001) both having increased fixation durations on target referents (i.e., those highest in referential transparency) relative to alternatives. Again, no significant differences were found between condition one and two of Experiment 3 (*β* = -98.58451, *SE* = 99.62288, Tukey 95% *CI* [-332.9578, 135.7887], *p* = 0.584). When examining age-related differences, 5-year-olds had slightly shorter fixation durations on targets relative to alternatives compared to 3-year-olds (*β* = -280.6892, *SE* = 131.3042, Tukey 95% *CI* [-538.8337, -22.54477], *p* = 0.033) and length of time- to-choice indicated increased fixation on targets relative to alternatives (*β* =.1465096, *SE* = .0216163, Tukey 95% [.1040119, .1890073], $p < 0.001$). However, the interaction between length of time-to-choice and age was non-significant (*β* = .0526684 , *SE* = .0358524, 95% *CI* [-.0178175, .1231544], *p* = .143) in the model meaning that there were no differences in fixation on target that related to both trial length and age. (See Figure 7).

Finally, Model 3.3. explained 19.9% of the variance in mean fixations on target (F(7, 986)) = 34.99, $p < .001$). Again, increased fixation durations on alternatives led to lesser fixation durations on target referent (*β* = -.2654233, *SE* = .0415769, 95% *CI* [-.3470127, -.1838338], $p < 0.001$). Both conditions of Experiment 3 were different when compared to Experiment 1 (baseline); condition one (*β* = 337.4578, *SE* = 66.42514, Tukey 95% *CI* [181.5413, 493.3744], $p < 0.001$) and condition two (*β* = 326.0369, *SE* = 61.43227, Tukey 95% *CI* [181.8399, 470.2339], $p < 0.001$) with no differences between conditions in Experiment 3 (*β* = -11.42096, *SE* = 69.52967, Tukey 95% *CI* [-174.6246, 151.7827], *p* = 0.985). Lastly, across both experiments, there were no significant differences in gaze behaviour toward target referents relative to alternatives between compound blocks (Block 1.1 and Block 1.2) (*β* = Top of Form 90.44238, *SE* = 50.81215, Tukey 95% *CI* [-9.27, 190.1548], *p* = 0.075).

Figure 6

Total Fixation Durations on Target and Alternative/Distractor Referent for Model 3.1 Aggregated Across Items with Example of "sommerfugl" (summer+bird // butterfly Across Experiments 1 and 2

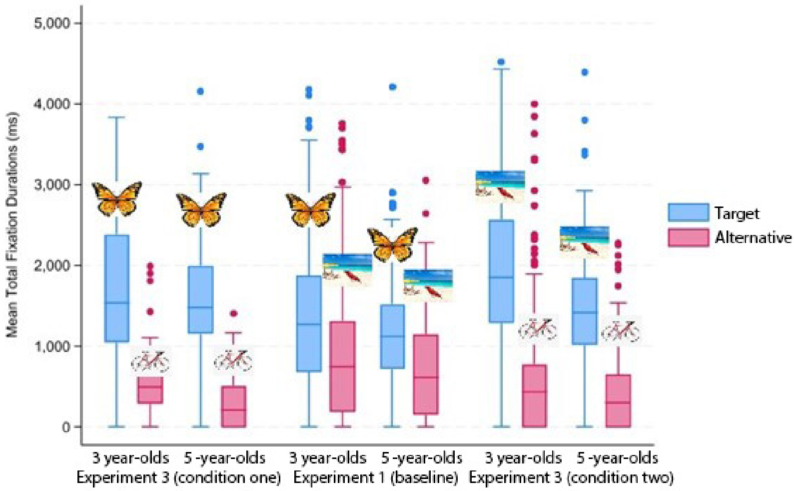
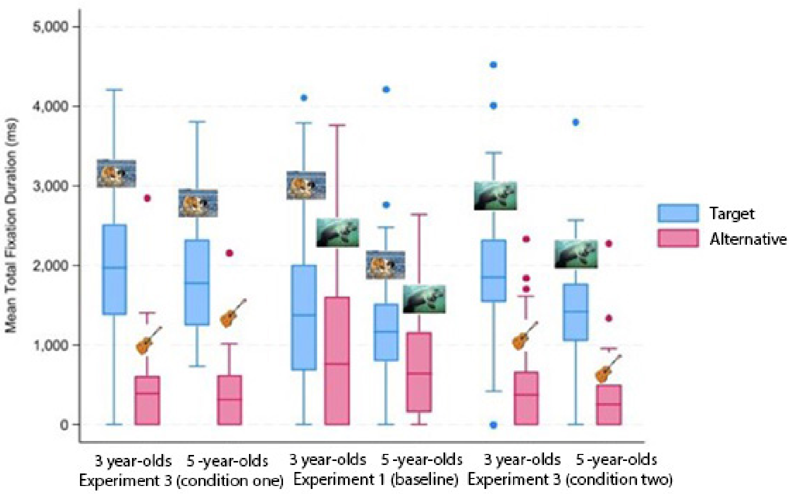


Figure 7

Total Fixation Durations on Target and Alternative/Distractor Referent for Model 3.2



Discussion

Firstly, choice behaviour was at ceiling in Experiment 3 for each of the referent-pair images used in Experiment 1 compared to the non-related distractors, working in support of assumed early AoA of compound terms and constituent nouns, as well as the acceptability of the unconventional and novel image-referents used for the compound terms in the study.

Secondly, gaze behaviour indicates that when an unconventional referent (alternative) possessing higher *referential transparency* than the conventional compound referent is shown, it significantly alters the fixation durations on the conventional referent (target). In other words, in Experiment 1 (baseline), the unconventional alternative attracts increased fixation durations in contrast to the non-related distractor of Experiment 3 (3.1). Age-related decreases in fixation on target conventional referents relative to onscreen alternatives could suggest that 5-year-olds compared to 3-year-olds show more flexibility in considering alternatives in Experiment 1 (baseline). However, 5-year-olds were also faster in referent selection overall and more time taken to choose a referent was significantly related to increased fixations on target relative to the alternatives across age groups, though these were relatively small unit changes (milliseconds in time).

The same general pattern held for polysemes and novel compounds that are higher in *referential transparency* when compared to those lower in *referential transparency*, where 5-year-olds also showed lesser fixation durations on referents highest in *referential transparency* relative to alternatives than 3-year-olds. No differences in gaze fixation distributions were found between conditions one and two in Experiment 3 and 5-year-olds again looked significantly less at targets relative to alternatives than 3-year-olds for this compound block. However, while length of time-to-choice indicated increased in fixation on targets overall, there was no interaction with age, suggesting more clearly that 5-year-olds considered alternatives slightly more than 3-year-olds in Experiment 1 (baseline) (3.2).

As such, across results from Models 3.1 and 3.2, the extra visual processing of the alternatives in Experiment 1 (baseline) compared to conditions one and two in Experiment 3, provides evidence that children *decompose* lexicalized compounds, while additionally showing that they do so in a similar way to polysemic and novel compounds. To further test this in Model 3.3, compound block (i.e., comparing processing across blocks testing conventionality and referential transparency) was added to the model and the results showed that there were no significant differences in gaze distribution suggesting that the presence or absence of conventionality did not influence how children processed the compound referents onscreen (3.3). Unlike studies with adult participants, which have suggested that priming the constituent nouns was necessary to bring about *decompositional* processing (Lamb, 1999; Libben, 2014), children seem to do so without any such priming. However, even within the adult literature, evidence for some *decomposition* of *entrenched* compound terms has been found without priming

(Bien et al., 2005), which suggests that some level of *decomposition* is always performed, possibly to handle unconventional meanings when and where they arise. This aligns with the relevance-theoretic framework of compound interpretation, whereby "local pragmatic processes of concept narrowing and broadening (ad hoc concept construction)" are thought to be necessary to resolve compound meanings (Bezuidenhout, 2019).

General Discussion

The data from Experiment 1 (baseline) document the strong role of *conventionality* in shaping young children's interpretations of compound words and their ability to rely on *referential transparency* when *conventionality* is held relatively constant (polysemes) or is absent (novel). This emphasizes the strong influence of linguistic norms on early language development showing how pragmatic development can be traced through language comprehension across early childhood, and more specifically, that with age, when a specific speaker is added to the context, young children are increasingly attuned to that individual speaker's intended meaning in referential communication.

The data show that, while children's interpretations are more generally shaped by *codification,* with age children are also beginning to appreciate the *scope restriction* of conventions becoming more flexible in their tolerance of unconventional reference across contexts (Kalish & Sabbagh, 2007). In other words, running parallel to a usage-based explanation of compound meaning acquisition that showed up in the choice data of the experiments (as a function of *conventionality*), is the generative aspect of children's acquisition that is demonstrated by the gaze behaviour data (as a function of *referential transparency*). Differences in how much children fixated on the target (/conventional) referent in Experiment 1 (baseline) where the alternative was higher in *referential transparency*, and in a context where the referential ambiguity was lower, as in conditions one (/conventional) and two (/unconventional) of Experiment 3, provided evidence for their sensitivity to the morphosemantics of compound words and their tendency to *decompose* both lexicalized and novel exocentric and endocentric compound terms in similar ways. Additionally, five-year-olds exhibited, on average, shorter fixation durations on target referents relative to alternatives suggesting they may be more open to alternatives and thus more flexible in their assessment of the referents on display. However, 5-year-olds also took less time to choose referents meaning that fixations on targets were marginally longer for 3-year-olds.

Furthermore, the data from Experiment 2 revealed that when a knowledgeable speaker is added to the context, both 3- and 5-year-olds shift their preferences in line with the speaker's choice of unconventional referents, and that 5-year-olds did this significantly more than 3-year-olds. This same tendency did not hold for polysemic or novel compounds where children's choice data was still determined by higher levels of *referential transparency*, as in Experiment 1. This finding is likely underpinned by key aspects of pragmatic development whereby a speaker's intended referent is given increased consideration at age 5 compared to age 3, in certain contexts when it diverges from the lexicalised meaning of the compound term, exemplifying growing sensitivity to *scope restriction* in conventional language with age.

Indeed, the age-related differences found across all three experiments also align with proposals such as the *lexical quality hypothesis* (Perfetti, 2007), which suggests that as children develop, their processing and comprehension of language becomes simultaneously more flexible, driven possibly most by the need to accommodate speakers' intended meanings and goals, and yet more rigid, owing to repeated exposure to specific conventional meanings (Schmidtke et al., 2018), making them more inclined to select conventional referents (Experiment 1 baseline) but also deviate from these when an individual speaker intends it (Experiment 2). Thus, young children show discriminatory ability within contexts where conventions are rigidly adhered to (*codification*) and increasingly across contexts with age, in situations where re-interpretation is permissible (*scope restriction*) (Experiment 3).

Limitations

While a subset of participants had exposure to additional languages at home, the current study recruited broadly from the local population without the intention of studying bilingual/multilingual effects. Rather, our sample simply reflects the natural linguistic diversity of children in a European capital city. However, in terms of the role that bilingualism/multilingualism could have played, the study's design explicitly anchored the tasks in a Norwegian *barnehage* (kindergarten) setting, with all experiments framed as activities where children were talking to a Norwegian-speaking teacher («Can you point to the image that best shows a [Norwegian compound]?»). This approach meant that the children reasoned and answered through their dominant L1 (Norwegian), minimizing cross-linguistic interference. In Experiment 2, the *barnehage* context also reinforced Norwegian linguistic norms, as children would associate the setting with conventional Norwegian usage making it unlikely that additional languages influenced choice. However, given at least some evidence that bilingualism/multilingualism can affect label-referent mapping strategies (Kalashnikova et al., 2019), this represents an important direction for future work that is focused on systematically comparing monolingual and bilingual/multilingual children in these tasks with sufficiently large sample sizes to do so.

Conclusion

In conclusion, the early ability children exhibit in combining nouns in meaningful ways enables them to communicate effectively without the need for more complex syntactic structures (Tomasello, 2003), while also allowing them to bridge so-called "gaps in the lexicon" (Clark, 1980). Moreover, this simple linguistic unit not only benefits young language learners when communicating but also serves as a valuable instrument for developmental scientists seeking to explore children's pragmatic development. Top of FormAs language learners, children need to acquire conventional and often semantically opaque meanings for compound words that they encounter in their social worlds, but they also need to acquire the relational knowledge to decipher new compound terms (Jackendoff, 2010). This study shows that both processes are connected, and their interplay is what points to the key pragmatic developments involved where, with age, the intentions of individual speakers are considered to a greater extent allowing for multiple and more flexible meanings to be integrated, while also facilitating quick referential access to well-established forms.

Authorship Details

The sole author was responsible for all aspects of the work.

Conflict of Interest Disclosure

The author reports no conflicts of interest.

Funding

This study was funded by the European Research Council Horizon 2020 grant no. 853211 (https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/projects-details/31045243/853211/H2020).

Research Ethics Statement

The ethics review for the research involving children was conducted by the Sikt - Norwegian Agency for Shared Services in Education and Research (study approval number: 596365). Ethical approval was granted to ensure the study minimized risks and protected the welfare of participants. Informed consent was obtained from parents or guardians, who were fully informed about the study's purpose, procedures, risks, and benefits. Participants were informed that their involvement was voluntary, and they could withdraw at any time without consequences. Researchers prioritized the children's safety, confidentiality, and well-being throughout the study.

Data Availability Statement

Study materials associated with this study are shared in a public, open-access repository: https://osf.io/4thjc/

References

Benczes, R. (2006). *Creative compounding in English*. John Benjamins Publishing Company. https://benjamins.com/catalog/hcp.19

Bezuidenhout, A. (2019). Noun-noun compounds from the perspective of Relevance Theory. In B. Clark, K. Scott, & R. Carston (Eds.), *Relevance, pragmatics and interpretation* (pp. 174–186). Cambridge University Press. https://doi.org/10.1017/9781108290593.016

Bien, H., Levelt, W. J. M., & Baayen, R. H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences*, *102*(49), 17876–17881. https://doi.org/10.1073/pnas.0508431102

Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2021). How young children integrate information sources to infer the meaning of words. *Nature Human Behaviour*, *5*(8), Article 8. https://doi.org/10.1038/s41562-021-01145-1

Clark, E. V. (1980). *Lexical innovations: How children learn to create new words.* Papers and Reports on Child Language Development, Number 18. https://eric.ed.gov/?id=ED192592

Clark, E. V. (1995). Language acquisition: The lexicon and syntax. In J. L. Miller & P. D Elmas (Eds.), *Speech, language, and communication* (pp. 303–337). Academic Press. https://doi.org/10.1016/B978-012497770-9/50011-X

Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Compound nouns and category structure in young children. *Child Development*, *56*(1), 84–94. https://doi.org/10.2307/1130176

Eiesland, E. A. (2016). *The semantics of Norwegian noun–noun compounds: A corpus-based study* [Doctoral dissertation, University of Oslo].

Eik, R. (2019). *The morphosyntax of compounding in Norwegian* [Doctoral dissertation, Norwegian University of Science and Technology]. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617105

Fares, M. (2019). *Noun–noun compound analysis: A holistic perspective*. https://www.semanticscholar.org/paper/Noun%E2%80%93Noun-Compound-Analysis%3A-A-Holistic-Perspective-Fares/fe3b766ac697ad3b11b3bedbc09b0c8883f313c1

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829. https://doi.org/10.1016/j.tics.2016.08.005

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences*, *14*, 497–505.

Jackendoff, R. (2010). The ecology of English noun–noun compounds. In R. Jackendoff, *Meaning and the lexicon: The parallel architecture 1972–2010* (pp. 413–451). MIT Press.

Kalashnikova, M., Mattock, K., & Monaghan, P. (2016). Flexible use of mutual exclusivity in word learning. *Language Learning and Development*, *12*(1), 79–91. https://doi.org/10.1080/15475441.2015.1023443

Kalashnikova, M., Oliveri, A., & Mattock, K. (2019). Acceptance of lexical overlap by monolingual and bilingual toddlers. *International Journal of Bilingualism*, *23*(6), 1517–1530. https://doi.org/10.1177/1367006918808041

Kalish, C. W., & Sabbagh, M. A. (2007). Conventionality and cognitive development: Learning to think the right way. *New Directions for Child and Adolescent Development*, *2007*(115), 1–9. https://doi.org/10.1002/cd.178

Kallen, R. W., Macpherson, M. C., Miles, L. K., & Richardson, M. J. (2022). Symmetries of social performance-environment systems. In K. Bicknell & J. Sutton (Eds.), *Collaborative embodied performance* (pp. 153–170). Bloomsbury Academic. https://doi.org/10.5040/9781350197725.ch-008

Köymen, B., Lieven, E., Engemann, D. A., Rakoczy, H., Warneken, F., & Tomasello, M. (2014). Children's norm enforcement in their interactions with peers. *Child Development*, *85*(3), 1108–1122. https://doi.org/10.1111/cdev.12178

Lamb, S. M. (1999). *Pathways of the brain*. John Benjamins Publishing Company. https://benjamins.com/catalog/cilt.170

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites*. Stanford University Press.

Li, S., Carlson, R., & Potts, C. (2022). *Systematicity in GPT-3's interpretation of novel English noun compounds* (arXiv:2210.09492). *arXiv*. https://doi.org/10.48550/arXiv.2210.09492

Libben, G. (2014). The nature of compounds: A psychocentric perspective. *Cognitive Neuropsychology*, *31*(1–2), 8–25. https://doi.org/10.1080/02643294.2013.874994

Lyster, S.-A., Horn, E., & Rygvold, A.-L. (2010). Vocabulary and vocabulary development of Norwegian children and youth. Results from a try out of British Picture Vocabulary Scale II, Second edition (BPVT II). *Spesialpedagogikk*, *74*, 35–43.

Mellenius, I. (1996). Children's comprehension of Swedish nominal compounds. In C. Johnson & J. H. V. Gilbert (Eds.), *Children's language* (Vol. 9, pp. 167–182). Lawrence Erlbaum.

Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science*, *21*(2), e12556. https://doi.org/10.1111/desc.12556

Nicoladis, E., & Krott, A. (2007). Word family size and French-speaking children's segmentation of existing compounds. *Language Learning*, *57*(2), 201–228. https://doi.org/10.1111/j.1467-9922.2007.00407.x

Ordbøkene.no – Bokmålsordboka og Nynorskordboka [Bokmål dictionary and Nynorsk dictionary]. *(2022, January 27). Ordbøkene.no. https://ordbokene.no*

Perfetti, C. (2007). Reading ability: lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Prendergast, C. (2024). Losing the thread: How three- and five-year-olds predict the outcome of a story when non-literal language is used to update events. *Cogent Psychology*, *11*(1), 2431422. https://doi.org/10.1080/23311908.2024.2431422

Ronfard, S., & Lane, J. D. (2018). Preschoolers continually adjust their epistemic trust based on an informant's ongoing accuracy. *Child Development*, *89*(2), 414–429. https://doi.org/10.1111/cdev.12720

Schmidtke, D., Gagné, C. L., Kuperman, V., & Spalding, T. L. (2018). Language experience shapes relational knowledge of compound words. *Psychonomic Bulletin & Review*, *25*(4), 1468–1487. https://doi.org/10.3758/s13423-018-1478-x

Sia, M. Y., & Mayor, J. (2022). Improvements of statistical learning skills allow older children to go beyond single-hypothesis testing when learning words. *Journal of Child Language*, *49*(6), 1268–1280. https://doi.org/10.1017/S0305000921000532

Snape, S., & Krott, A. (2022). The challenge of relational referents in early word extensions: Evidence from noun-noun compounds. *Journal of Child Language*, *49*(1), 131–163. https://doi.org/10.1017/S0305000920000793

StataCorp LLC. (2023). *Stata Statistical Software: Release 15* [Computer software]. StataCorp LLC.

Telzer, E. H., Flannery, J., Humphreys, K. L., Goff, B., Gabard-Durman, L., Gee, D. G., & Tottenham, N. (2015). "The cooties effect": amygdala reactivity to opposite-versus same-sex faces declines from childhood to adolescence. *Journal of Cognitive Neuroscience*, *27*(9), 1685–1696. https://doi.org/10.1162/jocn_a_00813

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Tomasello, M. (2009). The usage-based theory of language acquisition. In E. L. Bavin (Ed.), *The Cambridge Handbook of Child Language* (pp. 69–88). Cambridge University Press. https://www.cambridge.org/core/books/cambridge-handbook-of-child-language/usagebased-theory-of-language-acquisition/30A7D7DA467B600B632871678783A596

van den Herik, J. C. (2021). Rules as resources: an ecological-enactive perspective on linguistic normativity. *Phenomenology and the Cognitive Sciences*, *20*(1), 93–116. https://doi.org/10.1007/s11097-020-09676-0

van Renswoude, D. R., & Visser, I. (2017). *gazepath: Parse eye-tracking data into fixations* (Version 1.2) [Computer software]. https://CRAN.R-project.org/package=gazepath

Wellman, H. M. (2017). The development of theory of mind: historical reflections. *Child Development Perspectives*, *11*(3), 207–214. https://doi.org/10.1111/cdep.12236

Winner, E., Levy, J., Kaplan, J., & Rosenblatt, E. (1988). Children's understanding of nonliteral language. *Journal of Aesthetic Education*, *22*(1), 51–63. https://doi.org/10.2307/3332964